

Essays in Labour Economics

by

Neil Douglas Lloyd

B.Soc.Sc. (Hons), University of Cape Town, 2012

M.A., University of British Columbia, 2014

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate and Postdoctoral Studies

(Economics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

July 2021

© Neil Douglas Lloyd 2021

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation titled:

Essays in Labour Economics

submitted by Neil Douglas Lloyd in partial fulfillment of the requirements for

the degree of Doctor of Philosophy

in Economics

Examining Committee:

Professor Thomas Lemieux, Vancouver School of Economics, UBC

Co-supervisor

Associate Professor Florian Hoffmann, Vancouver School of Economics, UBC

Co-supervisor

Professor Nicole Fortin, Vancouver School of Economics, UBC

Supervisory Committee Member

Professor David Green, Vancouver School of Economics, UBC

University Examiner

Professor Sylvia Fuller, Department of Sociology, UBC

University Examiner

Additional Supervisory Committee Members:

Professor Kevin Milligan, Vancouver School of Economics, UBC

Supervisory Committee Member

Abstract

Chapter 2 demonstrates how individual income tax structures incentivize a more coordinated labour supply response to childbirth within married households: a joint selection out of the wage-paying sector and into self-employment. In a parallel analysis of longitudinal administrative and survey data from Canada, I show that the birth of a first child is associated with an increase in both maternal and paternal self-employment in married households; explained largely by an increase in co-employment. I develop a novel simulated instrument research design, which exploits exogenous tax variation, to show that this strategic re-organization of the household is partly incentivized by income splitting tax savings. Finding a reduced form elasticity of 0.5, these savings can account for half the increase in co-employment after childbirth. Beyond tax avoidance, this paper presents income splitting as a subsidy to the creation of flexible, tax-optimizing family firms that provide stable, long-run employment to households.

Chapter 3 provides the first causal evidence on the impact of incorporation on the labour supply and hiring practices of self-employed professionals. It exploits staggered reforms across professions in each province to permit the registration of professional corporations in Canada. I found no evidence of a labour supply response to the significant tax implications of incorporation. However, for female professionals, incorporation increases the likelihood of hiring at least one employee. This result is consistent with the cash flow benefits of retained corporate earnings that enable to business owners to ensure against uncertain revenue.

Chapter 4 extends the DiNardo, Fortin, and Lemieux (1996) study of the links between labour market institutions and wage inequality in the United States and updates the analysis to the 1979 to 2017 period. A notable extension quantifies the magnitude and distributional impact of spillover effects linked to minimum wages and the threat effects of unionization. A distribution regression framework is used to estimate both types of spillover effects. Accounting for spillover effects doubles the contribution of de-unionization to the increase in male wage inequality. It raises the explanatory power of declining minimum wages to two-thirds of the increase in inequality at the bottom end of the female wage distribution.

Lay Summary

This thesis consists of three chapters in Labour Economics. The first two chapters concern self-employed workers, while the final chapter addresses labour market institutions. Chapter 2 provides the first detailed account of selection into self-employment with the birth of a first child using Canadian longitudinal administrative and survey data. With family formation, there is a coordinated selection out of the wage-paying sector into self-employment by both mothers and fathers that is shown to be, in part, a response to the individual nature of Canada's tax structure. Chapter 3 complements Chapter 2 by providing the first causal evidence of how incorporation status affects the labour market behaviour of self-employed workers. Chapter 4 extends the debate on the role of labour market institutions in the rise of wage inequality in the US by accounting for the spillover effects of collective bargaining and minimum wages.

Preface

Chapters 2 and 3 of the thesis are pieces of original, unpublished, and independent work. Chapter 4 is joint work with professors Nicole Fortin (UBC and NBER) and Thomas Lemieux (UBC and NBER), a version of which is now published in the *Journal of Labor Economics* under the title “Labor market institutions and the distribution of wages: The role of spillover effects”. As leaders of this project, professors Lemieux and Fortin invited me to assist them at the outset of the research and soon after invited me to be a co-author. I completed the majority of the initial analysis which focused on the replication of an existing study using more recent data as well as a reduced form analysis of a set of recent legal reforms. Professor Lemieux provided the methodological contribution of the paper and I assisted him in the estimation of the model and preparation of results. Professor Fortin and I worked together to prepare figures and tables for presentations and publication. All members of the team contributed towards the writing and editing of the manuscript. Finally, this version of the paper includes results that were excluded from the final published version (see Section 4.3). I chose to include this additional analysis here as this work represents an area of ongoing research for which I am the lead analyst.

Contents

Abstract	iii
Lay Summary	iv
Preface	v
Contents	vi
List of Tables	ix
List of Figures	xi
Acknowledgements	xiv
Dedication	xv
1 Introduction	1
2 Strategic self-employment and family formation	7
2.1 Introduction	7
2.2 Self-employment and childbirth: flexibility and income splitting	13
2.3 Data & Measurement	16
2.3.1 Identifying childbirth in the LAD	16
2.3.2 Identifying self-employment in the LAD	18
2.4 Self-employment and family formation: event-study analysis	21
2.4.1 The self-employment ‘parent penalty’	24
2.4.2 On the <i>real</i> nature of co-employment after childbirth	27
2.4.3 Relative income and tax savings from self-employment	29
2.5 Tax incentives for co-employment	31
2.5.1 Identification	32
2.5.2 Testing income-splitting mechanism	39
2.5.3 Results	40
2.6 Conclusion	42

2.7	Tables and Figures	44
3	Does incorporation matter? Causal evidence from professional corporation re-	
	forms in Canada	66
3.1	Introduction	66
3.2	Why professional businesses incorporate	68
3.3	History of institutional reforms	73
3.4	Research design	77
3.5	Data	80
3.6	Rising incorporation among professionals	81
3.7	Labour supply and hiring among self-employed professionals	84
3.8	Conclusion	87
3.9	Tables and Figures	89
4	Labour market institutions and the distribution of wages: The role of spillover	
	effects	107
4.1	Introduction	107
4.2	Estimating spillover effects of the minimum wage	110
4.2.1	Distribution regressions	111
4.2.2	Empirical implementation	114
4.2.3	Identification	117
4.3	Union threat effects	118
4.3.1	Event study of the introduction of right-to-work laws	118
4.3.2	Measuring threat effects in a distributional context using the unioniz-	
	ation rates	120
4.4	Data and estimation results	123
4.4.1	Data	123
4.4.2	Minimum wage effects	124
4.4.3	Are spillover effects real or just due to measurement error?	128
4.4.4	Distribution regression estimates of the effect of unionization	130
4.5	Decomposition results	133
4.6	Conclusion	136
4.7	Tables and Figures	138
5	Conclusion	156
	Bibliography	158

Appendices

A	Chapter 2 Appendix	172
A.1	Finite horizon, unitary household model	172
A.2	Identifying childbirth in the LAD (continued)	180
A.3	Identifying self-employment in the LAD (continued)	182
A.4	Event-study-design coefficients	187
A.5	Accounting for Regional Differences in Co-employment	193
A.6	Characteristics of co-employment mothers	206
A.7	Relative earnings	209
A.8	Income-splitting: Regulatory setting and tax reforms	211
B	Chapter 3 Appendix	215
B.1	Proof of IV decomposition	215
B.2	IV decomposition with exogenous covariates	218
B.3	Alternative specifications	220
C	Chapter 4 Appendix	226
C.1	Right-to-Work results using NLRB data	226
C.2	Issues in the estimation of the direct and spillover effects of the minimum wage for distribution regressions	227
C.3	An illustrative example of the identification of minimum wage effects	231
C.4	RIF-regression estimates of the impact of unionization	233
C.5	Understanding the hump shaped effect of unionization on the wage distribution	236

List of Tables

2.1	Changes at the intensive margin of labour supply among employed women	56
2.2	Changes in taxable income per hour and working from home status among employed women	57
2.3	Reduced form elasticities: paternal and maternal self-employment	63
2.4	Family formation gap in male self-employment: LFS & CPS, 1988-2015	64
2.5	Relative probability of maternal self-employment conditional spousal self-employment	65
3.1	Years of regulation change	75
3.2	First stage by occupation based on event-study research design	96
3.3	Labour supply and hiring: OLS and IV results based on event-study first stage	99
3.4	Men: Labour supply and hiring: OLS and IV results based on event-study first stage	100
3.5	Women: Labour supply and hiring: OLS and IV results based on event-study first stage	101
3.6	Labour supply and hiring: gender differences in OLS before and after treatment spells	103
3.7	First stage by occupation based on event-study research design with treatment unit FEs	104
3.8	Labour supply and hiring: IV results based on event-study first stage with treatment unit FEs	105
4.1	Inequality measures and descriptive statistics	148
4.2	Difference-in-difference estimates of Right-to-Work laws	149
4.3	Minimum wage effects estimated from distribution regression models	150
4.4	Estimates of minimum wage impacts along the wage distribution	151
4.5	OLS estimates of the union threat effect on log wages with alternative definitions of the unionization rate	152
4.6	Unionization rate effects estimated from distribution regression models . .	153
4.7	Decomposition results: Men	154

List of Tables

4.8	Decomposition results: Women	155
A.1	Summary statistics and gender norms, by gender: WVS 1990, 2000, and 2006	201
A.2	Summary statistics and gender norms, by region: WVS 1990, 2000, and 2006	202
A.3	Summary statistics and gender norms, by employment: WVS 1990, 2000, and 2006	203
A.4	Reason for self-employment among prime-aged adults, by gender: GSS 2016	204
A.5	Reason for self-employment among prime-aged adults, by region: GSS 2016	205
A.6	Labour supply characteristics of employed, married mothers in two earner households	207
A.7	Labour market characteristics of self-employed and co-employed mothers, relative to paid employees in two earner households	208
B.1	First stage: generalized difference-in-difference, by occupation	222
B.2	Labour supply and hiring: IV results based on DiD first stage	223
B.3	First stage: break-in-trend, by occupation	224
B.4	Labour supply and hiring: IV results based on break-in-trend first stage . .	225

List of Figures

2.1	Participation parent penalty: class of employment	44
2.2	Trend in Canadian self-employment by incorporation and employer status.	45
2.3	Participation parent penalty: overall employment	46
2.4	Participation parent penalty: maternal self-employment and joint self-employment	47
2.5	Participation parent penalty: class of employment, by class of spouse	48
2.6	Participation parent penalty: overall employment, by class of spouse	49
2.7	Age profile: self-employment in Canada and the US	50
2.8	Comparison of event-time coefficients: participation in self-employment	51
2.9	Comparison of event-time coefficients: joint self-employment	52
2.10	Comparison of event-time coefficients: participation in self-employment by self-employment status of spouse	53
2.11	Share of two-earner couples employed in the same industry by age group and spouse's class of employment.	54
2.12	Comparison of event-time coefficients: log of annual hours worked and probability of working from home, by spouse's class.	55
2.13	Relative employment earnings of married women, by class of spouse	58
2.14	Changes in relative female earnings and average tax rate of the couple.	59
2.15	Trend in after-tax earnings for self-employed married households with a total income of \$80,000 (2016 CAD).	60
2.16	Simulated income-splitting tax savings (Δ) by birth cohort.	61
2.17	Comparison of event-time coefficients: participation in self-employment	62
3.1	Trend: Incorporation rate of self-employed professionals, by gender.	89
3.2	Female share of non-education professional employees.	90
3.3	Trend: Self-employment rate of professionals, by gender.	91
3.4	Event-study coefficients from model examining selection into self-employment.	92
3.5	Canadian corporate tax rate, including small business tax credit.	93
3.6	Trend: Incorporation by professional group and provincial treatment group.	94
3.7	Event-study coefficients from first stage: incorporation.	95
3.8	Trend: Outcome variables by provincial treatment group.	97

List of Figures

3.9	Trend: Average hours worked by gender: paid employees compared self-employed professionals.	98
3.10	Trend: Outcome variables by professional group and provincial treatment group.	102
3.11	Trend: Hiring status by professional group and provincial treatment group.	106
4.1	Real value (\$2017) of the minimum wage (MW) and fraction of workers in states with a higher minimum	138
4.2	Map of Right-to-Work adoption.	139
4.3	Coefficients from event-study of Right-to-Work laws.	140
4.4	Trends in unionization rates across Census Bureau division, and low- and high-unionization industries: Men and women combined	141
4.5	Change in union coverage and net certification over time.	142
4.6	Marginal effects of minimum wages	143
4.7	Marginal effects of a 1% increase in the unionization rate	144
4.8	Counterfactual densities: 1979-1988	145
4.9	Counterfactual densities: 1988-2000	146
4.10	Counterfactual densities: 2000-2017	147
A.1	Simulated <i>real</i> self-employment rate by event-time of childbirth	178
A.2	Simulated <i>real</i> co-employment by event-time of childbirth	179
A.3	Comparison of cohabitation rates across the LAD, LFS and SLID/CIS	180
A.4	Comparison of the self-employment rate in the LAD, LFS and SLID/CIS	183
A.5	Incorporation rate of self-employed workers in the LAD and LFS	184
A.6	Comparison of self-employment definitions across LAD and LFS for men and women aged 25-44.	185
A.7	Event-study-design coefficient: class of employment	188
A.8	Event-study-design coefficient: overall employment	189
A.9	Event-study-design coefficient: maternal self-employment and joint self-employment	190
A.10	Event-study-design coefficient: class of employment, by class of spouse	191
A.11	Event-study-design coefficient: overall employment, by class of spouse	192
A.12	Participation parent penalty: maternal class of employment by geographical region	197
A.13	Participation parent penalty: paternal class of employment by geographical region	198
A.14	Participation parent penalty: maternal self-employment and joint self-employment by geographical region	199

List of Figures

A.15 Region-specific child fixed effect.	200
A.16 Relative employment earnings (excl. dividends) of married women, ages 25-54, with and without children; SLID/CIS 1996-2016. Markers represent fraction of couples in binds of 2.5%.	210
A.17 Federal income tax brackets and Canadian Pension Plan contributions for 1993 and 2003; assuming self-employed filer.	214
C.1 Average of NLRB election data by event-time: Indiana, Michigan, and Wis- consin	226
C.2 Histogram of nominal wage bins	228
C.3 Estimates of minimum wages impact: 1979-88	229
C.4 Estimates of minimum wages impact: 1979-88	230
C.5 Illustrative example of the identification of minimum wages effects	232
C.6 OLS and RIF regression model: Spillover effects based on unionization rates	235
C.7 Understanding the hump-shaped effects of unionization	238

Acknowledgements

I would like to express my deepest gratitude to the members of my committee: Thomas Lemieux, Florian Hoffmann, Nicole Fortin, and Kevin Milligan. You encouraged me to pursue my own ideas, you were always generous with your time, and you were thoughtful with your advice. I owe a special thank you to Nicole and Thomas for the invitation to work with them on the joint work presented here. I learnt so much working alongside you and will forever be in your debt for this incredible opportunity. You have taught me to be a researcher: to quickly identify what matters, to wrestle with a problem, and to always hold the bigger story in mind.

I would like to thank the Social Sciences and Humanities Research Council of Canada for funding my research through a SSHRC Doctoral Fellowship from September 2018 to December 2020. I would like to acknowledge the unexpected, yet greatly appreciated, additional funding provided for the period of August-December 2020 in the wake of Covid-19.

I also want to thank David Green and Vadim Marmer for their personal leadership and encouragement in both my research and teaching at UBC. David, I will always remember your kindness in the wake of my brother's passing. Vadim, thank you for encouraging me into the programme and for encouraging me to stay. Thank you for the opportunities to teach and for demonstrating to me first-hand what it means to be an educator.

I owe a special thank you to much my rich community of friends and family who have supported and encouraged me all these years. Those in Vancouver, thank you for being a source of light and joy throughout this PhD journey. For the mountains we hiked and the conversations we shared. Those who I left behind to pursue this dream, thank you for your unending support and sacrifice. I hope that this new chapter sees our paths crossing more often.

Finally, I owe a special thank you to my wife, Stacey. Your commitment to me throughout the PhD has been unconditional. Through the doubt and insecurity, you have stood by me. You have celebrated each of my successes and remained steadfast throughout my setbacks. Thank you for the sacrifice that you have made by moving your own life and career to Vancouver with me. Your friendship has been the one constant these past 6 years and for that I am truly grateful.

Dedication

I dedicate this thesis to Robert Ernest Lloyd (June 1982 - November 2015). You were my big brother, then my idol, and, finally, my best friend. Thank you for dreaming this dream with me at a time when dreams were so fragile.

“Being his real brother I could feel I live in his shadows, but I never have and I do not now. I live in his glow.”

Michael Morpurgo, *Private Peaceful*

Chapter 1

Introduction

Inequality, be it within or between group, is a central focus of economists in the study of the labour market. A major debate concerning the rise of wage inequality in the labour market concerns the potential role of labour market institutions such as minimum wages and collective bargaining (Card et al., 2004; DiNardo et al., 1996).¹ While the discussion on the persistence of gender inequality has focused its attention on the labour market cost of parenthood (Bertrand et al., 2010; Kleven et al., 2019b). In all these debates, a common theme is the almost exclusive attention given to the wage paying labour market. This thesis makes a concerted effort to draw the attention to the self-employed labour market, while contributing further the debate on the role of labour market institutions in the wage paying sector.

Chapter 2 introduces selection between the self-employed and paid labour market as a central element of labour market adjustments to parenthood. The chapter identifies a coordinated and highly persistent path of joint selection from the wage-paying sector into self-employment by married parents in Canada. This decision is then causally linked to the individual nature of Canada's income tax structure. Chapter 2 is agnostic with regards to the legal structure of self-employed businesses, which take on either an unincorporated or incorporated form. For this reason, Chapter 3 explores the importance of incorporation for the labour supply of self-employed professionals, speaking directly to the debate on how these legal institutions affect inequality through their preferential tax treatment. Finally, Chapter 4 provides an important addition to the debate on the role of labour market institutions in the growth of wage inequality. This joint work, with Nicole Fortin and Thomas Lemieux, extends the existing analyses by developing a methodology to account for the spillover effects of both minimum wages and collective bargaining in the labour market.

The attention given to self-employed workers has *ebbed and flowed* over the last century. The dominant position prior to the 1970s is best summarized by Phillips (1962) as a “shrinking world within a growing economy”. Self-employment was perceived to be an increasingly irrelevant, pre-industrial mode of employment and expected to become a

¹The view that weakening labour market institutions explain the rise in wage inequality - in particular, in the US - competes with other narratives, including the role of skill-biased technical change (see, e.g., Acemoglu and Autor, 2011)

“refuge for older workers, the physically and mentally handicapped, and others with low personal productivity” (Aronson, 1991).² However, this perspective began to change in the late 1980s, as non-farming self-employment rose steadily throughout the 1970s and 80s across both the US and other industrial economies (see, e.g., Blau, 1987; Bögenhold and Staber, 1991; Hakim, 1988).³ Its re-emergence across industrial countries – including the US and Canada - during this period sparked two separate streams of literature. The first focuses on self-employment as a proxy for entrepreneurship, treating measured self-employment (commonly coded by class of employment in labour market survey data) as a signal of Schumpeterian risk-taking (Schumpeter, 1942).⁴ While this approach has been criticized, it is arguably the dominant lens through which self-employment is studied (see discussions in Bjuggren et al., 2010; Faggio and Silva, 2012).

The second strand of literature treats self-employed workers within a broader class of alternative work arrangements; focusing more on the contractual nature of self-employment in contrast to the dominant employer-employee relationship (Abraham, 1988; Barker and Christensen, 1998; Katz and Krueger, 2016). More recently, the alternative work arrangements literature has been revisited by Abraham et al. (2018) in relation to the emergence of gig-economy and platform-based work. From this perspective, selection into self-employment is perceived as a labour supply decision characterized primarily by a change in the relationship between worker and firm. While a paid employee typically has a contract stipulating their wage/salary in return for hours of work, a self-employed worker’s pay is more closely tied to their own individual output; be that because they are the owners of a business or because they operate as a freelance worker. This is a highly individual mode of employment, that can be best contrasted by the collective nature of unionized, paid employment. In return for the certainty and security of an employment contract, self-employed workers gain self-determination. It is this flexibility, and in particular the flexibility to determine hours and location of work, that appeals to many self-employed workers.⁵ It is also one of the reasons that self-employment increases steadily over the lifecycle: growing faster during years of family formation and during the onset of retirement.⁶

²This perspective is also seen in Oxenfeldt (1943), who describes self-employment as a viable path for the unemployed: “Most unemployed persons would probably be willing to take great risks for a chance to secure a small income and to erase the stigma associated with unemployment. ... It is not irrational-or unusual-for an individual who is unemployed to set up a business even if the chances of succeeding are small”.

³See also Kuhn and Schuetze (2001), Hipple (2004, 2010), Karoly and Zissimopoulos (2004), Lin et al. (2000), and Manser and Picot (1999a,b) for studies documenting later trends in self-employment.

⁴This practice is first evident in the works of Storey (1991) and Acs et al. (1994).

⁵A number of papers have attempted to measure the non-pecuniary benefits of self-employment using self-reported life satisfaction data (see, e.g., Blanchflower, 2000; Blanchflower and Oswald, 1998; Hurst and Pugsley, 2010; van der Zwan et al., 2018). Though, Hanglberger and Merz (2015) argue that the existing literature overestimates the satisfaction gains of self-employment, as they decline quickly over time with adaptation.

⁶Humphries (2019) adopts a novel clustering approach to group various paths of self-employment over the

In Canada, self-employment rates among prime-aged men rose sharply through the early 1990s, peaking in the late 1990s, and then declining during the early 2000s (Manser and Picot, 1999a,b). This wave of self-employment is best characterized by an increase in own account (without employees) unincorporated workers; although incorporation status is partly endogenous to regulations at the time (as discussed in Chapter 3). In contrast, US male self-employment rates consistently declined from the early 1990s, a trend largely explained by a stagnation in the growth of unincorporated self-employed workers (Hipple, 2010; Karoly and Zissimopoulos, 2004). It is this contrasting trend, between the US and Canada, that is central to the contribution of Chapter 2. A fundamental institutional difference between these two neighbouring countries is the way that household income is taxed. While Canada has elected to tax household income at the individual level, the US taxes households based on their combined income. I argue that this small, but important institutional detail, is the reason that Canadian (male) self-employment rates are higher and have followed their own trend since the early 1990s.

Chapter 2, *Strategic Self-employment and Family Formation*, centers the discussion on self-employment around the birth of a first child using longitudinal data from Canada. In this way, the paper contributes to a growing body of literature on the parent-penalty and its impact on gender inequality in the labour market (Angelov et al., 2016; Bertrand et al., 2010; Kleven et al., 2019b; Wilde et al., 2010; Zhang, 2010). While maternal paid employment declines with childbirth, childbirth is shown to be associated with a persistent increase in female self-employment (see also Jeon and Ostrovsky, 2019). In fact, the birth of a first child explains 80% of the lifecycle increase in female self-employment. For fathers, initial childbirth is shown to be associated with a highly persistent switching out the paid labour market and into self-employment, explaining a third of the lifecycle increase in self-employment. This is a novel finding that contrasts the absence of any employment loss associated with childbirth (as highlighted by other studies, including Kleven et al., 2019b). Thus, this infra-marginal response – selection into self-employment – is an important margin at which individuals adjust their labour supply with family formation.

The chapter then shows that a key characteristic of this decision is the increase in joint self-employment: households characterized by two self-employed parents. Multiple data sources are then used to show that this joint self-employment response is best described

lifecycle using Swedish administrative data. A notable finding is the presence of highly stable and long-lasting self-employed careers, which contrasts the association between self-employment and risk. These same stable paths are evident in Chapter 2. See also Zissimopoulos and Karoly (2007) for a discussion of selection into self-employment at older ages. Ahn (2010) discusses the relationship between risk tolerance and selection into self-employment early on the lifecycle, while Jeon and Ostrovsky (2019) provide a good introduction into the relationship between self-employment and family formation. Although this literature is almost exclusively focused on female self-employment as an alternative to part-time work. In this regard, Chapter 2 distinguishes itself from the existing literature as it focuses on maternal, paternal and joint self-employment responses to childbirth.

by co-employment: employment in the same closely held firm. One of the key advantages to self-employment and co-employment, at this stage of the lifecycle, is the flexibility it provides with regards to childcare. It is not surprising then that the increase in co-employment is associated with an increase in working from home by mothers. However, there are tax benefits too. Related to aforementioned tax differences between Canada and the US, Canadian households with a more equal income distribution pay lower taxes than those with the same total income, but more unequal distribution. Households therefore have an incentive to equalize their income where they can, but wage/salary income is reported by a third party to the tax authority and cannot be 'split'. Given the self-reported nature of self-employment income, self-employed households can more easily take advantage of these income smoothing strategies. This chapter shows that with childbirth, self-employed households disproportionately benefit from tax savings associated with a more equal income distribution.

Such tax avoidance behaviour by self-employed households is well documented in the literature, where a number of studies document the distortions to female labour supply and reported income in households where the male spouse is self-employed (Harju and Matikka, 2016; LaLumia, 2008; Schuetze, 2006; Stephens and Ward-Batts, 2004; Zinovyeva and Tverdostup, 2018). However, this paper is the first to show that both primary and secondary earner labour supply is endogenous to these savings. That is, the pattern of selection into self-employment by both fathers and mothers with childbirth is partly explained by the tax savings generated through income splitting. Put differently, without these savings the male self-employment response to childbirth in Canada would be approximately half of what it is. This conclusion is further supported by a joint analysis of paternal self-employment in Canada and the US using comparable survey data. Returning to the diverging US and Canadian self-employment trends, the wave of male self-employment in Canada, beginning in the 1990s, follows the evolution of these tax savings resulting from changes to the income tax structure. This paper therefore provides an explanation for the diverging trends in self-employment across the US and Canada, since the late 1980s (Manser and Picot, 1999a,b).

Chapter 2's analysis of self-employment does not take into account important tax differences between incorporated and unincorporated businesses. Unfortunately, the absence of a suitable instrument for this endogenous decision prevented me from doing so. It is for this reason that Chapter 3 provides an important complementary analysis. Chapter 3, *Does Incorporation Status Matter? Causal Evidence from Professional Corporation Reforms in Canada*, takes advantage of staggered legislative reforms governing regulated professionals across Canadian provinces to assess how allowing self-employed professionals to incorporate affects their own labour supply and decision to hire employees. Existing analyses of incorporated businesses in Canada, using administrative tax records, suggest

that corporate structures increase income inequality through preferential tax treatment (Wolfson and Legree, 2015; Wolfson et al., 2016). However, for data limitation reasons, these studies cannot account for any positive real labour market affects from incorporation.

The chapter documents a steady decline in professional hours worked since the early 1990s, as well as a dramatic decline in the share of self-employed professionals with employees (consistent with other studies, e.g., Crossley et al., 2009). The reforms present themselves as a unique opportunity to explore the relationship between corporate status and these labour market outcomes, as standard tax-based instruments for incorporation do not fulfill the exclusionary restriction required for to answer this question.⁷ Modelling the take up of incorporation in an event-study-design framework, the chapter shows that incorporation has had no affect on hours of worked. As suggested by Wolfson and Legree (2015), the wealth gains of incorporation are therefore a tax expenditure on the part of the government. However, the chapter does find that incorporation increases hiring among female professionals; a result that is consistent with the cash flow benefits of incorporation (Baron, 2013) and gender difference in risk-taking in the labour market (Bertrand, 2011). This result prompts the need for further analysis of gender differences in the response to incorporation using administrative records that can more accurately identify hiring practices and speak to gains in terms of gender inequality (of earnings) among professionals.

In return for their self-determination, self-employed workers must deal with the potential vulnerability of their labour market status; characterized by both the increased variance in income and the common absence of labour market policy protections. For example, self-employed workers are not protected by a minimum wage and until recent reforms in Canada were ineligible for employment insurance and parental leave benefits. The antithesis to this perceived vulnerability is the unionized worker, who benefits from increased job security and collectively bargained wages that pay a premium above non-unionized peers. While Canadian unionization levels remain robust, unionization has seen a steady decline in the US since the early 1980s (Card et al., 2020). Indeed, the share of unionized wage workers in the US is now comparable to that of self-employed workers; a shift that undoubtedly would have been incomprehensible to economists in the 1960s. This well documented shift has been tied to the changing environment of labour regulation in the US and, more importantly, linked to rising wage inequality (Card, 1996; Freeman, 1993; Nunn et al., 2019). One the founding papers in this literature is DiNardo et al. (1996)'s *Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach*. Prior to their seminal contribution, the primary means of assessing the impact of labour market institutions - namely, the minimum wage or collective bargaining laws - on wage inequality was lim-

⁷Tax based instruments are common in studies examining the relationship between the corporate tax rate or individual-corporate tax wedge and the incorporation status of firms (among others, Egger et al., 2009; Gordon and Slemrod, 1998; Liu, 2014).

ited to an understanding of how these institutions affected average wages, or possibly the variance. However, inequality – as measured by standard indicators such as the gini coefficient – is a characteristic of the full distribution. [DiNardo et al. \(1996\)](#) transformed this conversation by developing a methodology for assessing the impact of institutions on the full wage distribution; thereby allowing for a more complete counterfactual analysis of inequality in the absence of these institutions.

Chapter 4, *Labour Market Institutions and the Distribution of Wages: The Role of Spillover Effects*, represents joint work with Nicole Fortin and Thomas Lemieux, published in the *Journal of Labor Economics*.⁸ This paper extends the original analysis of [DiNardo et al. \(1996\)](#) by both incorporating more recent data and developing a new methodology that can account for spillover effects. In [DiNardo et al. \(1996\)](#), labour market institutions are assumed to exhibit no spillover effects. The minimum wage has no impact on the wages of those earning more than minimum and union bargaining does not affect non-union wages. In the latter case, this implies that the distribution of non-union wages forms a valid counterfactual for unionized workers. However, there is evidence that such spillover effects exist, implying that [DiNardo et al. \(1996\)](#) potentially underestimates the contribution of weakening labour market institutions to rising inequality. Indeed, we attempt to provide some reduced form evidence of a union threat effect using the recent expansion of Right-to-Work laws in the US (an approach adopted earlier by [Farber, 2005](#)).

The paper adopts a distribution regression approach to identify the spillover effects of the minimum wage and union coverage; building on [Chernozhukov et al. \(2013\)](#), [Foresi and Peracchi \(1995\)](#), and [Fortin and Lemieux \(1998\)](#). The methodology is flexible and can be applied in years where there is no variation in the minimum wage at the state level. As in ([Lee, 1999](#)), the model estimates suggest that the minimum wage does indeed increase wages above the minimum. Accounting for these spillovers increases the overall contribution of declining minimum wages to increasing wage inequality during the 1980s; particularly, for women. In the absence of a better alternative, we use the state-industry unionization rate as a proxy for the threat of unionization (an approach adopted in the older literature, e.g., [Freeman and Medoff, 1981](#); [Podgursky, 1986](#)). As such, we are able to separately identify both the direct impact of unionization as well as the threat effect of unionization on non-union wages. We find that accounting for the threat of unionization doubles the overall contribution of declining unionization to the growth in wage inequality among men. Overall, this chapter makes an important contribution both in terms of its conclusion as well as its methodology.

⁸The paper presented here includes additional results excluded from the published version. Some of these results feature in earlier, working-paper versions of the paper (see [Fortin et al., 2019](#)), but do not appear in [Fortin et al. \(2021\)](#).

Chapter 2

Strategic self-employment and family formation

2.1 Introduction

Adjustments to female labour supply at the point of family formation are arguably the most important factor determining the lifecycle trajectory of female labour force participation and gender inequality (Angelov et al., 2016; Bertrand et al., 2010; Kleven et al., 2019b; Wilde et al., 2010; Zhang, 2010). This paper makes a significant contribution to our understanding of this crucial stage in the lifecycle by providing a close examination of selection into self-employment with the birth of a first child. Applying an event-study of initial childbirth, I show that family formation is associated with an increase in both maternal and paternal self-employment; explained largely by an increase in co-employed households. This is the first paper to identify such a significant shift in male labour supply with childbirth. Moreover, I develop a novel simulated instrument research design, which integrates into the event-study-design, to identify a causal link between this co-employment response to childbirth and the value of income splitting under a progressive, individual income tax structure.

This link between parental self-employment and the income tax structure places this paper within an important literature on the taxation of household income. The choice over tax unit – individual or joint household – has important implications for the interdependence of labour supply in married households (Kleven et al., 2009). Indeed, numerous studies show that joint tax structures suppress female labour supply as a result of the higher marginal tax rate faced by women married to high income men (Crossley and Jeon, 2007; Eissa, 1995; Fuenmayor et al., 2018; LaLumia, 2008; Selin, 2013). Individual tax structures subvert this mechanism. Unless you are self-employed. There is a growing body of evidence showing that self-employed households split income: they shift taxable employment income between household members to reduce their tax liability (Harju and Matikka, 2016; LaLumia, 2008; Schuetze, 2006; Stephens and Ward-Batts, 2004; Zinovyeva and Tverdostup, 2018).⁹

⁹In addition to income splitting, the self-reported nature of self-employment earnings has been the topic of much discussion, with the literature highlighting issues of under-reporting (Hurst et al., 2014; Schuetze,

This gives self-employed workers a tax advantage in households with an unequal income distribution. As childbirth is associated with a loss in female earnings, it presents as a positive ‘shock’ to these savings. For this reason, under an individual tax structure, there are additional pecuniary benefits to parental self-employment that complement other non-pecuniary ones (most notably flexibility: [Jeon and Ostrovsky, 2019](#); [Joona, 2017](#); [Noseleit, 2014](#)).

These findings are based on a parallel analysis of administrative and survey data from Canada, an individual tax jurisdiction. The Longitudinal Administrative Database (LAD, 1982-2016) includes detailed income records based on individual and joint tax filings. As such, there is no self-reported identifier of an individual’s class of employment. Recent efforts to measure the size of the gig-economy using individual tax records emphasize the inconsistencies between administrative records and traditional labour force surveys ([Abraham et al., 2018](#); [Burtch et al., 2018](#); [Jeon et al., 2019](#); [Katz and Krueger, 2019](#)). I construct a proxy for self-employment based on individual tax records that tracks the key trends in Canadian self-employment, as measured by the Canadian Labour Force Survey (LFS, 1988-2016). I then incorporate this proxy into an event-study of initial childbirth following the methodology of [Kleven et al. \(2019b\)](#).

The event-study results show that switching between the wage-paying sector and self-employment provides an important infra-marginal response to childbirth (see Figure 2.1). Such changes in employment mode are consistent with existing evidence that women switch jobs with childbirth in search of more ‘family friendly’ employers, such as the public sector ([Fuller and Hirsh, 2019](#); [Goldin, 2014](#); [Hotz et al., 2018](#); [Kleven et al., 2019b](#)). I find that female self-employment increases discontinuously after the birth of a first child by ~ 5 %-points relative to the pre-period; approximately 80% of the lifecycle increase in Canadian self-employment among women. While this is the first paper to estimate the impact of childbirth on self-employment in this way, the results are consistent with other studies that highlight the important relationship between female self-employment and childcare ([Jeon and Ostrovsky, 2019](#); [Joona, 2017](#); [Noseleit, 2014](#)). It is also consistent with the long held view that gender differences in the self-employed labour market - which typically exceed those in the paid sector ([Clain, 2000](#); [Hundley, 2000](#)) - can be explained by differences in motivation, with female self-employment prioritizing flexibility and work-life balance ([Allen and Curington, 2014](#); [Boden, 1999](#); [Georgellis and Wall, 2005](#); [Lombard, 2001](#); [Saridakis et al., 2014](#)).¹⁰

Even more striking, I find that the event of initial childbirth is associated with a size-
[2002](#)), bunching at tax thresholds ([Kleven and Waseem, 2013](#); [Saez, 2010](#)), and income shifting ([le Maire and Schjerning, 2013](#)).

¹⁰This is not without consequence, as self-employed workers are disproportionately represented in the top end of the income distribution ([Smith et al., 2019](#)).

able and persistent selection of married men out of the wage-paying sector and into self-employment. Male self-employment increases by ~ 4 %-points with the birth of a first child in married households; about a third of the lifecycle increase in male self-employment.¹¹ This contrasts the existing literature which finds no significant evidence of a large extensive margin response to childbirth among men (Angelov et al., 2016; Bertrand et al., 2010; Kleven et al., 2019b; Wilde et al., 2010; Zhang, 2010) and mixed evidence on the intensive margin of hours worked (Astone et al., 2010; Glauber, 2008; Hodges and Budig, 2010; Lundberg and Rose, 2000, 2002). I then show that selection into joint self-employment can account for the majority of the increase in parental self-employment; it increases discontinuously by ~ 3 %-points after childbirth. Evidence from self-reported data in Survey of Labour and Income Dynamics (SLID, 1996-2010) shows that this joint self-employment response is consistent with a co-employment response: employment in the same business. These results suggest that the motivations of male and female self-employment in married households are interlinked (Bruce, 1999). Indeed, evidence from the SLID shows that, conditional on the self-employment status of the male spouse, the increase in female self-employment is proportional to the increase in working from home after childbirth; suggesting that co-employment represents the creation of a flexible, ‘family friendly’ firm.

I provide corroborating event-study-design estimates from a parallel analysis of self-reported class of employment data in the SLID (1996-2010). This is important for two reasons. First, it validates the measurement exercise and confirms that any measurement error in the LAD’s proxy for self-employment is uncorrelated with the timing of childbirth. Second, it strongly suggests that the LAD estimates – in particular, the extensive margin increase in joint self-employment – represents an actual labour supply response. This is crucial, as findings based on tax reported employment income can easily identify a tax reporting response that masks non-economic activity or misidentifies the true class of employment. Without these comparable survey results, it would be impossible to distinguish between a story in which childbirth induces an increase in co-employment and one in which it increases self-employment for a single, income splitting parent. Indeed, there is a tendency to view co-employment - or matching on industry in self-employed households - as evidence of income splitting tax avoidance (Schuetze, 2006). In the context of family formation, this paper contends that co-employment is a real response to the employment challenges of childcare, and that income splitting is a matter of income reporting, not ‘fake’ job creation.

The reported income profile of self-employed households strongly suggests that they do split income. I show that the average tax rate of self-employed households declines by

¹¹Humphries (2019) uses machine learning tools to cluster different life cycle trajectories of male self-employment in Sweden using similar administrative records. The author identifies two “waves” of selection into self-employment: one beginning at the start of the career and a second beginning during the early 30s: years typically associated with family formation.

an additional ~ 3.5 %-points with childbirth, conditional on total income. These savings arise from a more equal income distribution: the relative taxable income of mothers in self-employed households falls by ~ 3 %-points less than the average decline in non-self-employed households after childbirth. These results are consistent with the existing evidence on income splitting, but this is the first paper to demonstrate how this behaviour changes with the birth of a first child (Harju and Matikka, 2016; LaLumia, 2008; Schuetze, 2006; Stephens and Ward-Batts, 2004; Zinovyeva and Tverdostup, 2018). In particular, my results are in agreement with Zinovyeva and Tverdostup (2018), who argue that self-employed households explain the presence of a discontinuity at 0.5 in the distribution of relative earnings among married couples (the ‘bread winner’ gap: Bertrand et al., 2015; Binder and Lam, 2019; Codazzi et al., 2018; Doumbia and Goussé, 2019; Hederos and Stenberg, 2019; Lippmann et al., 2020; Wieber and Holst, 2015). While Zinovyeva and Tverdostup (2018) focus on the event of cohabitation, this paper shows that it is the event of initial childbirth that really increases the probability of co-employment.

This paper differentiates itself from existing income splitting literature by examining the potential endogeneity of both female and male self-employment. LaLumia (2008) shows that with the switch to a joint structure in the US, female self-employment fell in households where the male spouse was self-employed. Similarly, Schuetze (2006) finds that the level of female employment in self-employed households in Canada is disproportionately higher than in the US. In both these papers, as in the rest of the income splitting literature, the self-employment status of the male spouse is taken as given. This is the first paper to show that the both female *and* male self-employment in married households is increasing in the value of income splitting. We should expect employment to be higher in Canadian self-employed households because a significant share of these households are self-employed for the very purpose of being co-employed and benefiting from income splitting. In addition, it is the event of family formation which ‘activates’ these tax benefits.

I develop a unique simulated instrument research design (Gruber and Cullen, 1996) that allows me to estimate a reduced-form income-splitting elasticity within the event-study-design of Kleven et al. (2019b). To do so, I build on the taxable income elasticity framework of Gruber and Saez (2002). Naturally, the value of income splitting depends on both the absolute and relative income of a couple. In this instance, it also depends on the income tax structure which varies considerably over the estimating period of 1988-2016. I adopt a control function approach to isolate the variation in the simulated value of income splitting that arises from exogenous changes in the tax-structure across birth cohorts. As a result, the identifying variation does not rely on labour supply differences across households with different income profiles, but rather on tax shocks to *similar* households in different cohorts. This is an improvement on the existing family taxation literature which typically exploits

variation across households with a high- and low-income male earners at a point of policy reform (Crossley and Jeon, 2007; Eissa, 1995; Fuenmayor et al., 2018; LaLumia, 2008; Selin, 2013).

I estimate a reduced-form income-splitting elasticity of ~ 0.5 .¹² The elasticity is of a similar magnitude for both men and women, and stable through event-time after childbirth. Given a potential tax-savings of 3% this elasticity can explain a 1.5% increase in male self-employment and co-employment after childbirth: half the observed increase in co-employment. As a benchmark for the magnitude of the elasticity, I execute a simple difference-in-difference analysis of male self-employment in the US and Canada, using the comparable Canadian Labour Force Survey (1988-2016) and US Current Population Survey (1988-2016). The increase in male self-employment associated with family formation (presence of a child, conditional on age) is $\sim 2.5\%$ in Canada and 0.85% in the US; yielding a difference-in-difference estimate of 1.6% conditional on industry composition. As the US and Canada differ with respect to their joint and individual income tax structures, this is consistent with the estimated income-splitting elasticity. Thus, it is not just female employment in Canadian self-employed households that is elevated by income splitting, but indeed the level of male self-employment with family formation (Schuetze, 2006).

In addition, this paper provides an explicit test for the income-splitting hypothesis. For income splitting to increase joint selection into self-employment with family formation it should be that the same households explain both the maternal and paternal response to these tax shocks. A test of equality between the magnitude of the reduced form elasticities of mothers and fathers is insufficient to prove this relationship. However, this ‘symmetry’ hypothesis can be tested using an IV estimator. I estimate the probability of maternal self-employment conditional on spousal self-employment when the male spouse’s status is instrumented with the tax shock.¹³ If income splitting is indeed a mechanism that pushes households into co-employment with family formation, then the IV coefficient should be equal to 1: the marginal father who enters self-employment because of the tax shock should enter with his spouse. This is because income splitting assumes that both individuals will report self-employment income. Indeed, the conditional probabilities estimated by this IV estimator are all close to 1 and significantly different from 0.

Together these results provide a more complete perspective on the family taxation debate. This is, in part, because I examine the issue within the context of childbirth. The literature has emphasised the increase in female labour supply associated with a switch from

¹²This estimate should be comparable with those from the family income tax structure literature. Indeed, Eissa (1995) estimates an elasticity of 0.7, and Selin (2013) provide similar estimates in the range of 0.5-1.

¹³This instrumental variable approach has the advantage of acting as correction for measurement error in the self-employment proxy. For this reason, it is superior to the approach of estimating a reduced form elasticity for observed joint self-employment as measured by the interaction of both the maternal and paternal self-employment proxies.

a joint to individual tax structure (Crossley and Jeon, 2007; Eissa, 1995; Fuenmayor et al., 2018; LaLumia, 2008; Selin, 2013). This paper shows that a more complete counterfactual for a joint tax structure is not simply more female labour force participation, but also more male self-employment and co-employment with family formation. Individual tax structures give self-employed households a tax advantage that is amplified by the asymmetric ‘shock’ of childbirth resulting in an incentive to select out of the wage-paying sector with family formation. When combined with the non-pecuniary benefits of self-employment, this mechanism is shown to have long run consequences for both male and female labour supply.

I begin with a discussion motivating the examination of self-employment in the context childbirth. In addition to standard considerations of flexibility, I expand on the mechanism through which the tax structure may incentivize higher levels of paternal self-employment with childbirth. I then provide the details of the administrative and survey data used in this study and describe in detail the methodology used to identify both incorporated and unincorporated self-employed workers in the administrative records (Section 2.3). The first part of the analysis involves an event-study-design decomposition of selection into self-employment around initial childbirth (Section 2.4). This is conducted on both the administrative data and comparable survey data. The survey data findings both corroborate the measurement exercise and support the real nature of these self-employment responses to childbirth. The second half of the paper identifies the link between selection into self-employment at the point of initial childbirth and income splitting (Section 2.5). Here I build on the income tax elasticity literature to estimate a positive response by both fathers and mothers to exogenous changes in the value of income splitting. The magnitude of these elasticities is then shown to be proportional to a simple Canadian-US difference-in-difference estimate of the male self-employment response to family formation.

2.2 Self-employment and childbirth: flexibility and income splitting

This paper adopts the position that self-employment represents a mode of employment.¹⁴ It is a ‘catch all’ for a broad set of labour relations that differ from the standard employer-employee contract. This includes business ownership, but also extends to freelance and gig-economy workers.¹⁵ Nevertheless, a defining characteristic of self-employment, as a mode of employment, is flexibility: the flexibility to determine when you work, where you work, and how intensely you work. In a 2016 survey of Canadian self-employed workers approximately 40% referenced the non-pecuniary benefits of ‘work-life balance’, ‘freedom to be your own boss’, and ‘less stress’ as the primary reason for being self-employed.¹⁶ For women the proportion is higher at 45%, while for men it is 37%. In contrast, 14% of men reference ‘earning potential’ as their primary motivating factor, compared with 9% of women. This simple gender difference has long been promoted as the main explanatory factor for the larger gender earnings gap in the self-employed labour market (Allen and Curington, 2014; Boden, 1999; Clain, 2000; Georgellis and Wall, 2005; Hundley, 2000; Lombard, 2001; Saridakis et al., 2014).

Flexibility is the primary reason we might expect *female* self-employment to increase with childbirth. Indeed, the notion that there are two, gendered, paths to self-employment is one of the oldest in the self-employment literature (Carr, 1996). A number of more recent studies examine the relationship between female self-employment and childcare responsibilities, emphasising the value of mothers place on the ability to balance employment and childcare responsibilities (Jeon and Ostrovsky, 2019; Joona, 2017; Noseleit, 2014). This fits within a broader discussion on the importance of flexibility in the wage-paying sector and its

¹⁴In many empirical applications self-employment, as identified in survey and administrative dataset, is used as a proxy for entrepreneurship or firm/business ownership. This research trend dates back to the early 1990’s and now represents a large body of research (Acs et al., 1994; Storey, 1991). It is problematic as it equates the creative and risk taking characteristics of Schumpeter (1921)’s entrepreneur with the observed characteristics of self-employed workers, many of whom are self-employed out of necessity and do not demonstrate a drive for innovation and growth (Bjuggren et al., 2010; Faggio and Silva, 2012; Hurst and Pugsley, 2010). Own-account self-employment (those without employees) is also widely used as a proxy for the informal sector; in particular, in developing world labour markets. This too is a problematic perspective as it both limits the set of workers in the informal sector to own account workers, when many are employers and employees, and places self-employed workers outside of the regulated economy. The position here - of self-employment as a mode of employment - fits better within an older, established literature on alternative work arrangements (Abraham, 1988; Barker and Christensen, 1998). This literature has recently been revisited in reference to the growing gig-economy (Abraham et al., 2018; Katz and Krueger, 2019; Stanford, 2017). However, it would be mistaken to equate all self-employed workers with the same vulnerability that has become associated with gig work (Kuhn, 2016).

¹⁵For this reason, it is a mode of employment that may be more common within specific occupations; for example, regulated professionals and tradesman for whom it is common to set up an independent practice/business.

¹⁶These results can be found in Table A.4 of Appendix A.5.

consequences for employment and wage equality (Fuller and Hirsh, 2019; Goldin and Katz, 2016). However, the problem with a gendered comparison of self-employed workers is that within married households self-employment often begets self-employment (Bruce, 1999). A male self-employed worker may emphasise higher earnings because his co-employed spouse emphasises work-life balance. Flexibility may therefore be an important indirect motivator behind male self-employment.¹⁷ Any analysis of the relationship between self-employment and childcare should account for co-employed households. In this paper, I show that joint selection into self-employment in married households is even more important at the point of family formation.

If it is a joint labour supply decision then it may also be one that is influenced by the ‘jointness’ of the tax structure. There is a large body of evidence supporting the claim that joint tax structures, which pool a household’s income, reduce female labour supply (Crossley and Jeon, 2007; Eissa, 1995; Fuenmayor et al., 2018; LaLumia, 2008; Selin, 2013). These studies are all based on natural experiments involving substantial tax reforms. In the case of Selin (2013) and Fuenmayor et al. (2018) the authors examine instances in Europe of a switch from a joint to individual structure, while LaLumia (2008) examines the US’s switch from an individual to joint structure in 1948. These studies all exploit a difference-in-difference research design using a treatment and control group based on the income of the male spouse. While they certainly account for family structure (i.e. presence of children), the analyses are not focused on labour market responses before and after childbirth. What this paper highlights is the interaction between the tax structure and decisions taken at the event of childbirth. I argue that individual tax structures incentivize co-habiting couples to coordinate their labour supply around the event of childbirth, while joint tax structures do not. In what follows I explain why, in addition to flexibility, we should expect an increase in both male and female self-employment with childbirth under an individual tax structure.

Consider a two-earner household approaching the birth of their first child. All the evidence points to the fact this event is associated with a long lasting decline in female labour market earnings (Angelov et al., 2016; Bertrand et al., 2010; Kleven et al., 2019b; Wilde et al., 2010; Zhang, 2010). Let $\{z_t^1, z_t^2\}$ denote the labour market earnings of each earner in period t , with $t = 0, 1$ denoting the periods before and after childbirth. Assume for simplicity that $z_0^1 = z_1^1$: paternal earnings do not change with childbirth, a fact which is supported by the parent penalty literature (Kleven et al., 2019b). The inequality below relates the discrete change in after-tax earnings of the household with childbirth under a

¹⁷Co-employment may also provide an important path through which fathers can increase his hours of work (Astone et al., 2010; Glauber, 2008; Hodges and Budig, 2010; Lundberg and Rose, 2000, 2002). However, I find little evidence to support this claim. With the exception that in cross-sectional Canadian time-use survey, self-employed men do spend more time engaged in their primary vocation and spend less time engaged in domestic activities compared with paid employees (not shown here).

joint and individual tax structure,

$$\underbrace{z_1^2 - z_0^2 + 2T(\bar{z}_0) - 2T(\bar{z}_1)}_{\text{joint structure}} \geq \underbrace{z_1^2 - z_0^2 + T(z_0^2) - T(z_1^2)}_{\text{individual structure}} \quad \text{for } T'(z) \geq 0, T''(z) \geq 0 \text{ and, } z_1^2 \leq z_1^1$$

where \bar{z}_t represents the average income of the household in period t . The joint tax structure generates a automatic income effect by lowering the average tax rate of both parents, while under the individual structure only the mother's marginal tax rate falls (Kabátek et al., 2014; Kleven et al., 2009). The inequality holds in cases where the tax structure is progressive, and the father earns more than the mother before childbirth.¹⁸ From a purely pecuniary perspective then, a household with such an income profile would prefer to reside in joint tax jurisdiction.

The same outcome can be generated under an individual structure through a transfer from the high to low earner; however, third-party reporting of employment income typically ensures that individuals are taxed on their individual earnings.¹⁹ The exception being self-employed workers who self-report their employment earnings, providing them with the unique opportunity of splitting income with a spouse. While of the evidence points to such tax avoidance within self-employed households, no studies have shown households become self-employed for this reason (Harju and Matikka, 2016; LaLumia, 2008; Schuetze, 2006; Stephens and Ward-Batts, 2004; Zinovyeva and Tverdostup, 2018).

The interaction between the asymmetric event of childbirth and individual income tax structure generates an additional pecuniary benefit to self-employed at a point in the lifecycle when the non-pecuniary benefits are particularly acute. The same tax incentive to select out of the wage-paying sector does not exist in a joint income tax jurisdiction. We could therefore expect to observe switching between the wage-paying sector and self-employment among married men with childbirth under an individual tax structure. This should also be followed by a conditional increase in *tax reported* female self-employment income. This maternal response need not be real, as the tax benefits simply necessitate a tax-reporting response. Indeed, the tax savings would be larger under a pure accounting response, where the mother is economically inactive. However, the presence of a now self-employed spouse also creates a co-employment opportunity, which may be preferred to a pure tax reporting response if the new enterprise can meet the flexibility requirements

¹⁸In derivative form, the inequality holds if $T'(\bar{z}) \geq T'(z^2)$ which under a progressive tax structure, $T''(z) > 0$, requires that $z^2 \leq z^1$

¹⁹Of course non-labour market income, such as rental income, can more easily be reallocated within a household in a tax efficient manner. Indeed, I provide evidence that even in wage-employed households (households where the male spouse is a paid employee) total income has a more equal distribution to employment income; particularly, among couples with children (see Appendix A.7).

of the new family.

In Appendix A.1 I demonstrate this exact outcome in a simple model. This model depicts the optimal labour supply paths of married households as they jointly choose both their mode of employment (self-employed or paid employee) and hours of work along a predetermined path of family formation. The timing of childbirth and the earnings under each mode of employment are known. Self-employed individuals automatically split income, even outside of co-employment, while *real* co-employment has childcare benefits. The model demonstrates that income splitting increases the share of *real* co-employed households with childbirth. Paternal self-employment increases even before childbirth, but the increase in maternal self-employment (and co-employment) is postponed till after the event.²⁰ Without childbirth, one observes a small increase in male self-employment, but no increase in co-employment.²¹ This simple model demonstrates the fact that with childbirth we can expect an increase in real co-employment from a tax incentive typically associated with tax avoidance. A result that is driven by the complementary childcare benefits of co-self-employment and the tax benefits of income splitting.

2.3 Data & Measurement

This paper is based on a parallel analysis of Canadian administrative tax records and household survey data. The Longitudinal Administrative Dataset (LAD, 1988-2016) forms the basis of the event-study of initial childbirth and the simulated instrument research design. I use the monthly Labour Force Survey (LFS) and two household income surveys - the Survey of Labour and Income Dynamics (SLID, 1996-2011) and the Canadian Income Survey (CIS, 2012-2016) - to benchmark my measurement of self-employment in the LAD. In addition, I replicate the event-study of initial childbirth using the longitudinal version of the SLID (1996-2010), and exploit this survey's richer data to explore qualitative changes to employment, self-employment, and co-employment with childbirth.

2.3.1 Identifying childbirth in the LAD

The LAD is a longitudinal administrative dataset based on a 20% sample of the T1 Family File (T1FF), the master file of individual tax returns beginning in 1982.²² In this paper I

²⁰This difference in timing is in part because of the disruption of childbirth, but also because of the employment insurance (maternity leave) structure, which only insures paid employment.

²¹This male self-employment response is driven largely by the fact that the average female worker is assumed to earn 80 cents on the dollar. As such, there are fewer households where the wife would benefit from income splitting where she to become self-employment.

²²The LAD is sampled at the individual, not household level, from the T1FF. Once sampled, an individual remains in the sample so long as they file an individual tax return or a joint tax return in which case individual

restrict my analysis to the years 1988-2016, while utilizing the full sample to consistently identify the event of initial childbirth. I do this for two reasons. First, the 1988 federal tax reform introduced several changes to the way employment income is reported that improve my ability to identify self-employed workers. Second, the 1988 reform removed a family tax benefit that allowed married households to shift a limited amount of income from a high to low earner.²³ Thus, for period 1988-2016 Canada's income tax structure can be described as an individual tax structure.

I use self-reported information on the age of the seven youngest children in the household to identify the year of initial childbirth in each household. Given the longitudinal nature of the data this can be applied retroactively, to identify event-time prior to initial childbirth and account for missing records. In Canada various federal and provincial child support grants provide an incentive for parents to complete this information when filing their tax returns; thereby improving the reliability of such data.²⁴ However, one major limitation is that in the T1FF (and subsequently the LAD) this information is attached to a single parent. I find that in the vast majority of cases information on the age of children is attached to the mother in the LAD dataset. This rules out the possibility of estimating a separate analysis for men and women, except in the case of married couples. I therefore limit my analysis to joint-filing married and common-law couples, and do not include a control group of women who never have children.²⁵

For this reason it is important that I observe a representative sample of co-habiting couples; both married and common-law.²⁶ Figure A.3 in Appendix A.2 plots the trend in cohabitation among men and women aged 25-44 in LAD. The comparable series estimated using the LFS and SLID/CIS suggests that the LAD sample does not significantly mis-

records are constructed through imputation. Individual sampling means that a sample member's spouse may not be in the LAD, but for all married couples one observes the tax details of the spouse for each variable. The LAD includes basic demographics such as age, gender, and census family composition.

²³This reform is studied by [Crossley and Jeon \(2007\)](#) in relation to female labour supply; although, self-employed households are not included in their analysis. A similar tax credit was introduced for a limited period in 2014 (retroactively) and 2015.

²⁴The key federal tax benefits include the Family Allowance (1982-1992), the Child Tax Benefit (1993-2005), the Universal Child Care Benefit (2006-July 2016), and the Canada Child Benefit (July 2016 onwards). Changes to the Family Allowance program in 1988, and its replacement with Child Tax Benefit further incentivized reporting of children's ages. This provides a third motivation for focusing on the post-1987 subsample.

²⁵Selecting on married households is relatively common in the parent penalty literature (as in [Angelov et al., 2016](#)). In general, I do not find that selecting on marriage changes the results of the event-study-design in a significant way. With the exception that the pre-trend of employment is flatter among married women (see Figure 2.3). This paper primarily concerns decisions made by co-habiting couples. I choose not include a control group of women who never have children for the aforementioned reasons related to the construction of the LAD. There is a chance this sample would include women who are in fact mothers, but for reasons outside of my control have not been assigned with this information when constructing the LAD. [Kleven et al. \(2019b\)](#) provide a discussion on this topic, and generally find that including such a control does not significantly change the results.

²⁶The LAD does include information on whether a joint filing couple is a common-law or married couple.

identify cohabiting couples. Further details on how the sample is constructed are provided in Appendix A.2.

2.3.2 Identifying self-employment in the LAD

In survey data the self-employment status of a worker is typically identified by an individual's self-reported class of employment. This is the case in the LFS and SLID/CIS which include a self-reported class of employment of an individual's main and second/additional jobs.²⁷ However, most administrative records do not include such a self-reported identifier, which adds an additional measurement exercise to this paper. This is non-trivial as survey data and administrative records can provide vastly different estimates on the level of self-employment. Consider the recent attempts to measure the size of the gig-economy using both administrative and survey data (Abraham et al., 2018; Burtch et al., 2018; Jeon et al., 2019; Katz and Krueger, 2019). In this paper I create a proxy for self-employment using the various sources of employment income in the LAD that closely tracks the key trends in male and female self-employment found in comparable survey data.

A particular challenge will be ensuring that the proxy identifies both unincorporated and incorporated self-employed workers. The LFS shows that there has been a clear shift towards incorporated and own-account (without employees) self-employment in Canada over the past three decades (see Figure 2.2).²⁸ There is also a distinct break in trend around the turn of the century where unincorporate numbers begin to decline. Two important factors explain this shift towards incorporation. These are discussed in greater detail in Chapter 3. First, there has been a 'race to the bottom' of corporate tax rates both at a federal and provincial level (see Figure 3.5). Second, reforms to provincial legislation as well as acts governing professions within each province have allowed for the establishment of *professional* corporations (see also discussion in Wolfson and Legree, 2015). What is important for this chapter is that I find no evidence that the professional corporation reforms discussed in Chapter 3 affected the self-employment rate among professionals, only the incorporation rate among self-employed professionals (see Figure 3.4). This is significant, as the turn of the century also is also characterized by significant tax reforms that will feature as a source of identifying variation in this chapter.

As there is no self-reported identifier of class of employment in the LAD, I allocate each individual a self-employed, paid employee, or not employed (neither self-employed nor

²⁷The SLID and CIS surveys collect information on all jobs held in the last year, including class of employment. In the case of the SLID, the survey administrators then assign one of these jobs as the individual's main job based on hours of work in the past year. The CIS has seized this distinction.

²⁸These figures exclude unpaid family workers, who are typically women married to self-employed men. The number of unpaid family workers has been in a steady decline for the last four decades, and since the turn of century is largely obsolete.

paid employees) status based on their primary source of taxable employment income. In this way I am not only identifying self-employment, but implicitly assigning each individual a ‘main job’ based on their annual tax records. This requires separating out self-employment income from wage/salary employment income, and then developing a set of rules that can be applied uniformly across the sample period and within a married household.²⁹³⁰

In the LAD one observes separate filings of unincorporated self-employment income: grouped by business, professional, farming, fishing, and other self-employment income.³¹ Both gross and net income are observed, with the latter taxable.³² Incorporated self-employment income is not explicitly identified and ownership of a Canadian Control Private Corporation (CCPC) is only observed in the LAD from 2000.³³ However, corporate income enters the household through one of two channels: dividends or wage/salary income. In the case of wages, I rely on the fact that in Canada wages paid to a shareholder of a CCPC owner (with equity > 40%) are uninsured: ineligible for employment insurance (EI) contributions. Similarly, wages paid to an individual within an ‘arm’s length’ of the business owner - such as a spouse or child - *may* not be insured.³⁴ I therefore use the observation of *large* uninsured wage receipts, as well as *large* dividend receipts, to identify instances of incorporated self-employment.³⁵ A more detailed discussion on the thresholds

²⁹Between 1982 and 2016 there are several changes to the way self-employment income is recorded, as well as additions to the LAD dataset that undoubtedly improve the identification of self-employed workers in more recent years. For example, from 2000 the LAD has included an indicator from T2 tax records on for whether the individual is a 10% shareholder in a Canadian Control Private Corporation (CCPC). However, many of these changes occur alongside important policy reforms; for example, the 2000 federal tax reform. Allowing my definition of self-employment to change during the sample period could therefore introduce a spurious correlation with these reforms. To avoid this outcome, I apply a consistent definition of self-employment throughout the period and where possible use any updates to the tax files as a means of benchmarking my estimates.

³⁰For almost all records the LAD provides individual values as well as the total value for the joint-filing couple. Prior to cohabitation, this variable measures the total value of the parents in the Census family, which is equal to the individual value in a single household. In almost all circumstances, this variable can be used to calculate the spouse’s value. For this reason, I follow the advice of Statistics Canada, and define an individual’s marital status based on the census family variable, and not self-reported marital status.

³¹From 1988 I separately observe taxable profits (or losses) from a limited liability partnership. Prior to 1988 this information was not well recorded.

³²For gross records, the LAD provides only the maximum value within a joint-filing couple, while the sum of the net values is included. For this reason, the gross value cannot be used in the definition of self-employment, and the net value which can be negative must be used.

³³This variable is not observed for a spouse, which in addition to the fact that it is only observed from 2000 onwards, prevents it from being used.

³⁴The following [link](#) outlines Canada’s policy concerning the hiring of family and contribution to EI. There are circumstances in which the employment of a relative may be deemed at arm’s length.

³⁵From 2005 one can separately identify dividends received from a small and large, as after 2005 the two types receive a different mark-up to account for the small business tax exemption. As such, dividends from small and large firms are filed separately, helping to separate out dividends from smaller CCPC businesses and larger (potentially publicly traded) companies. As this cannot be done for years prior to 2005 I do not use this variable, but instead use it as a benchmark. I find that my identifier of *large* captures approximately 70% of small business dividend receipts.

used to denote a ‘large’ value can be found in Appendix A.3.

This definition of incorporated self-employment income has two important implications. First, I will likely underestimate incorporated self-employment if individuals incorrectly contribute to EI or if they have insured wages from other paid work.³⁶ Incorrect contributions can be expected as any over payment of EI premiums is simply deducted from federal taxes owed upon submission of a tax return.³⁷ Second, this definition of self-employment may include spouses who are employed in a family business. Such individuals may not report as self-employed in a household survey. This possibility extends to the case of unincorporated businesses owners, who may also employ a spouse in an uninsured capacity.³⁸ For this reason, my measure of self-employment in the LAD is more broadly one of employment within a household firm.

Once I have identified potential sources of self-employment income, I define an individual as self-employed if their total self-employment income is more than half their total employment income.³⁹ To account for the fact that taxable self-employment income can be negative I use the definition below that incorporates the absolute value of self-employment income. In this way if an individual has insured wage receipts of \$40,000, but a self-employment loss of \$100,000 they are considered self-employed.

$$Q = \mathbf{1} \left\{ \frac{|income_q|}{|income_q| + income_w} > 0.5 \right\}$$

Figure A.4 depicts this measure of self-employment from the LAD with a comparable measure using the LFS. As predicted, the LAD’s self-employment rate for men underestimates that of the LFS by ~ 2 %-points. Moreover, this is because I underestimate the number of incorporated self-employed men (see Figure A.5). For women, there is less evidence of an underestimation of self-employment, after 1990. This reflects the broader definition of self-employment within the LAD which incorporates employment within a household firm, thereby counteracting any negative bias from the underestimation of incorporated self-employed workers. Indeed, this is confirmed by Figure A.6 which depicts a higher measure of joint self-employment among married individuals within the LAD. Despite

³⁶From 2000 onwards I observe the number of T4 returns (wage/salary income) submitted. However, I only observe the total value of all T4 returns and total employment insurance premiums paid. With these it is not possible to calculate the share of T4 income that is insured.

³⁷Unfortunately, this variable is not included in the LAD.

³⁸This introduces the possibility changes in self-employment status around family formation could simply capture changes in EI contributions for the same job. Fortunately, the parallel evidence provided by the longitudinal analysis of the SLID (see Section 2.4.2) suggests increases in employment in a household firm are driven by a change in industry, and not EI contributions.

³⁹Note, this definition of employment income includes dividends which are typically included under investment income. This is to account for incorporated income streams. Other important market income sources, such as rental income, are excluded from this definition.

these slight differences the LAD's proxy for self-employment does track the rise in both male and female self-employment through the 1990's, and the shift towards incorporation beginning in the early 2000's. For more details, please consult Appendix A.3.

2.4 Self-employment and family formation: event-study analysis

This paper uses an event-study-design (ESD) methodology to describe selection into self-employment as a function of family formation, where family formation is described by the birth of a first child. This methodology has been popularized by Henrik Kleven and co-authors in a series of recent papers on the parent penalty (Kleven et al., 2019a,b). One of the strengths of this methodology is that it can be applied in settings where there is no control group, so long as the timing of treatment is staggered. This lends itself to the analysis of childbirth, as different cohorts of mothers present as a staggered 'treatment' and women who never have children may not be a reasonable control group.

In this section I follow Kleven et al. (2019a)'s methodology and estimate the simple, but flexible, ESD model described in Equation 2.1. It includes a set of dummy variables for event-time (denoted by subscript s), age, and calendar year (denoted by subscript t). Event-time is defined as years since year of first birth (or birth cohort, b_i): $s = t - b_i$. In this paper, I focus on the outcome of self-employment, denoted by the indicator function Q_{its}^g , where g indicates the gender of the parent.⁴⁰

$$Q_{its}^g = \sum_{j \neq -1} \alpha_j^g \cdot \mathbf{1}\{j = s\} + \sum_k \beta_k^g \cdot \mathbf{1}\{k = age_{its}\} + \delta_t^g + \varepsilon_{its}^g \quad g = 1, 2 \quad (2.1)$$

The event of childbirth does not lend itself easily to causal inference. It is typically anticipated and more often than not planned. Kleven et al. (2019a) justify the use of this methodology on the basis that childbirth generates sharp discontinuities in the labour market outcomes of parents that are "arguably orthogonal to unobserved determinants of those outcomes as they should evolve smoothly over time". Their argument is more robust in the initial few years after childbirth. For longer lags (and leads) the 'smoothness' assumption becomes more tenuous and their approach relies more heavily on the non-parametric estimation of the age-profile and year fixed effects (FEs).

Staggered-adoption research designs have been the topic of discussion in number of recent papers (Abraham and Sun, 2018; Athey and Imbens, 2018; Borusyak and Jaravel, 2018; Goodman-Bacon, 2018). They are also a special case of the broader class of two-way FEs methodologies (see discussion in de Chaisemartin and D'Haultfœuille, 2020). The literatures highlight two issues which I wish to address: treatment effect heterogeneity and

⁴⁰Throughout the chapter I use $g = 1$ to denote the male spouse and $g = 2$ to denote the female spouse.

underidentification.

Abraham and Sun (2018) discuss the issue of treatment effect heterogeneity in the context of ESDs. By defining a measure of the cohort-specific average treatment effect on the treated ($CATT_{b,s}$ where b denotes birth-cohort and s event-time), the authors demonstrate that in standard event-study design specification (which includes two-way year and unit/cohort FEs) the population regression event-time coefficient is a non-convex weighted average of all $CATT_{b,s}$'s: across both cohort and *all* event-times.⁴¹ Notably, these weights may be negative, but do sum to 1 for the event-time of interest ($s = j$) and 0 for all other event-times ($s = j'$). As a result, an assumption of treatment homogeneity is required to interpret the population regression event-time coefficient as the average treatment effect on the treated.⁴²

Equation 2.1 deviates from the standard event-study specification as it does not include unit/cohort FEs.⁴³ Instead, the model includes age dummies. Unfortunately, we cannot simply reframe the discussion of cohort-specific ATT to one of age-specific ATT, because cohorts are intrinsically linked to event- and calendar time, while age is not. For α_j in Equation 2.1 to identify the $CATT_j$ we still require that $CATT_{b,j} = CATT_j \forall b$; that is, homogeneity across birth cohorts.

With the inclusion of the non-parametric age-profile, this assumption implies that average impact of childbirth on the decision to enter self-employment can only vary across birth cohorts in so far as the age composition of each cohort changes. The underlying age-profile of self-employment must therefore be independent of birth cohort and stable over time; as demonstrated by the absence of an interaction term in Equation 2.1. This latter assumption may be problematic if the education and occupation attainment of women changes over time, resulting in different self-employment life-cycle trajectories. However, it can also be relaxed through the inclusion of an interaction between age and year dummy FEs.⁴⁴

The exclusion of cohort (or individual) FEs in Equation 2.1 relates to a secondary issue concerning event-studies: that of underidentification. It is standard to exclude the event-time dummy corresponding to the period before treatment ($s = -1$, as shown in Equation 2.1). With the inclusion of two-way FEs - year and treatment unit (individual/cohort) FEs - one must exclude at least one more event-time dummy (see discussions in Abraham and Sun, 2018; Borusyak and Jaravel, 2018). If not, the event-time dummies are only identified up to a linear trend in event-time. The issue of underidentification can be solved through

⁴¹ A similar argument is made in de Chaisemartin and D'Haultfœuille (2020) concerning the broader class of two-way FE models.

⁴² A direct result of the fact that the weights sum to zero for event-times $s = j'$.

⁴³ This special case of event-study designs is commonly used in contexts where the timing of the event is random. Abraham and Sun (2018) address this special case and demonstrate that the population regression event-time coefficient remains a non-convex weighted average of $CATT$ across all cohorts and event-times.

⁴⁴ The results reported here do not include such an interaction.

the use of a ‘never-treated’ control group, which help to pin down the year FEs (Borusyak and Jaravel, 2018). Other common solutions are to exclude unit FEs entirely or exclude a larger set of event-time dummy variables. For example, if you assume that there is no pre-emptive behaviour ($\alpha_j = 0 \forall j \leq -1$) then the pre-treatment period can be used to pin down the unit FEs independent of event-time and year FEs. Borusyak and Jaravel (2018) describe such a set up as semi-dynamic specification; a methodology which I apply in Chapter 3 where the institutional setting permits such an assumption.

In this chapter, I follow Kleven et al. (2019b) and exclude treatment unit FEs (either individual or cohort FEs). I do so primarily as a means to test for pre-emptive self-employment responses to childbirth, but secondly, because women who never have children are arguably not a reliable ‘never treated’ control group.⁴⁵ This naturally, requires an additional assumption regarding the role of cohort effects. Specifically, Equation 2.1 assumes that holding the age profile of different cohorts fixed, differences in selection into self-employment across cohorts are independent of responses to childbirth. For the estimation of immediate event-time lags (and leads), which makes use of differences across ‘close’ cohorts, such an assumption is plausible. However, for longer lags (and leads) this assumption is more tenuous. For example, a reduction in wage-discrimination in paid labour market over time may affect parent’s decisions to enter the self-employment with childbirth. In practice, I find that the self-employment response to childbirth has remained remarkably stable over time.⁴⁶ Moreover, it has declined in accordance with the policy changes discussed later in Section 2.5 and will be accounted for in later specifications (see Section 2.5.1).

After estimating the model each event-time coefficient α_j is rescaled by the expected value of the outcome variable under the counterfactual that the event did not occur. In order to compare selection into self-employment with employment in the wage paying sector, I normalize both outcomes by the predicted level of employment (denoted below by \tilde{Y}_{its}^g).⁴⁷ Plots of the unadjusted event-time coefficients can be found in Appendix A.4.

$$P_t^g = \frac{\alpha_j^g}{E[\tilde{Y}_{its}^g|t]}$$

⁴⁵As emphasised by Abraham and Sun (2018), testing for pretends requires an assumption of homogenous treatment effects.

⁴⁶This statement is based on a specification of equation 2.1 which interacts event-time and year FEs (Kleven et al., 2019b, as in). These results are not presented here.

⁴⁷This is the standard methodology for estimating relative changes in such models, as opposed to the use of a log transformed outcome variable. It has the added benefit that the model can be estimated using level outcomes - such as annual taxable income - without the exclusion of 0 values. This is useful in the context of the LAD data as I do not observe weeks or hours worked, part-time or full-time status. I can only observe participation based on the filing of employment income. When modelling level employment income (including 0 outcomes) the estimates will encapsulate both the wage and labour supply response to childbirth.

2.4.1 The self-employment ‘parent penalty’

Figure 2.1 plots the self-employment ‘parent penalty’ alongside the comparable paid employment penalty, both normalized by the same employment counterfactual.⁴⁸ The maternal response suggests that in 5 years after a mother’s first child is born self-employment increases by approximately 5 %-points relative to the level in the base period ($s = -1$). There is some evidence of a decline beyond year 8. In contrast, employment falls sharply in the wage paying sector, and even 10 years after initial childbirth is still 18% lower than the predicted employment counterfactual.⁴⁹ While the increase in maternal self-employment is a fraction of the decline in wage-employment, when normalized by the predicted level of self-employment the 5 %-point increase represents a 65% in the predicted value of female self-employment. It is also approximately 80% of the lifecycle increase in female self-employment, once you account for the broader definition of self-employment (see Figure 2.7).⁵⁰ This is consistent with the flexibility benefits of self-employment that appeal to new mothers wishing to balance work and childcare responsibilities (Hundley, 2000; Jeon and Ostrovsky, 2019; Joona, 2017; Noseleit, 2014).

More striking then is the evidence of switching between the wage-paying and self-employment among married men. This trend partly pre-empts the event of childbirth, beginning as early as 5 years before the birth of a first child (see Figure A.7). By event-time 10, paternal self-employment has increased by 3 %-points relative to the base period. Between event-time -5 to 5 paternal self-employment increases by over 4 %-points. This represents about a third of the increase in male self-employment over the life-cycle in Canada, as it increases from around 8-20% between the ages of 25 and 50 (see Figure 2.7). Consistent with the literature there is no evidence of a drop in paternal employment after childbirth (see Figure 2.3).

A striking feature of this shift into self-employment with family formation is its persistence and stability. Afterall, small businesses typically face a high failure rate (see discussions in, Cader and Leatherman, 2011; Headd, 2003; van Praag, 2003). However, there are many

⁴⁸While the self-employment indicator is based on the share of employment income from a self-employment source, the indicator for paid employment used in this and similar figures simply identifies individuals with insured T4 earnings. Thus, an individual can be identified as both a paid employee and self-employed. However, these cases are infrequent and in this application the difference between self-employment and paid employment ESD results is explained by an overall employment difference.

⁴⁹The pre-trends suggest that employment in the wage-paying sector increases marginally between event-time -3 and -2. This is consistent with the fact that maternity benefits are issued through the employment insurance structure in Canada and until 2011 self-employment income was not insured. Thus, there is an added incentive to be a wage-employee before childbirth. I discuss this in more detail below.

⁵⁰As measured by the LFS, the increase of 5 %-points is approximately 100% of the increase in female self-employment over the life cycle. However, given the broader definition of self-employment in the LAD that includes co-employed mothers who report as paid employees in survey data, I estimate that explains closer to 80% of the lifecycle increase.

stable paths to self-employment that are not associated with such risk; notably those in the trades and professions.⁵¹ Humphries (2019) finds a similar path of stable self-employment for men in Sweden. The author uses machine learning methods to cluster different life-cycle trajectories of self-employment using longitudinal administrative data. They find an initial wave at the start of the career, followed by a second wave in the early 30s: years typically associated with increased family formation. These stable self-employed careers are also evident in Zinovyeva and Tverdostup (2018).

Consistent with the hypothesis of this paper, most the increase in maternal self-employment can be explained by a discontinuous increase in the incidence of joint self-employment: both husband and wife are self-employed (Figure 2.4). Moreover, while maternal self-employment declines after year 8, the increase in joint-self-employment is stable at 3%. Thus, joint self-employment is more persistent than maternal self-employment. This increase in joint self-employment is also depicted in Figure 2.5, which shows how the conditional probability of maternal self-employment changes dependent on the self-employment status of the spouse. This is done by interacting the event-time and age-dummies with an indicator for spousal self-employment. Here, I use the contemporaneous spousal status and not the pre-treatment status in event-time $s = -1$. This is done for three reasons. First, to avoid selecting on measurement error in the self-employment proxy. Second, this is the conditional probability that underlies the joint self-employment response in Figure 2.4. Third, in Section 2.5.2 I instrument the male spouse's contemporaneous self-employment status from this same specification with an exogenous shock to the value of income splitting. While this approach incorporates the changing selection of men into self-employment, it is striking to see the discontinuous jump in this conditional probability. Moreover, I find that the change in conditional probability of maternal self-employment is remarkably similar between households where the male spouse was self-employed in the year before childbirth and those where the household is currently self-employed, suggesting that childbirth is the primary factor changing the conditional probability of maternal self-employment and not selection.⁵²

Maternity leave benefits likely play a big role in determining the discontinuous increase in conditional probability maternal self-employment immediately after the first child is born. Prior to a 2011 Federal reform, self-employment earnings were not insured under Canada's employment insurance structure. The exception here is Quebec, which introduced its own Quebec Parental Insurance Plan (QPIP) in 2006 which extended coverage to self-

⁵¹Evidence from survey data confirms that a disproportionate share of self-employment associated with family formation can be linked to trades (i.e., construction) and professional occupations. Similarly, uninsured wages (a primary identifier of incorporated business ownership) are closely associated with professional industries (healthcare, legal, business services, etc.) and construction. This data is only available from 2000 onwards. Unfortunately, industry data is not collected for unincorporated business income.

⁵²These results have not been included in this version of the paper.

employed workers. Since 2011, self-employed workers in the rest of Canada have had the option to *voluntarily* contribute to the federal EI system. However, for most Canadians, and the majority of the time period used to estimate these figures, self-employment income remained uninsured. For this reason, a mother must be paid employee leading up to the event of childbirth to receive EI benefits. In 1991, the Federal government increased its policy of 15 weeks of maternity leave (established in 1971) with the addition of 10 weeks of shareable parental leave. In 2001, it was again extended to 35 weeks; covering individuals up to a maximum of 1 year. Throughout this period, the Federal employment insurance structure has an income replacement rate of 55%.⁵³

The length of job protected leave matters too. Baker and Milligan (2008) provide a detailed discussion on the province-specific expansion of job protected in Canada. Alongside the federal parental leave reform of 1991, most provinces expanded job protected leave from 17-18 weeks to 29-34 weeks in 1991. Quebec increased their coverage to a full year, while Alberta and Saskatchewan delayed any reform.⁵⁴ Between 2000-2001 all provinces extended their policies to at least 52 weeks, while Quebec moved to 70 weeks. The absence of extended job protected leave in Alberta and Saskatchewan in the early 1990's is consistent with the fact that these are the only two regions where joint self-employment appears to pre-empt childbirth (see Appendix A.5). Thus, the EI structure is an important factor in determining the discontinuous increase in the probability of female self-employment with childbirth; in particular, in households where the male spouse is self-employed. It may also be an important factor in determining the gradual rise in parental self-employment after childbirth, as parents may wish to re-enter the wage-paying sector before a second child is born.

For women married to non-self-employed men the increase in self-employment associated with childbirth is around 3% but declines after year 8 to 1.5% by year 15. Paid employment in these households increases faster after year 8. This suggests that for some mothers *individual* self-employment plays a temporary employment role during years of intensive childcare for some mothers, but in the long-run paid employment is preferred.⁵⁵ The fact that maternal self-employment does not decline with childbirth, as with paid employment, speaks to importance of flexibility as a source of resilience in the labour market (Goldin, 2014; Jeon and Ostrovsky, 2019).

In Appendix A.5 I provide a more detailed discussion on the regional variation in the selection into self-employment and co-employment with childbirth. In parallel analysis of

⁵³This coverage is up to a maximum threshold. Under QPIP, Quebec's system is more generous.

⁵⁴Saskatchewan moved from 18 to 30 weeks in 1996, while Alberta jumped from 18 to 52 weeks in 2001.

⁵⁵In Appendix A.6 I provide evidence that *individually* self-employed mothers earn an average hourly wage that is 80% less than the comparable wage-employed mother. In comparison, this gap does not exist in joint self-employed households. Thus, maternal self-employment is associated with a large compensating wage variation, which likely explains the return to the wage-paying sector.

male self-employment using the LFS, I show that even after accounting for individual and labour market characteristics there is a distinct East-West pattern in Canada: men in British Columbia and Alberta are more likely to be self-employed after childbirth, conditional on observed characteristics. Both the parent penalty and relative income literatures have established a correlation between their associated labour market outcomes and gender norms across geographic regions (Bertrand et al., 2015; Kleven et al., 2019a). Here, I find that the self-reported importance of work and flexibility across regions is more consistent with the observed selection into self-employment with family formation. However, these are preliminary results and require further investigation.

2.4.2 On the *real* nature of co-employment after childbirth

As the identification of self-employment in the LAD is based on income tax filings, it is certainly possible that the observed selection into joint self-employment is partially an accounting phenomenon. In jointly self-employed households only one individual need be self-employed while the other parent may indeed be economically inactive (or a paid employee outside of the home). This joint self-employment response would then be a tax avoidance response; consistent with the literature on income splitting (LaLumia, 2008; Schuetze, 2006). More generally, we know from the literature that self-employed workers more readily engage in tax avoidance behaviour (Kleven and Waseem, 2013) and, in Canada, there exists strong evidence that incorporated business owners will even allocate dividends to children (Bauer et al., 2015). Here, I contend that in the context of family formation this is not the case. At the extensive margin, the reported increase in joint self-employment is a real labour supply phenomenon and is best described as a co-employment response: employment within the same firm. While income splitting is indeed a characteristic of these households, it is a matter of compensation and not employment.

To make this argument, I execute a parallel event-study of initial childbirth using self-reported class of employment data from the SLID (1996-2010). The SLID matches self-reported labour supply information with annual income tax records. It also attaches a self-reported class of employment status to each job held in the last year. Together with information provided on the year of first birth, I can replicate the above ESD. As discussed in Section 2.3, the LAD's proxy for self-employment may identify wage-employment within the business of a self-employed spouse. I can identify such potential cases in the SLID through the identification of paid employees who have the same employer characteristics as their self-employed spouse (regardless of gender), as well as individual's who are 'unpaid family workers' married to a self-employed spouse.⁵⁶ This gives me two definitions of

⁵⁶While the SLID does not contain a unique employer ID with which to identify co-employment, I am able to match on four employer characteristics: industry (3-digit), sector (public-private), number of employees

self-employment in the SLID: an adjusted and unadjusted definition, where the former includes a proxy for co-employed spouses.

Figure 2.8 compares the event-study coefficients from the LAD (1988-2016) and SLID (1996-2010) for married men and women. It is evident that my identification of self-employment in the LAD does a good job of capturing the change in self-employment around childbirth that corresponds to a change in the self-reported status of an individual's main job.⁵⁷ The coefficients virtually overlap one another. This also true for the increase in joint self-employment when the adjustment is made for co-employment (see Figure 2.9). Without the adjustment, the LAD estimates over state the rise in joint self-employment. This difference between the adjusted and unadjusted series is made starker by a comparison of the conditional probability of maternal self-employment in Figure 2.10. The adjustment for co-employment is driven by a change in the probability of industry-matching with childbirth. Over the lifecycle, industry-matching among two-earner couples increases from 16% to 20% in Canada. This 4 %-point increase is explained entirely by self-employed households (see Figure 2.11). These event-study estimates suggest that 75% of this increase can be explained by childbirth. These results confirm that the event-study analysis of the LAD's proxy for self-employment captures a real change in the self-reported status of parental labour supply. It is therefore not a tax reporting phenomenon.

Using the SLID's additional information on the intensive margin, I show that income splitting is a matter of compensation and not labour supply. In addition, the rise in maternal self-employment (and co-employment) corresponds to a proportional increase in the likelihood of working from home among employed mothers. Table 2.1 depicts the event-time change in the intensive margin outcomes of annuals weeks worked and (log of) average weekly hours change among employed women, conditional on the self-employment status of the spouse. Similarly, Table 2.2 looks at the (log of) average hourly (taxable) market income and the likelihood of working from home. The coefficients for hours and working from home are also plotted in Figure 2.12. The data on weeks and hours of work refute the notion that women married to self-employed men disproportionately decrease their labour supply at the intensive margin with childbirth. In fact, the evidence suggests they work more weeks in the year *after* childbirth.⁵⁸ Hours show no difference but their total annual taxable income per hour is around 15-20% higher. In Appendix A.6 I provide a more de-

(categorical), and number of firm locations (categorical). Almost all unpaid family workers are married to a self-employed spouse. I find that these adjustments largely affect the self-reported class of employment of women. I make this adjust for the individual's main job only.

⁵⁷In the case of self-employment this is particularly important as a large share of individuals report some self-employment income despite being a paid employee. Similarly, many have some self-employment work each year, while their main job is paid employment. In the SLID (1996-2010) 15.6% of adults had some self-employment work, while only 12.8% of employed individuals were self-employed in their main job.

⁵⁸The level difference is negative before childbirth.

tailed discussion of the differences across paid employee, self-employed, and co-employed mothers. A striking result is that conditional on covariates, wages of co-employed mothers are not disproportionately higher than paid-employees. In contrast, the effective wage of self-employed mothers is significantly lower.

A defining characteristic of these self-employed households is that employed mothers are significantly more likely to work from home after childbirth. In fact, the increase in working from home in both types of households (based on the self-employment status of the spouse) is directly proportional to the increase in self-employment. Five years from initial childbirth it increases by 3.8% in households where the spouse is not self-employed, and by 11% in household where he is (see Table 2.2).⁵⁹ Together with the results on wages, it appears that co-employed mothers enjoy the same non-pecuniary benefits of self-employment, while avoiding the compensating wage variation. Moreover, if income splitting is a characteristic of self-employed households, it likely concerns compensation and not the extensive margin of employment. In what follows I provide evidence that self-employed households do, indeed, lower their taxes through income splitting after childbirth.

2.4.3 Relative income and tax savings from self-employment

One of the many distinguishing factors between paid employees and self-employed workers is the self reported nature of their taxable income. While the incomes of a wage/salaried workers are typically reported by a third party (employer), self-employed workers must self report their income to authorities. The literature has long acknowledged both the measurement and public policy challenges this creates (Hurst et al., 2014). For instance, it is hard to know if observed differences between wage and self-employed earnings identify compensating wage variation, changes to reporting (e.g. income shifting; le Maire and Schjerning, 2013; Slemrod, 2007) or pure tax avoidance - revenue under-reporting, cost inflation, income splitting. The splitting of business income between individuals within an economic family is a particular form of tax avoidance that is beneficial only under an individual tax structure (Stephens and Ward-Batts, 2004). In Appendix A.8, I provide an overview of the legal and tax institutions governing the value of income splitting in Canada. Here, I examine how the average tax rate of a married couple - as observed in the data - varies with childbirth, conditional on the self-employment status of the male spouse and the total income of the household.

Figure 2.13 plots the relative income distribution of married women in two-earner

⁵⁹This consistent with the fact that 43% of employed, married women (aged 20-54) who work from home in their main job report being self-employed, and 63% of self-employed women work from home. In households where couples are jointly self-employed, the likelihood that the mother works from home is around 60%. (Own calculations based on SLID (1996-2010). Excludes any adjustment for (un)paid employment in a household firm.)

Canadian households, by the self-employment status of their spouse.⁶⁰ The estimates are based on the pooled SLID and CIS cross-sections, covering the years 1996-2016. I implement the same methodology as [Bertrand et al. \(2015\)](#); estimating the LOWESS smoother separately on each side of the 0.50 divide. In support of [Zinovyeva and Tverdostup \(2018\)](#), I find strong evidence that any break in the relative income distribution at 0.50 is explained by the presence of self-employed households.⁶¹

This difference in relative income distribution provides self-employed households with significant tax savings. In the LAD, I find that the average tax rate of self-employed households (where the male spouse is self-employed) falls by an additional 3.5% after childbirth, relative to non-self-employed households (see Figure 2.14).⁶² There is no evidence of a pre-trend, and the tax differences remain stable past event-time 5; as with the conditional increase in maternal self-employment (Figure 2.5). This decline in average tax rate corresponds to a smaller decline in the female share of taxable income in self-employed households: by 3% in event-time 5. As in Figure 2.5, these estimates are based on an interaction with contemporaneous male self-employment status. Selection on the income profile of households where the male spouse is self-employed may therefore explain these differences. For this reason, I include a specification that flexibly controls for total income using a ten-decile spline in the log of total taxable income. Differences in total incomes explain the level difference between paid and self-employed households in the base period (not depicted in the figures), but not the change in relative earnings with childbirth. As you would expect with a progressive tax structure, total income explains some of the drop in average taxes across both types of household but not the difference between paid and self-employed households.

Together, these event-study-design results suggest that the event of initial childbirth is associated with a significant increase in the share of co-employed households. It is not just that female self-employment increases in households where a male spouse is already self-employed, as men are shown to select out of the wage-paying sector with childbirth. This suggests that co-employment is a joint, and not conditional (female), response to childbirth: there is a subpopulation of households that jointly select out of the wage-paying sector

⁶⁰I use the SLID in this application to avoid any measurement error associated with my proxy in the LAD. In this instance the spouse is self-employed if he self-reports any self-employment work in the past year. This is therefore a broader category than those whose main job is self-employed. However, identifying an individual's main job is subjective, and this category can be more easily identified in both the SLID and CIS surveys.

⁶¹Even in wage-employed households the relative distribution of total taxable income is more equal than that of employment income, reflecting the fact that there is an incentive to allocate non-employment income in a tax-minimizing manner (see Figure 2.13 in Appendix A.7). For example, a household may choose to report rental income under the name of the individual with less employment income. The asymmetric shock of childbirth will only increase the gains from non-employment income reallocation; demonstrated here by the difference between wage-employed households with and without children.

⁶²This tax wedge is proportional to the simulated savings from income splitting after childbirth (see Figure 2.16).

with family formation. In the remainder of the paper, I show that savings from income splitting can account for half of the increase in co-employed couples in Canada. In this way, I establish that individual tax structures increase the level of male self-employment and co-employment in the labour market.⁶³

2.5 Tax incentives for co-employment

In this section, I develop a simulated-instrument research design to identify the relationship between parental self-employment and income-splitting tax savings. This path of enquiry faces several empirical challenges, the foremost being that income splitting is unobserved. However, I show that despite this limitation one can still estimate a plausible reduced form elasticity and test the hypothesis that mothers and fathers respond equally to exogenous shocks in the unobserved value of income splitting. I show that the savings from income splitting do indeed increase paternal self-employment, even before childbirth. I estimate a *reduced form* elasticity of 0.5 that is equal across mothers and fathers. With a simulated savings of $\sim 3\%$, consistent with the estimated decline in average tax rate (Figure 2.14), the estimated *reduced form* elasticity of ~ 0.5 suggests that the income splitting could increase paternal self-employment and co-employment by 1.5%. This is shown to be the difference between the male self-employment ‘family formation gap’ in Canada and the US (Table 2.4).

Section 2.5.1 begins by defining that value of income splitting as $\tilde{\Delta}_{pt}$: a function that depends on the tax structure in a given province (p) and year (t). As a potential instrument for this endogenous variable, I provide a plausible simulated equivalent Δ_{pt} . Unfortunately, neither of these values is observable as they depend on the unobserved self-employment earnings of each household. I will therefore introduce a third value Δ^S which applies the simulated functional form of Δ_{pt} to a projected earnings profile of each household based on pre-childbirth earnings in the wage-paying sector. Δ^S will then be used to estimate a reduced form income-splitting elasticity within the same event-study setting described by Equation 2.1. As $\tilde{\Delta}_{pt}$ is unobserved, Δ^S cannot be applied as a simulated *instrument* in the traditional sense. One can only estimate the reduced form equation: the relationship between self-employment and Δ^S .

Given the strong exclusion-restriction assumption required to identify the reduced form elasticity using Δ^S , as well as other omitted variable concerns, I introduce a simulated counterfactual Δ^C . By including Δ^C as a control function in the reduced form equation I

⁶³Note, I intentionally do not refer to this as an increase in female self-employment given the broader definition of self-employment in the LAD, which disproportionately affects the measurement of female self-employment.

can estimate the reduced-form elasticity using exogenous variation in Δ^S that arises from changes in the tax structure. Note, Δ^C is not an instrument for Δ^S , but rather a control function. This is because in Section 2.5.2, I apply Δ^S as an instrument for paternal self-employment in an equation relating maternal self-employment to paternal self-employment with childbirth.⁶⁴ As will be explained, this IV estimator provides an explicit test for the income-splitting mechanism.

2.5.1 Identification

There are numerous challenges in identifying the relationship between income splitting and selection into self-employment around childbirth. To assist in understanding the source of these challenges it will help to have in mind a simple Roy-model empirical framework. Suppose two-earner households select between joint employment in the paid and self-employed labour markets based on differences in the value of employment in each sector. Let V_{its}^w be the value of joint employment in the paid labour market and V_{its}^q be the value of joint self-employment (co-employment) for household i in period t and event-time s . To simplify notation, I focus here on the case of joint employment in both sectors. Appendix A.1 provides a more complete model of sectoral choice, including non-employment.

$$V_{its}^q - V_{its}^w = \gamma \left[\ln \left(y_{its}^1 + y_{its}^2 - T_{pt}(y_{its}^1) - T_{pt}(y_{its}^2) \right) - \ln \left(z_{its}^1 + z_{its}^2 - T_{pt}(z_{its}^1) - T_{pt}(z_{its}^2) \right) \right] \\ + g(X_{its}, \theta_{its}, t, s) + \underbrace{\sigma \tilde{\Delta}_{pt}(y_{its}^1, y_{its}^2, \tilde{\delta}_{its})}_{\text{income splitting}}$$

In this additively separable utility specification, households concern themselves with the (log) difference in after-tax earnings in each sector. Earnings in the paid and self-employed labour markets are denoted by z_{its} and y_{its} respectively. In addition to earnings, households may select into joint self-employment based on the relative non-pecuniary benefits of employment in each sector, denoted here by the function $g(\cdot)$. These non-pecuniary factors will depend on the characteristics of the household (X_{its} , including age and geography), idiosyncratic preferences (θ_{its}), and factors that vary directly with calendar time (or real business cycle, t) and event-time (s). The direct inclusion of event-time (s) demonstrates the importance of childbirth in determining the relative non-pecuniary benefits of co-employment. For example, we expect the relative flexibility benefits of co-employment to increase discontinuously with family formation.

Crucially, under an individual tax structure self-employed households may benefit from

⁶⁴I therefore avoid having an instrument for an instrument. The approach should be equivalent given the well-established equivalence of control functions and instrumental variables.

income splitting. Denoted here by the value $\tilde{\Delta}_{pt}(y_{its}^1, y_{its}^2, \tilde{\delta}_{its})$. I allow the coefficient on income splitting (σ) to vary from that on after-tax income (γ) as households may place a discount (or premium) on the value of savings generated through income splitting. As these benefits are exclusive to self-employed households, the savings they generate are a function of self-employment earnings only. The value of income splitting can formally be defined as the increase in after tax income as a result of income reallocation in the household,

$$\begin{aligned} \tilde{\Delta}_{pt}(y_{its}^1, y_{its}^2, \tilde{\delta}_{its}) &= \ln(y_{its}^1 + y_{its}^2 - T_{pt}(y_{its}^1 - \tilde{\delta}_{its}) - T_{pt}(y_{its}^2 + \tilde{\delta}_{its})) \\ &\quad - \ln(y_{its}^1 + y_{its}^2 - T_{pt}(y_{its}^1) - T_{pt}(y_{its}^2)) \end{aligned}$$

where $\tilde{\delta}_{its}$ is an endogenous transfer between earners for the purposes of lower their average tax rate. The value $\tilde{\Delta}$ represents the proportional change in after tax earnings from income splitting and takes on a similar structural form to a reduced form income effect modelled in [Gruber and Saez \(2002\)](#). The Δ -function depends on calendar year (t) and province (p) as it depends on the time-dependent tax structure in a given province.

Consider that as you move through event of childbirth (from event-time $0 \rightarrow s$) the value of income splitting changes along two dimensions.⁶⁵

$$\begin{aligned} &\tilde{\Delta}_{pt}(y_{its}^1, y_{its}^2, \tilde{\delta}_{its}) - \tilde{\Delta}_{p,t-s}(y_{it-s,0}^1, y_{it-s,0}^2, \tilde{\delta}_{it-s,0}) \\ &= \underbrace{\tilde{\Delta}_{pt}(y_{its}^1, y_{its}^2, \tilde{\delta}_{its}) - \tilde{\Delta}_{p,t-s}(y_{its}^1, y_{its}^2, \tilde{\delta}_{its})}_{\text{tax structure}} \\ &\quad + \underbrace{\tilde{\Delta}_{p,t-s}(y_{its}^1, y_{its}^2, \tilde{\delta}_{its}) - \tilde{\Delta}_{p,t-s}(y_{it-s,0}^1, y_{it-s,0}^2, \tilde{\delta}_{it-s,0})}_{\text{income component}} \end{aligned} \tag{2.2}$$

First, $\tilde{\Delta}$ will vary with changes to the income profile of the household. No doubt, this latter ‘income component’ changes dramatically during the event of childbirth, and likely increases the value of income splitting. This is arguably the main driver behind the potential relationship between income splitting and selection into self-employment with childbirth. The implication then is that the value of income splitting is endogenous to labour supply, and in particular, changes to labour supply with childbirth.

Second, the value of $\tilde{\Delta}$ will vary depending on tax structure. Between periods in which the tax structure remains constant this term will be zero, but after a reform may be negative or positive depending on the nature of the tax changes and how they interact with the

⁶⁵ Assuming that other couple specific characteristics that might determine the ‘take up’ of income splitting remain constant. For example, the perceived risk of being audited by the tax authority.

income profile of the household. It is this complex interaction between the tax structure, income profile of the household, and event of childbirth that will allow me to identify the relationship between self-employment and income splitting, apart from childbirth.

In Appendix A.8, I provide a more detailed discussion of the various income tax reforms which have changed the value of income splitting for different households between 1988 and 2016. To summarize the cumulative effect of these reforms - both federal and provincial - I plot the difference between the after-tax income of two households with the same total income (\$80,000 in 2016 CAD), but different relative incomes (see Figure 2.15). This difference represents the value of income splitting in the household with a 65-35 income profile. It is evident that the level differs by province, and that there is considerable variation across time.

We can reframe our path of enquiry in terms of the reduced form relationship between the discrete outcome of self-employment and the value of income splitting, conditional on event-time. To simplify the notation, I absorb the set of age-dummy variables in [Kleven et al. \(2019a\)](#)'s event-study framework into the set of individual covariates (X_{its}).

$$Q_{its}^g = \sum_{j \neq -1} \alpha_j \cdot \mathbf{1}\{j = s\} + \sum_{j \neq -1} \tilde{\beta}_j^g \tilde{\Delta}_{pt} \left(y_{its}^1, y_{its}^2, \tilde{\delta}_{its} \right) \cdot \mathbf{1}\{j = s\} + X'_{its} \gamma^g + \delta_t + \tilde{\varepsilon}_{its}^g \quad (2.3)$$

$g = 1, 2$

In Equation 2.3, the stand-alone event-time dummies capture the base level increase in self-employment associated with childbirth because of unobserved (non-pecuniary) factors. As the income-splitting elasticity may vary with event-time, I deliberately interact the value of income splitting with event-time. The error term in equation 2.3 includes both unobserved preference factors and sectoral earnings differentials that matter for the self-employment decision but are unobserved to the researcher.

If income splitting is an important mechanism through which the event of family formation leads to an increase in co-employment we should find that,

$$H_0 : \tilde{\beta}_j^1 = \tilde{\beta}_j^2 > 0 \quad \forall j$$

That is, the paternal and maternal elasticities are equal. This simple hypothesis is non-trivial. A symmetric response would suggest that income splitting coordinates the selection of married couples into co-employment. In addition, the dynamics of this relationship are a consideration. As suggested by the event-study-design, the timing of paternal and maternal responses to the event of childbirth may differ, with male self-employment decisions pre-empting childbirth.

The endogeneity $\tilde{\Delta}$ is in part mitigated by the fact that $\tilde{\delta}_{its}$ is unobserved. In the LAD, I observe $\{y_{its}^1 - \tilde{\delta}_{its}, y_{its}^2 + \tilde{\delta}_{its}\}$ for self-employed households, while for paid employees, these earnings reflect an unobserved counterfactual. Suppose for a moment that $\tilde{\delta}_{its}$ was observed. The value of these savings would be endogenous. Not only because of self-selected ‘take up’, but also because they are highly dependent on $\{y_{its}^1, y_{its}^2\} = \{\mu_{its}, y_{its}\}$: the relative and absolute earnings of the couple.⁶⁶ These are undeniably correlated with unobserved factors that may themselves determine selection into self-employment and co-employment around childbirth, including the unobserved sectoral earnings differential.

One solution to the endogenous take up of income splitting would be to use a simulated instrument of income-splitting tax savings (Gruber and Cullen, 1996). In line with the evidence that self-employed households bunch at 0.5, I propose a *potential* savings function equal to the proportional difference between the after-tax income of the household under a joint and individual tax structure. This happens to be the transfer that maximizes the household’s after-tax income under a progressive tax structure.⁶⁷ Note, I introduce province (p) as an additional subscript to Δ at this point to highlight the fact that the tax structure varies both with province and time.

$$\Delta_{pt}(y_{its}^1, y_{its}^2) = \underbrace{\ln(y_{its}^1 + y_{its}^2 - 2T_{pt}(0.5(y_{its}^1 + y_{its}^2)))}_{\text{joint structure}} - \underbrace{\ln(y_{its}^1 + y_{its}^2 - T_{pt}(y_{its}^1) - T_{pt}(y_{its}^2))}_{\text{individual structure}}$$

As $\tilde{\delta}_{its}$ is unobserved, it is not possible to execute the simulated instrument research design. If $\{y_{it}^1, y_{it}^2\}$ was observed across all households, the implicit reduced form of the simulated IV would be feasible were. In the absence of observing this income path, the next best solution is to substitute in a predicted path: $\{\hat{z}_{its}^1, \hat{z}_{its}^2\}$. The switch here to z_{its}^g is deliberate as it denotes the fact that self-employment earnings must be predicted based on earnings in the wage-paying sector.⁶⁸

Were this a paper on selection into self-employment at single point in time we could simulate $\Delta_{pt}(z_{it-1}^1, z_{it-1}^2)$ - the savings based on wage-earnings in the previous period.

⁶⁶We could write $\tilde{\Delta}_{pt}(y_{its}^1, y_{its}^2, \tilde{\delta}_{its}) = \Delta_{pt}(\mu_{its}, y_{its}, \tilde{\delta}_{its})$, where μ is the relative income of earners in the household, and y the total income of the couple. The value of Δ therefore depends on both the absolute and relative incomes within the household. This makes the following exercise more difficult than the standard individual tax response exercise (for example, Gruber and Saez, 2002).

⁶⁷Note, the transfer implicit in the above savings may be larger than the minimum transfer required to minimize the households tax liability. With a stepped income tax structure, the minimum required transfer need not split the couple’s income equally to equalize a couple’s marginal tax rates. Regardless, under a progressive tax structure, the household’s tax liability will be minimized at the equal split.

⁶⁸Observed self-employment earnings are endogenous to income splitting. For this reason, an individual’s earnings in the wage-paying sector will be assumed to be the best predictor of their productivity in self-employment. This must be the case for both the absolute and relative productivity of a married couple.

However, this would be inadequate for equation 2.3 which differs in two important respects: the model estimates selection into self-employment over a much longer period and the relevant time (subscript) is event-time (s), not calendar time (t). Naturally, one could then consider $\Delta_{pt}(z_{it-s-1,-1}^1, z_{it-s-1,-1}^2)$ - the savings based on wage-earnings in the period before childbirth (and possibly adjust for inflation and real wage-growth). This too would be insufficient, as it would not capture the increase in savings associated with the highly unequal event of childbirth. The predicted income path of the household must capture both real wage-growth, as well as the ‘shock’ of childbirth to labour market earnings. To achieve this outcome, I adopt a two-stage methodology and first predict,

$$\{\hat{z}_{its}^1, \hat{z}_{its}^2\} = \{\hat{a}_s^1 z_{i,t-s-2,-2}^1, \hat{a}_s^2 z_{i,t-s-2,-2}^2\}$$

where

$$a_s^g = \frac{E[z_{its}^g | X_{its}, t, s]}{E[z_{it,-2}^g | X_{its}, t, s = -2]}$$

The event-time coefficients (a_s^g) are estimated in first stage event-study design using the outcome of (level) taxable income. For fathers, the event of childbirth has a negligible affect on the predicted path of earnings, which accounts for the age-profile of the individual and includes an adjustment for inflation. In contrast, mothers face a persistent shock to real income in the order of 30%. To identify the income-splitting elasticity before and after childbirth, I shift the base period from event-time -1 to -2. As such, the analysis will be limited to households where neither spouse is self-employed two years before their first child is born, with the result that the reduced-form elasticity is unidentified in this based period ($s = -2$). This yields a simulated vale of income-splitting savings $\Delta^S = \Delta_{pt}(\hat{z}_{its}^1, \hat{z}_{its}^2)$ and a feasible reduced form specification,⁶⁹

$$Q_{its}^g = \sum_{j \neq -1} \alpha_j \cdot \mathbf{1}\{j = s\} + \sum_{j \neq -2} \beta_j^g \underbrace{\Delta^S \cdot \mathbf{1}\{j = s\}}_{\text{simulated savings}} + X'_{its} \gamma^g + \delta_t + \delta_p + \epsilon_{its}^g \quad (2.4)$$

$s \neq -2 \quad \text{and} \quad g = 1, 2$

While $\beta_j^g \neq \tilde{\beta}_j^g$ from Equation 2.3, the fact that Δ^S is common within a couple means that,⁷⁰

⁶⁹In practice, Δ^S is simulated using Milligan (2016)’s Canadian Tax and Credit Simulator (CTaCS).

⁷⁰The unobserved first stage from the simulated instrument would be the same for husband and wife. As such, the reduced form of the IV would yield the coefficient $\beta_j^g = \rho \tilde{\beta}_j^g$. Where ρ is the coefficient from the first stage. Neither the value nor sign of ρ are relevant to the test, only that $\rho \neq 0$.

$$H_0 : \beta_j^1 = \beta_j^2 \iff H_0 : \tilde{\beta}_j^1 = \tilde{\beta}_j^2$$

Thus, the unobserved nature of $\tilde{\delta}_{its}$ does not prevent me from testing the hypothesis that the *true* income-splitting elasticity ($\tilde{\beta}_j^g$) is symmetric within married households at the point of family formation.

Under what assumptions are the reduced form elasticities (β_j^g) in equation 2.4 identified? As it stands, the relative and absolute income of the couple affect the discrete choice of self-employment only through the value of Δ^S .⁷¹ This is strong exclusionary restriction as there may be a few direct channels through which a couple's income profile informs selection into self-employment, even joint self-employment. For one, the presence of a high earning spouse may provide insurance or investment capital for a lower earning spouse to engage in entrepreneurial activity. I could relax this assumption by including a sufficiently flexible function of income, $g(\hat{z}_{its}^1, \hat{z}_{its}^2)$, that when interacted with event-time captures any direct relationship between a household's income profile and selection into self-employment. Doing so would require sufficient variation in the tax structure over time and across provinces to ensure that β_j^g is still identified.

In this paper I adopt a different, control function, approach. I use a counterfactual value of Δ to isolate the variation in Δ^S that arises solely from changes in the tax structure. Recall equation 2.2, one can decompose the change in the value of income splitting with event-time using the counterfactual $\Delta^C = \Delta_{p,t-s-2}(\hat{z}_{its}^1, \hat{z}_{its}^2)$: the value of income splitting were the tax structure to have remained unchanged since the base period before childbirth. The difference between the 'realized' value of income splitting and this counterfactual represents a tax 'innovation' in the value of income splitting. I could use this difference - $\Delta^S - \Delta^C$ - as an instrument for the value of Δ^S , or adopt the equivalent control-function approach and include Δ^C as an additional control in the model. Here, I adopt the latter control-function approach, because in Section 2.5.2 I use Δ^S as an instrument for spousal self-employment (Q_{its}^1). The control-function can be more readily applied in this setting, whereas an instrumental variable approach at this stage would lead to a double instrument in Section 2.5.2.

$$Q_{its}^g = \sum_{j \neq -1} \alpha_j \cdot \mathbf{1}\{j = s\} + \sum_{j \neq -2} \beta_j^g \underbrace{\Delta^S \cdot \mathbf{1}\{j = s\}}_{\text{simulated savings}} + \sum_{j \neq -2} \mu_j^g \underbrace{\Delta^C \cdot \mathbf{1}\{j = s\}}_{\text{control function}} \quad (2.5)$$

$$+ X'_{its} \gamma^g + \delta_t + \delta_p + v_{its}^g \quad s \neq -2 \quad \text{and} \quad g = 1, 2$$

⁷¹A secondary consideration is that the LAD contains few individual level covariates, which may be a source of omitted variable bias. For example, I do not observe the education level, industry, and occupation of the individual. These likely determine both the income profile and the self-employment decision of the household.

Once again, β_{-2}^g is unidentified as the control function and ‘realized’ simulated savings are equal in the base period. In equation 2.5, the identifying variation now arises from exogenous variation in the value of income splitting as the income-profile remains constant across both simulations.⁷² In particular, given the additional controls for province and year fixed effects, it is the tax variation across birth cohorts that identifies the β_j^g ’s. As an additional identifying assumption it must be that these perturbations in the value of income splitting remain uncorrelated with the counterfactual value.⁷³ Conditional on a more parametric control for the income profile of the household, this is not too onerous an assumption.⁷⁴

There is a strong intuition to this identification strategy. From the perspective of a wage-employed couple in the base period ($s = -2$) the value of Δ^C represents a predictable path of tax savings, given an expected shock to the mother’s income. However, any realized deviation in that value, resulting from changes in the tax structure is unpredictable. Figure 2.16 depicts how both the realized and counterfactual values of Δ evolved for six simulated birth cohorts. Ontario households with a pre-birth income ratio of 45-55 and total earnings of \$80,000’s (2016 CAD) experienced positive shocks if the had their first child in the early to mid 1990’s. Beginning in the late 1990’s birth cohorts received negative shocks, while from the mid 2000’s cohorts experienced insignificant changes.

For this reason, the research design has the added benefit of mitigating concerns of reverse causality between male self-employment the timing of childbirth. By incorporating of the asymmetric projection of childbirth into the projected income path, this potential source of endogeneity is incorporated into the simulation. Apart from event-time -1, this research design ensures that the β_j^g ’s are identified off variation that arises after the event of childbirth and therefore cannot be associated with the timing of childbirth.⁷⁵ This ensures that the estimated elasticity represents a mechanism in which income splitting determines labour supply decisions conditional on event-time, and not one in which it directly affects the timing of childbirth itself.

⁷²This is not far from the identification strategy used by Gruber and Saez (2002) to account for the endogeneity of the income effect. They use a counterfactual measure of the income effect (holding income constant, while varying the tax structure) as an instrument. In this way they isolate the variation in the change in after tax income that arises from a change in the tax structure alone. In this application, to isolate that same variation one requires a control function instead of an instrument.

⁷³That is,

$$\Delta_{pt} \left(\hat{z}_{its}^1, \hat{z}_{its}^2 \right) = \Delta_{t-s-2} \left(\hat{z}_{its}^1, \hat{z}_{its}^2 \right) + \omega_{its} \quad \text{where} \quad E \left[\Delta_{t-s-2} \left(\hat{z}_{its}^1, \hat{z}_{its}^2 \right) \omega_{its} \right] = 0$$

⁷⁴In practice, I include $\{\ln(\hat{z}_{its}^1), \ln(\hat{z}_{its}^2)\}$ as additional terms in the model. I find allowing for too much flexibility in the income control function removes what limited identifying variation there is in the income-splitting tax shocks.

⁷⁵For event-time -1 I need to assume that the probability of giving birth in the next period is not correlated with a tax increase in the value of income splitting.

2.5.2 Testing income-splitting mechanism

The symmetry hypothesis can be implemented using an IV estimator. I reframe the null hypothesis that the reduced form elasticity is equal across mother and father as the hypothesis that the ratio of the two coefficients is equal to 1. This ratio is given by the IV estimator.

$$H_0 : \beta_j^1 = \beta_j^2 \Rightarrow H_0 : \frac{\beta_j^1}{\beta_j^2} = \eta_j^{IV} = 1$$

The implied second stage equation models maternal self-employment as a function of paternal self-employment, interacted with event-time. In Equation 2.6 the coefficient η_j denotes the difference in selection into self-employment relative to the based period, for women married to self-employed men.⁷⁶

$$Q_{its}^2 = \sum_{j \neq -2} \alpha_j \cdot \mathbf{1}\{j = s\} + \eta Q_{its}^1 + \sum_{j \neq -2} \eta_j Q_{its}^1 \times \mathbf{1}\{j = s\} + X'_{its} \gamma + \delta_t + \delta_p + \varepsilon_{its} \quad (2.6)$$

When implemented as an IV estimator, each interaction $Q_{its}^1 \times \mathbf{1}\{j = s\}$ is instrumented for using the corresponding simulated instrument interaction $\Delta^S \cdot \mathbf{1}\{j = s\}$, while the control function - $\Delta^C \cdot \mathbf{1}\{j = s\}$ - remains in the second stage. As Δ^S and Δ^C are equal in the base period there is no instrument for ηQ_{its}^1 : the indicator for spousal self-employment in the base period. Once again, there is no loss here, as the above identification strategy already requires selection on couples who were not self-employed in the base period.

This set of instruments does not fulfil the exclusion restriction required to interpret the resulting coefficient as a Local Average Treatment Effect. Nevertheless, the intuition behind the LATE interpretation fits the null hypothesis being tested. For a household to benefit from income splitting it must be that both individuals report some self-employment income. Therefore, the marginal father who is pushed into self-employment by a shock to the value of income splitting will only benefit from having done so if his spouse follows him into self-employment. Thus, under an income-splitting mechanism we expect a LATE that is equal to 1.

Following the initial step, in which I predict a path of income for each couple based on taxable income in the base period, I simulate the contemporaneous and counterfactual

⁷⁶This is in essence an unidentified difference-in-difference coefficient, where women married to wage-employed men are the control group.

$$\eta_j = \left[E[Q_{itj}^2 | Q_{itj}^1 = 1, X_{itj}] - E[Q_{i,t-j-2,-2}^2 | Q_{i,t-j-2,-2}^1 = 1, X_{i,t-j-2,-2}] \right] \\ - \left[E[Q_{itj}^2 | Q_{itj}^1 = 0, X_{itj}] - E[Q_{i,t-j-2,-2}^2 | Q_{i,t-j-2,-2}^1 = 0, X_{i,t-j-2,-2}] \right]$$

Note, in this instance I do not interact the self-employment status of the spouse with the age-dummy variables.

value of Δ . The latter incorporates adjustments made for inflation.⁷⁷ I restrict the estimating sample to couples who have their first child between 1990-2010, and include event-times -4 through 6 in the sample. In addition, I select on couples where neither spouse is self-employed in the base period and the male spouse is paid employee.⁷⁸ For now, I report the results with standard errors clustered at the individual/couple level.⁷⁹

2.5.3 Results

Panel A of Figure 2.17 plots the paternal ('first-stage') and maternal ('reduced form') reduced form income-splitting elasticities corresponding to Equation 2.5 (see also Table 2.3). Prior to the base period the paternal elasticity is not significantly different from zero. This provides a placebo test for the exogeneity of the underlying policy variation. The estimated elasticities are stable across event-time, with $\beta_j^1 \approx 0.5$. The elasticities are even positive before childbirth, in accordance with the pre-emptive selection of spouses in self-employment.

The maternal elasticities show some evidence of a correlation in the pre-period. However, comparing the specifications in Table 2.3 we see that these point estimates are not robust to the additional various income controls. As with men, the maternal elasticity is stable across event-time, with $\beta_j^2 \approx 0.5$. In previous versions of this paper, the maternal elasticity was insignificant in the year before childbirth ($s = -1$). These earlier results were based on a later subsample of children born after 1992.⁸⁰ This subtle difference is import given the expansion of parental leave and job protected leave in 1991 (Baker and Milligan, 2008). As previously discussed, until 2011 the Federal employment insurance structure excluded self-employed workers. Protected leave and employment insurance in the wage-paying sector incentivize households to delay co-employment till after childbirth.

⁷⁷While Canada's federal and provincial income tax brackets have not always been pegged to inflation, the number of years being predicted here suggest that adjusting for inflation will be necessary.

⁷⁸In the sample selection I do not require that the female spouse is also a paid employee in the pre-period so as to account for changes in female labour supply over the sample period. In practice, the elasticities are slightly larger for the subsample including two-earner wage-paying sector couples. I make the additional income restriction that neither spouse earned more than half a million (2016 CAD) in the base period.

⁷⁹Block-bootstrapped standard errors, clustered at the individual level, have also been estimated and are available for these estimates. For the event-time-specific elasticities, the bootstrapped standard errors match the reported standard errors to the second, even the third, decimal place, with the result that they have no bearing on the statistical significance of these estimates. This is not surprising given the large sample size. For the IV estimates, the estimated block-bootstrapped standard errors are significantly larger. This is likely due to a weak instrument problem given the limited residual variation in Δ^S , conditional on Δ^C . With a weak instrument, a small first stage can lead to very large IV estimate, generating outliers in the bootstrap sample. This is consistent with the fact that the 95% bootstrapped confidence intervals are far closer to those depicted in Figure 2.17, while the standard errors are significantly larger. In general, inferences made using the 95% bootstrapped confidence intervals do not differ from those presented here. The bootstrapped confidence intervals do not reject an IV estimate of 1 but do also contain 0 for early and late event-time lags (i.e., event-times -1, 0, 5 and 6).

⁸⁰Future versions of the paper will included such a specification. Unfortunately, changes to the way I identify self-employed workers in this updated version mean that I cannot directly compare these results to earlier specifications.

Alberta and Saskatchewan did not immediately expand their job protected leave provisions in 1991, and it is in these two regions where one observes some pre-emptive selection into co-employment before childbirth (see Figure A.14 in Appendix A.5).

Recall that these coefficients are not the same as the elasticity coefficient $\tilde{\beta}_j^g$. Instead, $\beta_j^g = \tilde{\beta}_j^g \cdot \rho_j$ where ρ_j is the unobserved first-stage coefficient. As previously discussed, this does not affect the symmetry hypothesis but does limit my interpretation of the magnitude of the coefficient. For β_j to be an underestimate of the true income-splitting elasticity it must be that $\rho_j < 1$. This would suggest that on average the simulated savings should overestimate the true, unobserved savings. There are good reasons to think that this is the case. First, under a progressive tax structure,

$$\tilde{\Delta}_{pt} \left(y_{its}^1, y_{its}^2, \tilde{\delta}_{its} \right) \leq \Delta_{pt} \left(y_{its}^1, y_{its}^2 \right)$$

as the joint income tax structure minimizes the household's tax liability.⁸¹ Thus, simulated savings at the unobserved income levels will be at least great as the actual, unobserved savings. I require then that, on average, the simulated savings at the predicted income profile are greater or equal to those under the unobserved income profile.

$$E \left[\Delta_{pt} \left(y_{its}^1, y_{its}^2 \right) \right] \leq E \left[\Delta_{pt} \left(\hat{z}_{its}^1, \hat{z}_{its}^2 \right) \right]$$

There are two reasons to believe this is the case. First, the predicted path of income assumes a proportional change in earnings, while the realized change could be significantly smaller given that the evidence suggests these co-employed mothers do supply positive hours of work. Second, if selection into self-employment is associated with a short-term negative income shock (for example, a start up cost) the total income in the household will be lower, which will likely reduce the value of income splitting.⁸²

Empirically, I find that self-employed households reduce their taxes by an additional $\sim 3\%$ with childbirth (by event-time 5; see Figure 2.14). These savings are proportional to the simulated savings I have made suggesting that the unobserved first stage could be close to 1. The reduced form income-splitting elasticity of 0.5 would suggest that a tax savings of 3% from income splitting would generate an increase of 1.5% in male self-employment, and co-employment after childbirth. This would account for $\sim 50\%$ of the increase in

⁸¹Note, Canada's income tax structure is not strictly progressive. Under the CPP reform of 1998 the marginal income tax rate becomes regressive in parts (for self-employed workers). As such, income splitting can result in marginally higher taxes. However, these differences are small and for the majority the progressivity assumption still holds.

⁸²This does not hold for households with very high incomes as at the point where the total income of the household is more than double the top income tax bracket the dollar value of savings is fixed. The relative savings are declining in income at that point. For such households, the simulated savings will increase with a fall in total income.

co-employment with childbirth (see Figure 2.4), and about a third of the increase in either male or female self-employment between from event-time -5 and 5 (see Figure 2.1).

As benchmark, I re-examine the family-formation gap in male self-employment estimated using the Canadian LFS. This is the same methodology I apply in my discussion of regional variation in Appendix A.5. In Canada, employed men with children are on average 2.1% more likely to be self-employed relative to men without children, conditional on age and other covariates (see Table 2.4). Estimating the same specification using the US CPS, I estimate a family formation gap of only 0.66% (see Table 2.4). Pooling both surveys, the difference between Canada and the US's family formation gap is 1.6%, inline with the above prediction of 1.5% based on an elasticity of 0.5. As the US has a joint tax structure, this difference is consistent with the impact of income splitting.⁸³

Panel B of Figure 2.17 plots the sample OLS and IV estimates corresponding to Equation 2.6 and Table 2.5. The OLS model depicts the increase in the conditional probability of self-employment for women married to self-employed men. This probability is non-zero in the year before childbirth and in the pre-period.⁸⁴ When you instrument the spouse's self-employment status with the realised value of income splitting one estimates an IV estimate that is indistinguishable from 1. The confidence intervals on these estimates are fairly large, reflecting the limited variation in the simulated savings conditional on the control function. Nevertheless, the point estimates are all close to 1 and significantly different from zero (except for event-time -1). This confirms the symmetry of the labour supply response to income splitting and proves that these savings work to coordinate households into co-employment with childbirth.

2.6 Conclusion

Adjustments to female labour supply with childbirth are a primary concern for gender inequality (Bertrand et al., 2010; Kleven et al., 2019b). It is therefore of critical importance that we identify the various economic, institutional, and cultural drivers of parental labour supply at the point of family formation. This paper is the first to demonstrate that individual tax structures incentivize a more coordinated response to childbirth within married households. This strategic response is best characterized by a joint selection out of the wage-paying sector and into self-employment, where co-employed households enjoy both

⁸³Other policy differences may also play an important role in explaining this difference in Canadian-US male self-employment; in particular, differences in the provision of healthcare. In the US, the cost of private medical aid and its attachment to employment in the wage-paying sector, may dissuade households from becoming self-employed around family formation.

⁸⁴This could reflect measurement error in the self-employment proxy, which allowed for some couples who were self-employed in the base period to remain in the sample.

flexibility and tax benefits.

The subsidization of male self-employment and co-employment under an individual tax structure provides an important new insight into the nature of the parent penalty and the relationship between taxation and labour supply. For both men and women, selection into self-employment is an important margin through which parents balance the joint challenges of employment and childcare. While this is expected for women, it has not been for men, who demonstrate no extensive margin response to childbirth. Moreover, the endogeneity of this decision to the value of income splitting under a progressive, individual tax structure provides several new insights. It shows that individual tax structures are not neutral, at the point of family formation, and portrays income splitting as an incentive for, and not just a consequence of, self-employment. Beyond tax avoidance, income splitting is a subsidy to the creation of flexible, tax-optimizing family firms that provide stable, long-run employment to households after childbirth.

2.7 Tables and Figures

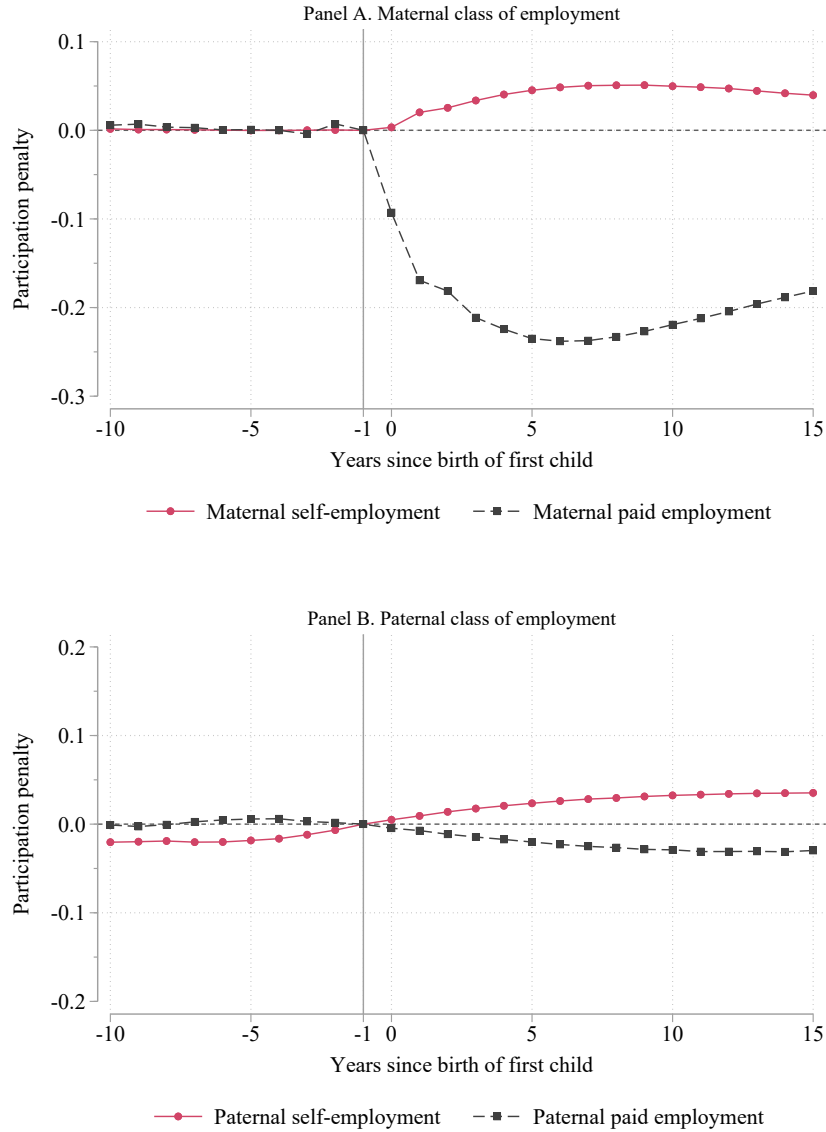


Figure 2.1: Participation parent penalty: class of employment

Note: Estimated using the Longitudinal Administrative Dataset (1988-2016). Sample includes all married couples where the female spouse is at least 20 years of age. Each data point corresponds to the rescaled event-time coefficient from an event-study-design (ESD) of initial childbirth. In each panel the wage-employment and self-employment curves, correspond to separate ESD models with the discrete outcome of participation in each sector. In both cases the coefficients are normalized by the same value: the predicted level of employment under the counterfactual that the event of childbirth did not occur. See discussion in Section 2.4. Confidence intervals are not shown as they are too small to plot. Panel A corresponds to the maternal response, while Panel B to the paternal response. In both instances the age-dummies included in the underlying models are those of the mother. See Figure A.7 for underlying event-time coefficients with 95% confidence intervals.

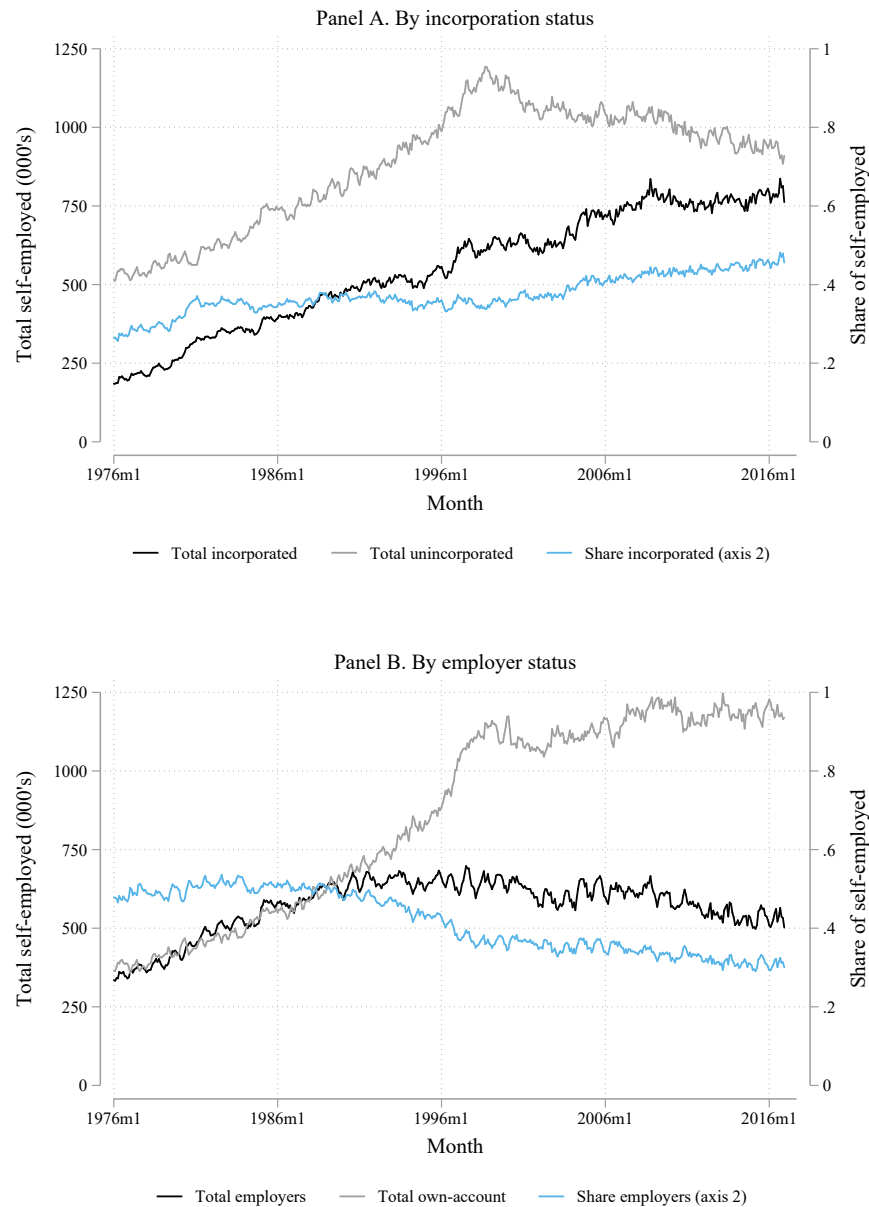


Figure 2.2: Trend in Canadian self-employment by incorporation and employer status.

Note: Estimated using the Canadian Labour Force Survey (1976-2016). Sample includes all self-employed workers, aged 25-54. Both figures plot the monthly trend in the total number of self-employed workers estimated using post-stratified sampling weights provided by Statistics Canada. In Panel A self-employed workers are grouped by incorporation status, while in Panel B they are grouped by employer status. In each figure, a third line depicts the share of self-employed workers in the first group and corresponds to the second right-hand y-axis.

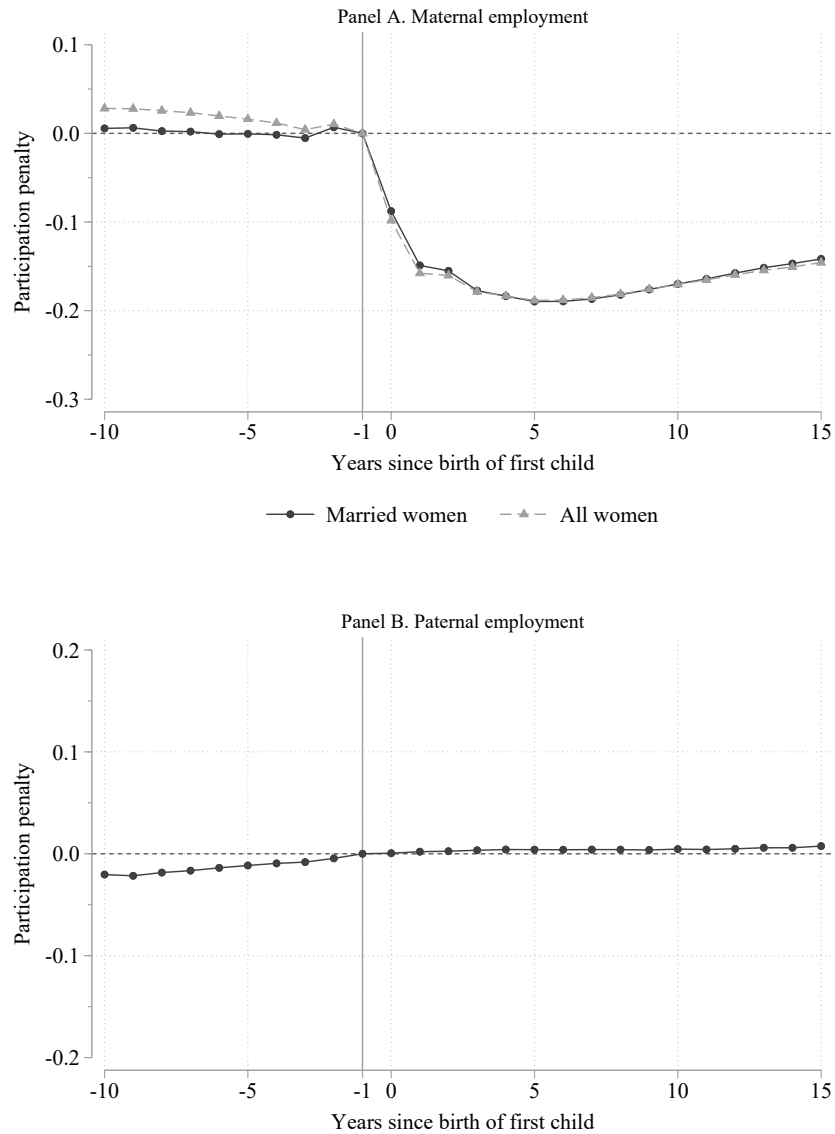


Figure 2.3: Participation parent penalty: overall employment

Note: Estimated using the Longitudinal Administrative Dataset (1988-2016). Sample includes all married couples where the female spouse is at least 20 years of age. Each data point corresponds to the rescaled event-time coefficient from an event-study-design (ESD) of initial childbirth. The curves represent the expected change in the probability of employment, normalized by the predicted level of employment under the counterfactual that the event of childbirth did not occur. See discussion in Section 2.4. Confidence intervals are not shown as they are too small to plot. Both the paternal and maternal curves correspond to a model in which the included age-dummies are those of the mother. See Figure A.8 for underlying event-time coefficients with 95% confidence intervals.

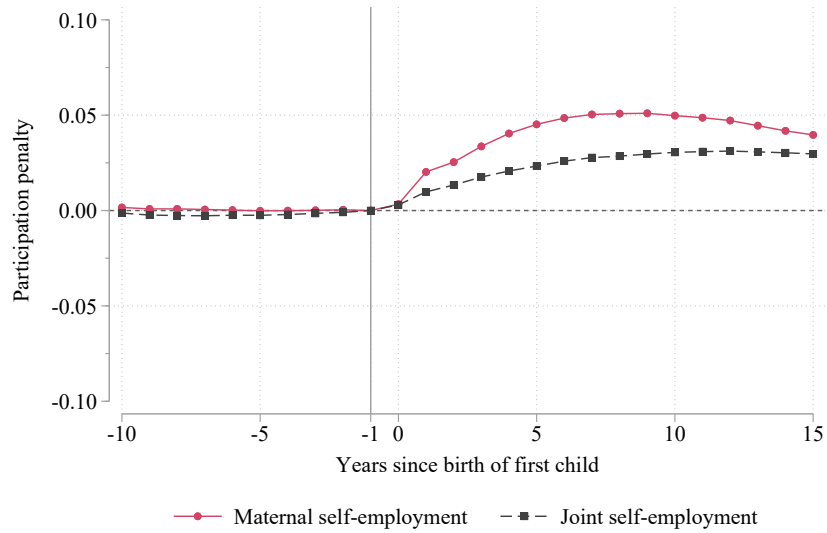


Figure 2.4: Participation parent penalty: maternal self-employment and joint self-employment

Note: Estimated using the Longitudinal Administrative Dataset (1988-2016). Sample includes all married couples where the female spouse is at least 20 years of age. Each data point corresponds to the rescaled event-time coefficient from an event-study-design (ESD) of initial childbirth. The curves represent the expected change in the probability of employment, normalized by the predicted level of employment under the counterfactual that the event of childbirth did not occur. As in Figure 2.1, each coefficient is normalized by the expected level of employment under the counterfactual of no childbirth, unconditional on sector of spouse. See discussion in Section 2.4. Confidence intervals are not shown as they are too small to plot. Both the paternal and maternal curves correspond to a model in which the included age-dummies are those of the mother. See Figure A.9 for underlying event-time coefficients with 95% confidence intervals.

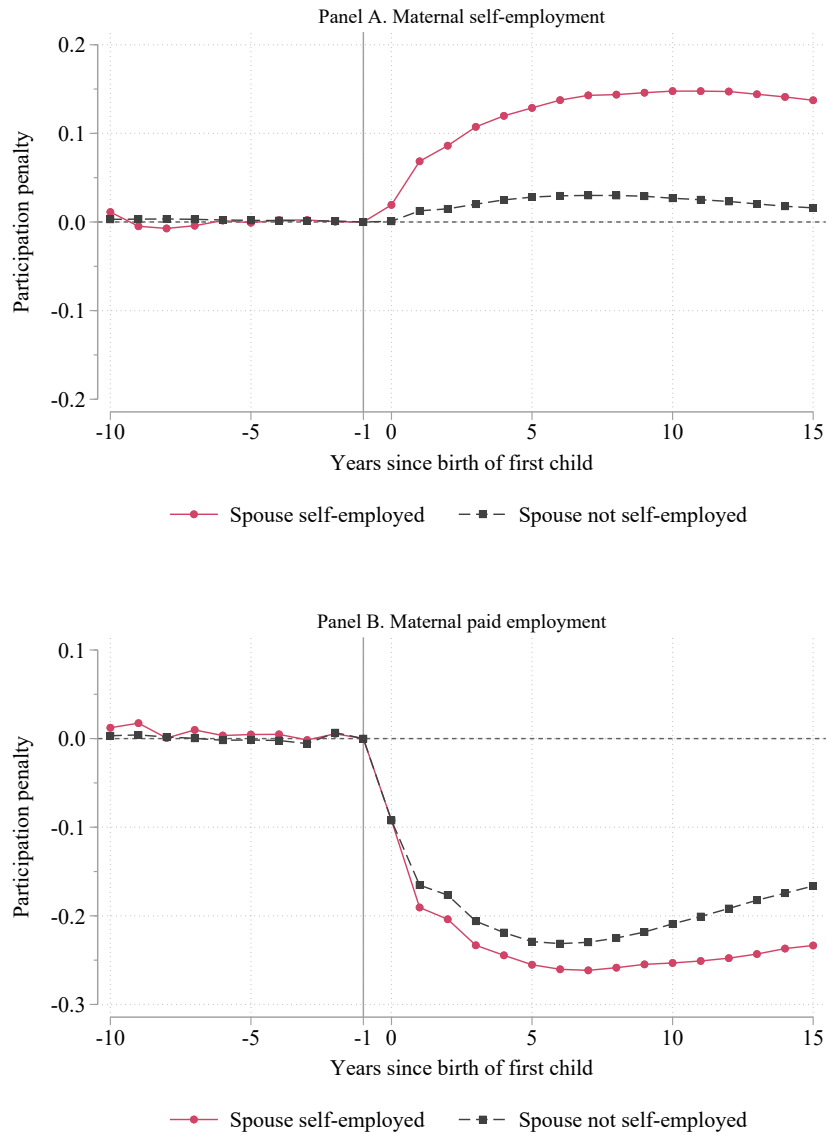


Figure 2.5: Participation parent penalty: class of employment, by class of spouse

Note: Estimated using the Longitudinal Administrative Dataset (1988-2016). Sample includes all married couples where the female spouse is at least 20 years of age. Each data point corresponds to the rescaled event-time coefficient from an event-study-design (ESD) of initial childbirth. Panel A represents the outcome of self-employment, while Panel B the outcome of wage-employment. As in Figure 2.1, each coefficient is normalized by the expected level of employment under the counterfactual of no childbirth, unconditional on sector of spouse. See discussion in Section 2.4. The coefficients are estimated by interacting the event-time and age dummy variables with the self-employment status of the spouse. The year fixed effects are not interacted. See Figure A.10 for underlying event-time coefficients with 95% confidence intervals.

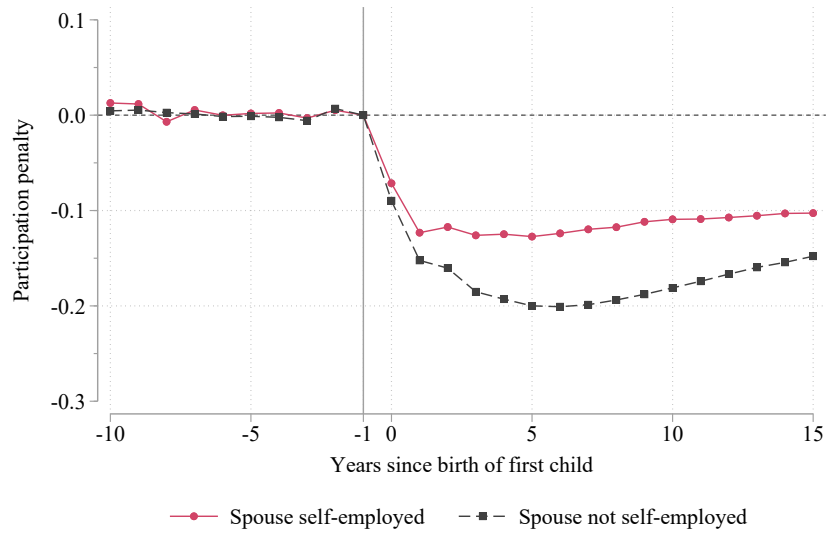


Figure 2.6: Participation parent penalty: overall employment, by class of spouse

Note: Estimated using the Longitudinal Administrative Dataset (1988-2016). Sample includes all married couples where the female spouse is at least 20 years of age. Each data point corresponds to the rescaled event-time coefficient from an event-study-design (ESD) of initial childbirth. Panel A represents the outcome of self-employment, while Panel B the outcome of wage-employment. As in Figure 2.1, each coefficient is normalized by the expected level of employment under the counterfactual of no childbirth, unconditional on sector of spouse. See discussion in Section 2.4. The coefficients are estimated by interacting the event-time and age dummy variables with the self-employment status of the spouse. The year fixed effects are not interacted. See Figure A.11 for underlying event-time coefficients with 95% confidence intervals.

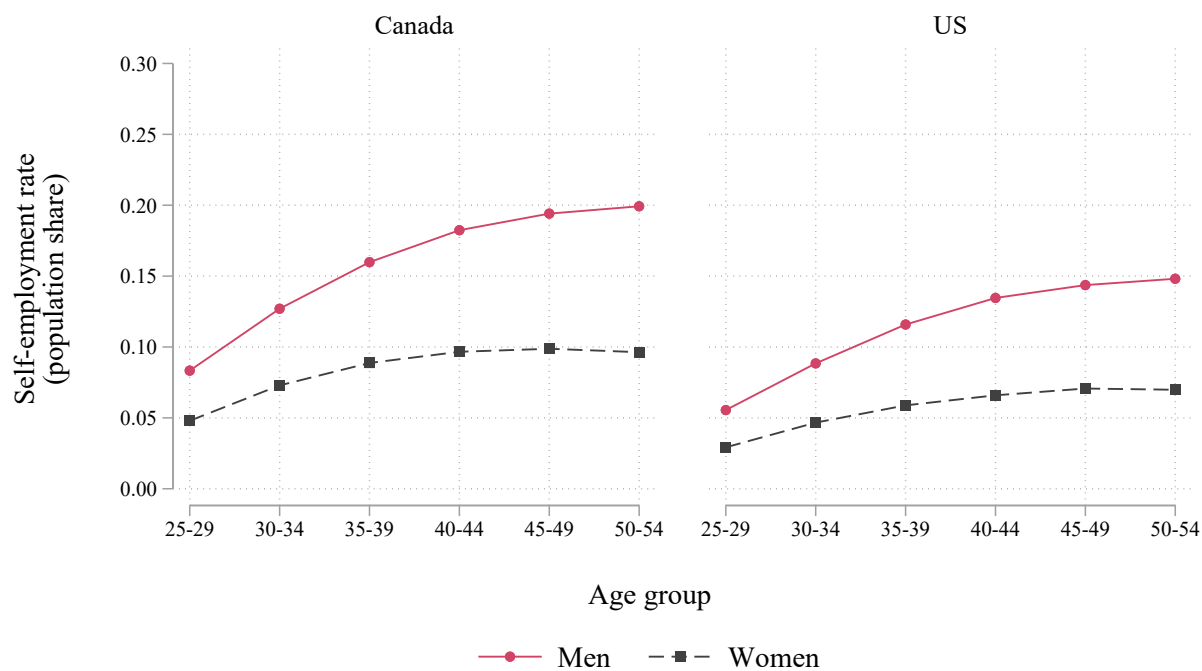


Figure 2.7: Age profile: self-employment in Canada and the US

Note: Estimated using the Canadian LFS (1988-2016) and US CPS (1988-2016). Sample all individuals aged 25-54. Each data point corresponds self-employment rate (share of population) within a five-year age bracket. Estimates are calculated using cross-section weights.

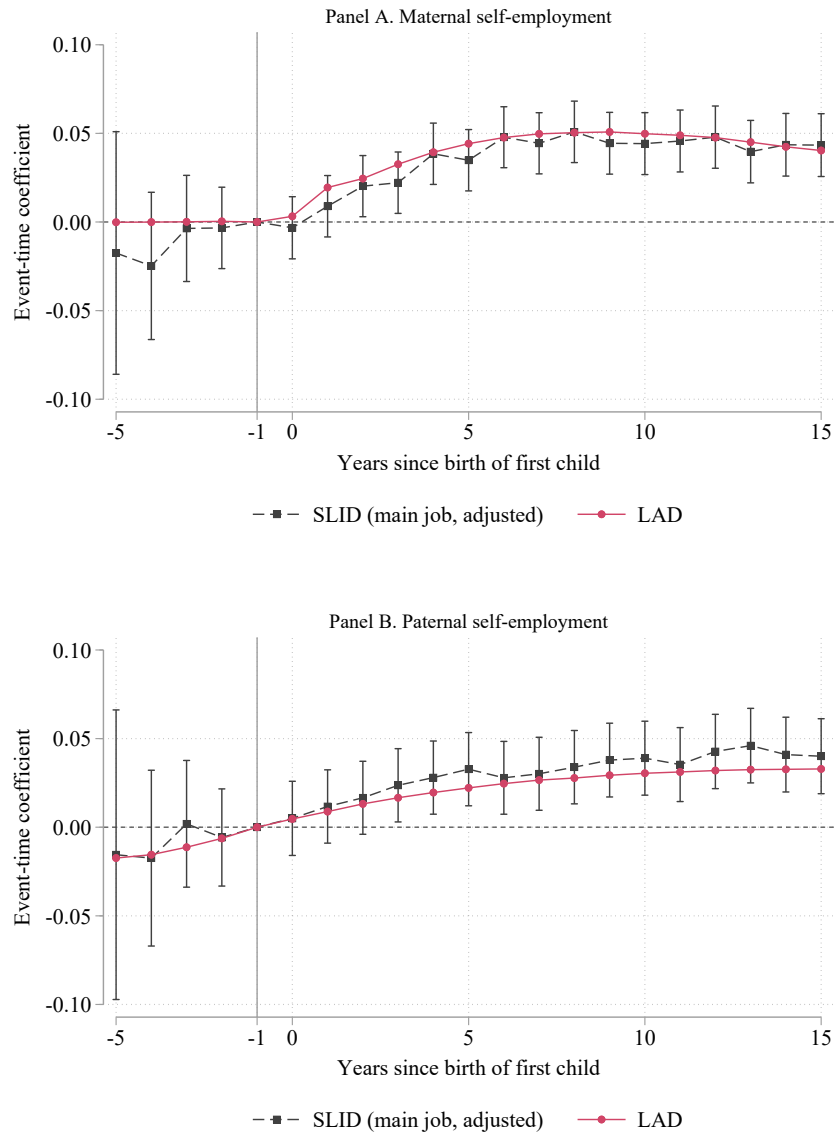


Figure 2.8: Comparison of event-time coefficients: participation in self-employment

Note: Estimated using the Survey of Labour and Income Dynamics (1996-2010). Sample includes all married/common-law couples, where the female spouse is at least 20 years of age. Estimates are unweighted, and 95% confidence intervals are shown. For LAD sample details see Figure 2.1.

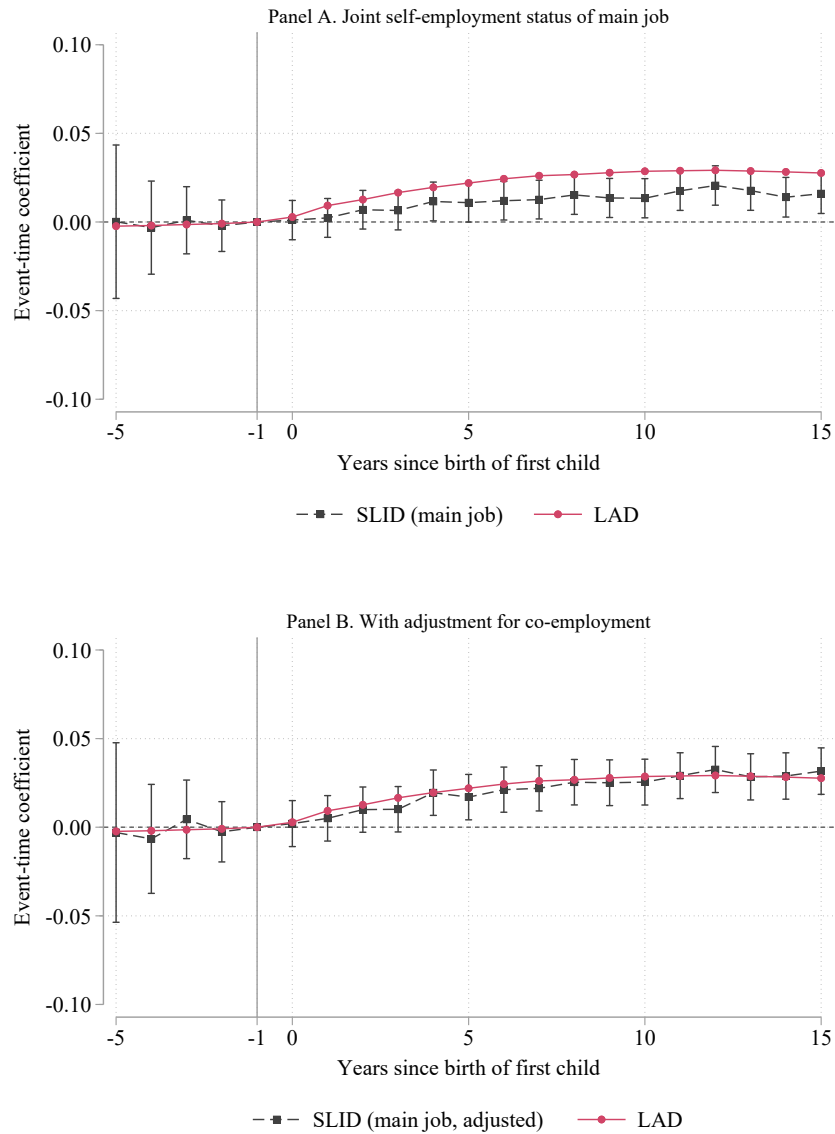


Figure 2.9: Comparison of event-time coefficients: joint self-employment

Note: Estimated using the Survey of Labour and Income Dynamics (1996-2010). Sample includes all married/common-law couples, where the female spouse is at least 20 years of age. Estimates are unweighted, and 95% confidence intervals are shown. For LAD sample details see Figure 2.4.

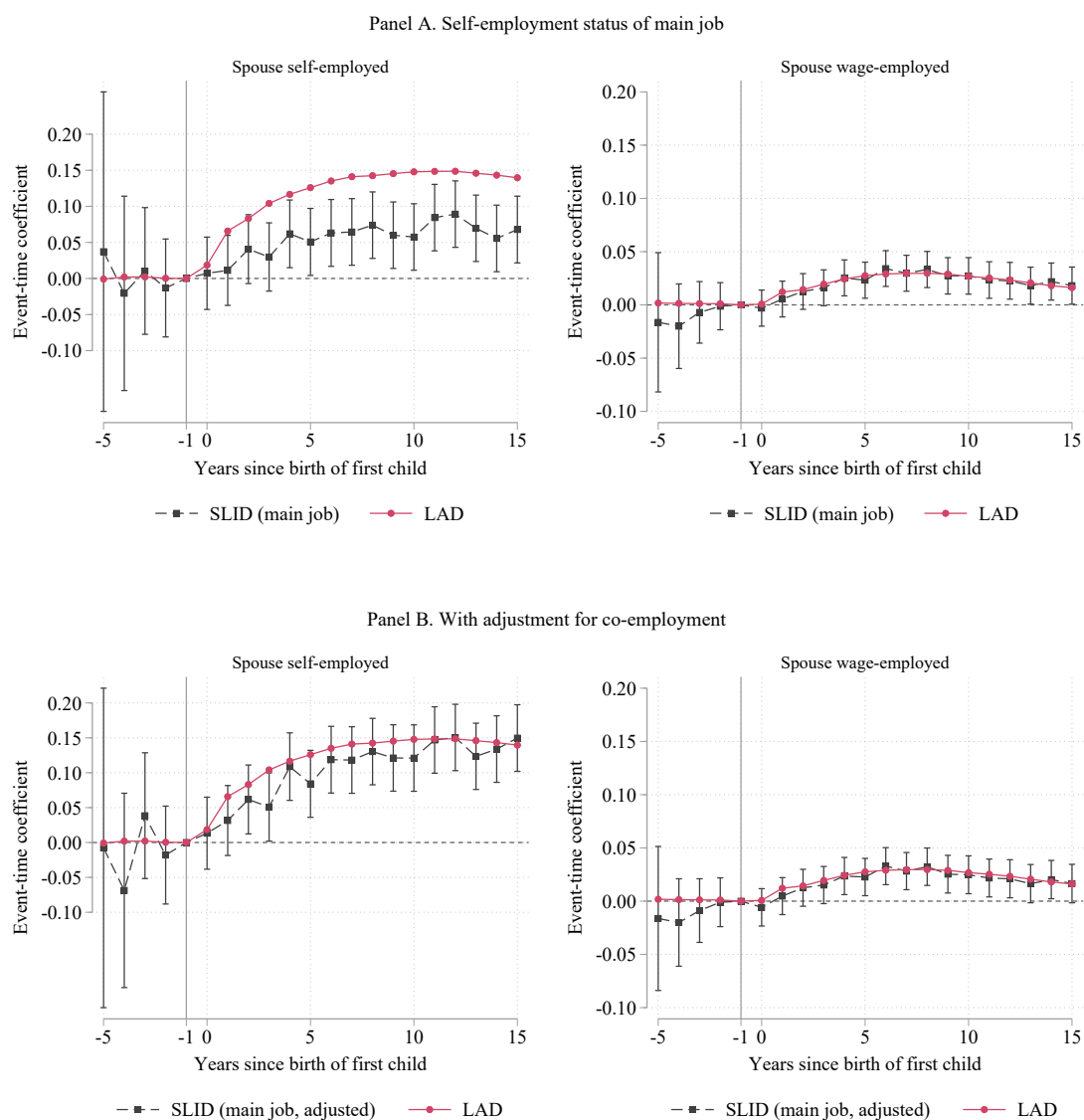


Figure 2.10: Comparison of event-time coefficients: participation in self-employment by self-employment status of spouse

Note: Estimated using the Survey of Labour and Income Dynamics (1996-2010). Sample includes all married/common-law couples, where the female spouse is at least 20 years of age. Estimates are unweighted, and 95% confidence intervals are shown. For LAD sample details see Figure 2.5.

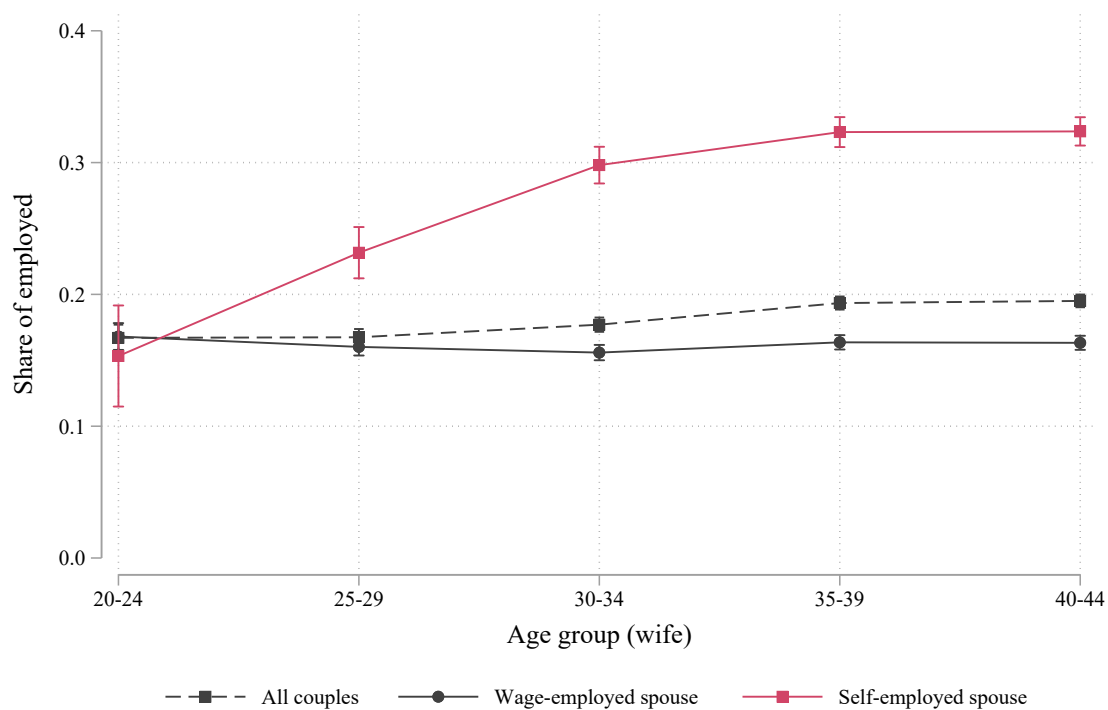


Figure 2.11: Share of two-earner couples employed in the same industry by age group and spouse's class of employment.

Note: Estimated using the public release version of the Survey of Labour and Income Dynamics (1996-2010). Sample includes all couples, where the female spouse is aged 20-44, and male spouse is aged 20-54. Each marker represents the share of two-earner couples in that age-group that are employed in the same industry. Estimates are unweighted, and 95% confidence intervals are shown.

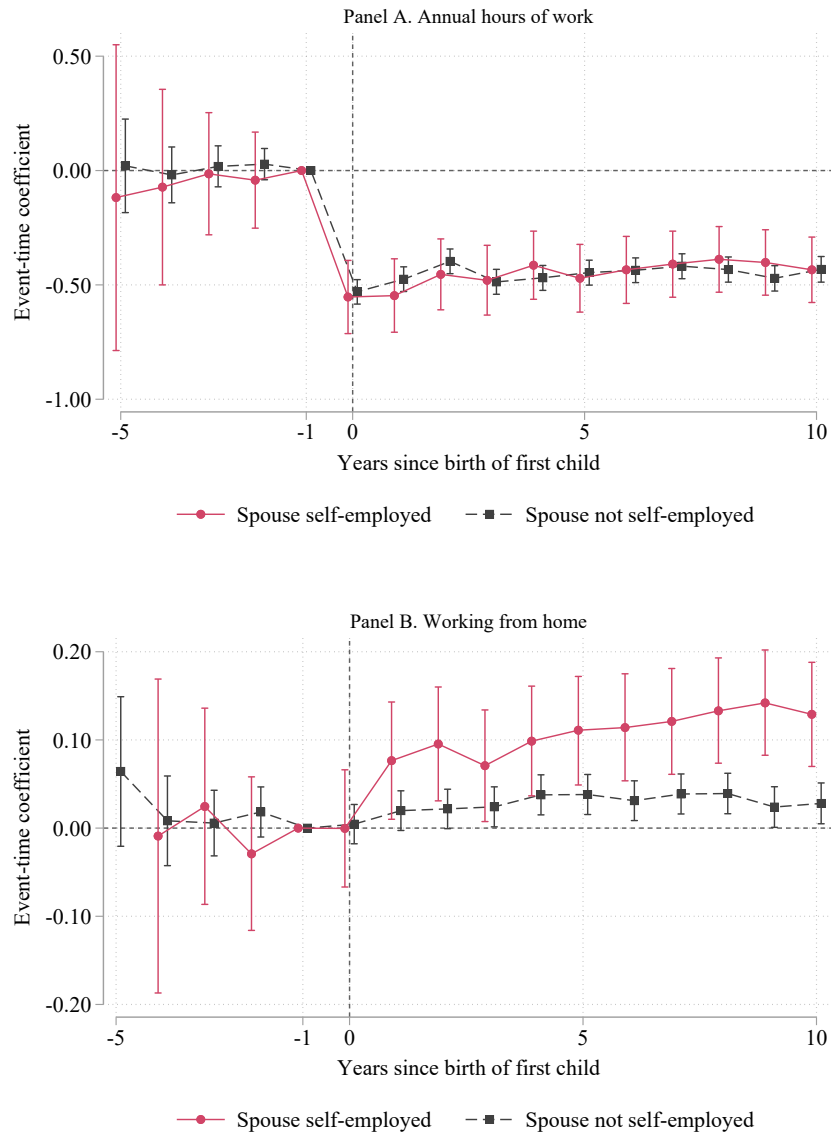


Figure 2.12: Comparison of event-time coefficients: log of annual hours worked and probability of working from home, by spouse's class.

Note: Estimated using the Survey of Labour and Income Dynamics (1996-2010). Sample includes all married/common-law couples, where the female spouse is at least 20 years of age. Estimates are unweighted, and 95% confidence intervals are shown. Markers estimate the change relative to the base period for households where the male spouse is and is not self-employed. The level difference is not shown. Details of the model are provided in Table 2.1.

Table 2.1: Changes at the intensive margin of labour supply among employed women

	<i>Annual weeks worked</i>			<i>Log of annual hours</i>		
	Spouse not self-empl.	Spouse self-empl.	Difference	Spouse not self-empl.	Spouse self-empl.	Difference
<i>Event-time coefficients</i>						
s=-5	0.743 (1.336)	1.855 (4.373)	1.113 (4.570)	0.0207 (0.104)	-0.118 (0.341)	-0.139 (0.356)
s=-4	0.116 (0.800)	2.388 (2.799)	2.272 (2.910)	-0.0192 (0.0624)	-0.0725 (0.218)	-0.0533 (0.227)
s=-3	0.239 (0.585)	-0.762 (1.746)	-1.001 (1.840)	0.0181 (0.0456)	-0.0144 (0.136)	-0.0325 (0.144)
s=-1	0.196 (0.447)	-0.413 (1.373)	-0.609 (1.443)	0.0281 (0.0349)	-0.0421 (0.107)	-0.0702 (0.113)
s=0	-2.683*** (0.351)	-3.413*** (1.045)	-0.730 (1.102)	-0.530*** (0.0275)	-0.553*** (0.0815)	-0.0226 (0.0860)
s=1	-4.073*** (0.353)	-2.126** (1.045)	1.948* (1.103)	-0.475*** (0.0276)	-0.547*** (0.0818)	-0.0719 (0.0863)
s=2	-3.201*** (0.352)	-1.411 (1.014)	1.790* (1.072)	-0.397*** (0.0275)	-0.454*** (0.0791)	-0.0570 (0.0836)
s=3	-4.228*** (0.357)	-1.086 (0.996)	3.142*** (1.056)	-0.487*** (0.0278)	-0.480*** (0.0778)	0.00688 (0.0824)
s=4	-4.868*** (0.357)	-1.623* (0.975)	3.245*** (1.035)	-0.469*** (0.0279)	-0.414*** (0.0761)	0.0548 (0.0807)
s=5	-4.691*** (0.358)	-1.809* (0.969)	2.882*** (1.028)	-0.446*** (0.0279)	-0.471*** (0.0756)	-0.0247 (0.0802)
s=6	-4.967*** (0.355)	-1.330 (0.956)	3.636*** (1.013)	-0.436*** (0.0277)	-0.434*** (0.0746)	0.00176 (0.0791)
s=7	-4.761*** (0.357)	-1.756* (0.946)	3.005*** (1.004)	-0.418*** (0.0279)	-0.409*** (0.0738)	0.00921 (0.0783)
s=8	-4.734*** (0.361)	-1.739* (0.940)	2.995*** (0.997)	-0.433*** (0.0281)	-0.388*** (0.0733)	0.0447 (0.0778)
s=9	-5.254*** (0.362)	-1.808* (0.936)	3.446*** (0.993)	-0.471*** (0.0283)	-0.402*** (0.0730)	0.0696 (0.0774)
s=10	-5.246*** (0.364)	-2.016** (0.932)	3.230*** (0.987)	-0.432*** (0.0284)	-0.434*** (0.0727)	-0.00191 (0.0770)
<i>Level difference</i>						
Spouse self-empl.	-1.831** (0.852)			-0.0103 (0.0664)		
Covariates	Yes			Yes		

Note: Estimated using the Survey of Labour and Income Dynamics (1996-2010). Sample includes all married/common-law couples, where the female spouse is at least 20 years of age. Estimates are un-weighted and standard errors are shown. Covariates include province fixed effects, a quadratic in age of spouse, and an indicator for common-law couples.

Table 2.2: Changes in taxable income per hour and working from home status among employed women

	<i>Log of effective wage</i>			<i>Work from home</i>		
	Spouse not self-empl.	Spouse self-empl.	Difference	Spouse not self-empl.	Spouse self-empl.	Difference
<i>Event-time coefficients</i>						
s=-5	0.0642 (0.112)	0.0245 (0.365)	-0.0397 (0.382)	0.0643 (0.0433)	-0.150 (0.142)	-0.215 (0.148)
s=-4	0.0574 (0.0672)	0.217 (0.234)	0.159 (0.243)	0.00827 (0.0259)	-0.00903 (0.0908)	-0.0173 (0.0944)
s=-3	0.0668 (0.0489)	0.0437 (0.146)	-0.0231 (0.154)	0.00571 (0.0190)	0.0245 (0.0566)	0.0188 (0.0597)
s=-1	0.0500 (0.0374)	0.0872 (0.116)	0.0372 (0.122)	0.0183 (0.0145)	-0.0292 (0.0445)	-0.0475 (0.0468)
s=0	-0.136*** (0.0296)	-0.0580 (0.0882)	0.0781 (0.0930)	0.00452 (0.0114)	-0.000351 (0.0339)	-0.00487 (0.0357)
s=1	-0.238*** (0.0298)	0.0367 (0.0884)	0.275*** (0.0932)	0.0198* (0.0115)	0.0765** (0.0339)	0.0567 (0.0358)
s=2	-0.210*** (0.0296)	0.0327 (0.0852)	0.242*** (0.0901)	0.0218* (0.0114)	0.0954*** (0.0329)	0.0737** (0.0348)
s=3	-0.283*** (0.0300)	-0.0320 (0.0842)	0.251*** (0.0892)	0.0241** (0.0116)	0.0708** (0.0323)	0.0467 (0.0343)
s=4	-0.269*** (0.0301)	-0.0660 (0.0821)	0.203** (0.0871)	0.0377*** (0.0116)	0.0986*** (0.0316)	0.0609* (0.0336)
s=5	-0.291*** (0.0301)	-0.109 (0.0818)	0.181** (0.0867)	0.0381*** (0.0116)	0.111*** (0.0314)	0.0724** (0.0334)
s=6	-0.341*** (0.0299)	-0.108 (0.0805)	0.233*** (0.0853)	0.0311*** (0.0115)	0.114*** (0.0310)	0.0833** (0.0329)
s=7	-0.385*** (0.0301)	-0.156* (0.0796)	0.229*** (0.0844)	0.0387*** (0.0116)	0.121*** (0.0307)	0.0824** (0.0326)
s=8	-0.387*** (0.0303)	-0.254*** (0.0790)	0.132 (0.0839)	0.0392*** (0.0117)	0.133*** (0.0305)	0.0941*** (0.0324)
s=9	-0.384*** (0.0305)	-0.289*** (0.0787)	0.0950 (0.0835)	0.0239** (0.0117)	0.142*** (0.0304)	0.118*** (0.0322)
s=10	-0.432*** (0.0306)	-0.315*** (0.0783)	0.117 (0.0830)	0.0280** (0.0118)	0.129*** (0.0302)	0.101*** (0.0320)
<i>Level difference</i>						
Spouse self-empl.	-0.0728 (0.0714)			0.0871*** (0.0276)		
Covariates	Yes			Yes		

Note: Estimated using the Survey of Labour and Income Dynamics (1996-2010). Sample includes all married/common-law couples, where the female spouse is at least 20 years of age. Estimates are un-weighted and standard errors are shown. Covariates include province fixed effects, a quadratic in age of spouse, and an indicator for common-law couples.

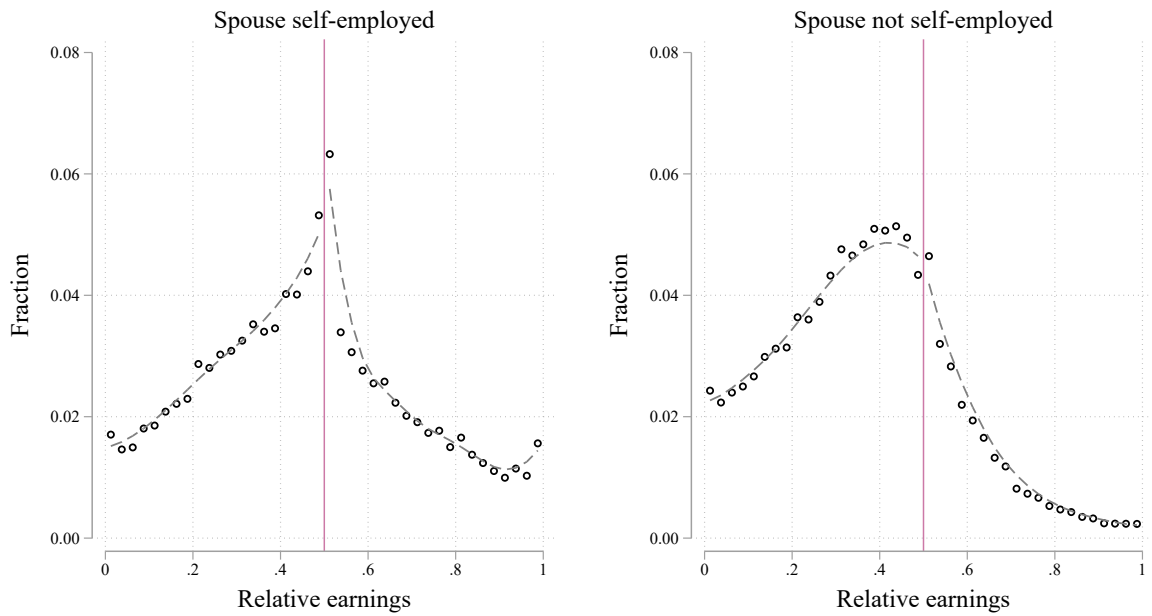


Figure 2.13: Relative employment earnings of married women, by class of spouse

Note: Estimate using public release version of the Survey of Labour and Income Dynamics (1996-2011), and Canadian Income Survey (2012-2016). Sample includes married, two-earner households where the female spouse is working aged (25-54). Employment earnings include dividend income. The male spouse's self-employment status is based on a self-reported indicator for any self-employed work in the past year. This work need not reflect the individual's main job. Each marker depicts fraction of individuals/couples within a 0.025 bin. The dashed curve depicts the LOWESS smoother; estimated separately for households on either side of the 0.5 cutoff. Cross-sectional weights have been applied.

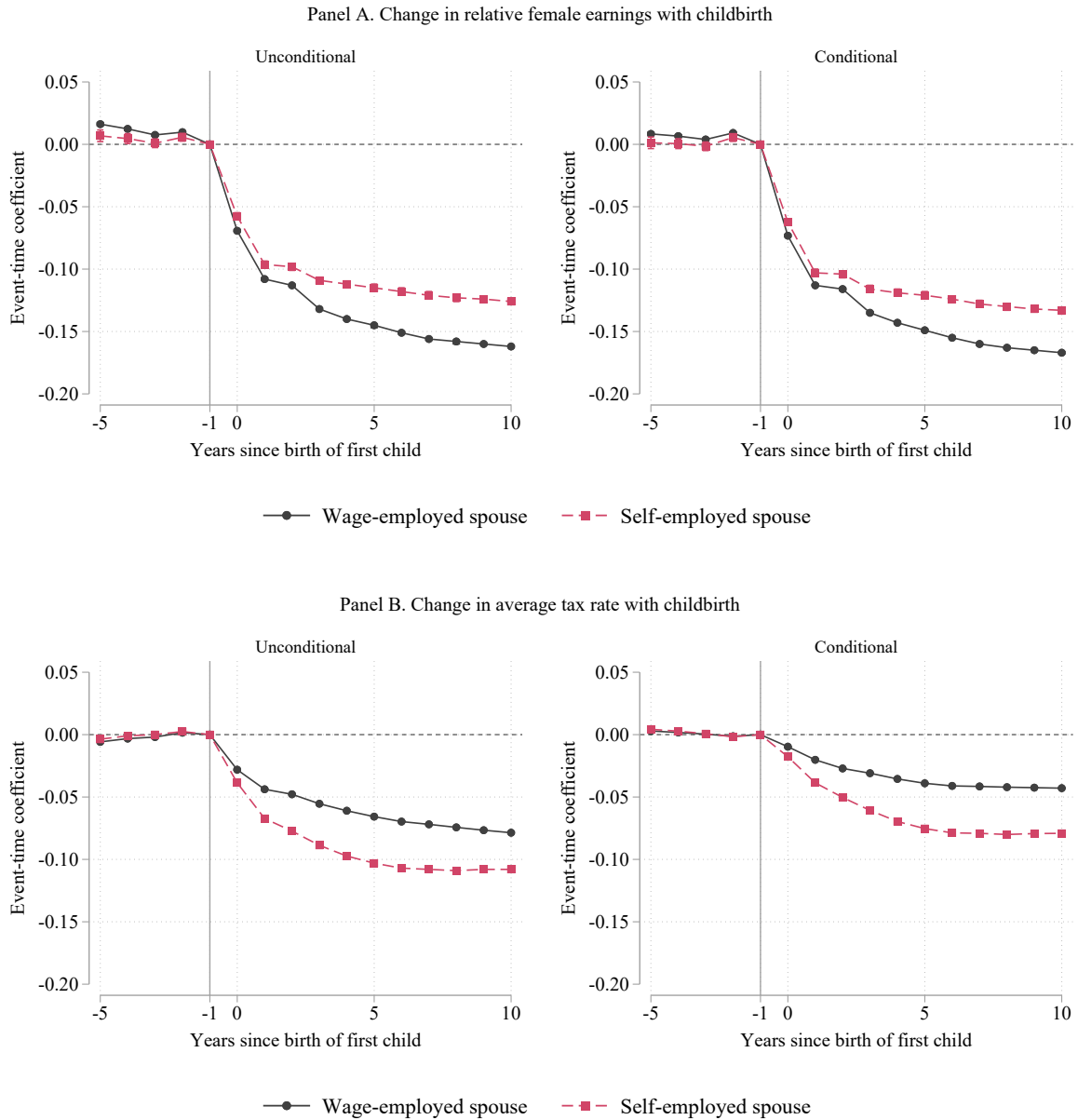


Figure 2.14: Changes in relative female earnings and average tax rate of the couple.

Note: Estimated using the Longitudinal Administrative Dataset (1988-2016). Sample includes all married couples where the female spouse is at least 20 years of age. Each data point corresponds to the event-time coefficient from an event-study-design (ESD) of initial childbirth. Panel A represents the change in relative earnings of the female spouse relative to the based period, by self-employment status of the spouse. Similarly, Panel B represents the change in the average tax rate of the couple relative to the based period, by self-employment status of the spouse. As in Figure A.10 the coefficients are estimated by interacting the event-time dummy with an indicator for the self-employment of the spouse. In the unconditional graph (left-hand side) includes year and age fixed effects in the estimating equation; where the latter is not interacted with the self-employment status of the spouse (as in A.10). The conditional graph includes province fixed effects, a quadratic in the spouse's age, indicator for common-law status, and log of total couple taxable income. The plot includes 95% confidence intervals, but as they are very small, they do not show.

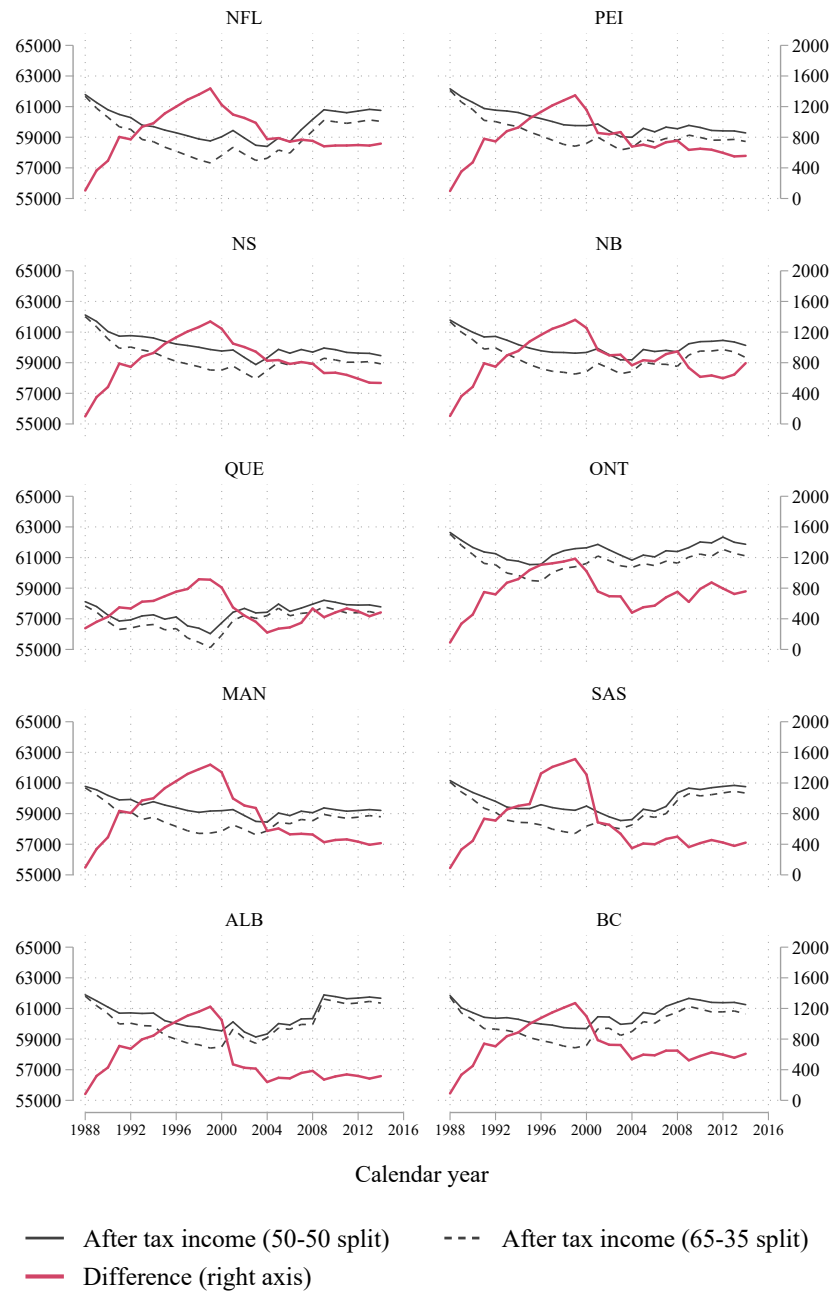


Figure 2.15: Trend in after-tax earnings for self-employed married households with a total income of \$80,000 (2016 CAD).

Note: Simulated using CTaCS simulator (Milligan, 2016). Shown by province of Canada with the difference plotted on the right-hand side axis. Both axes are denoted in 2016 Canadian dollars. The difference in between the after-tax income of the two curves approximates the value of income splitting for such a household.

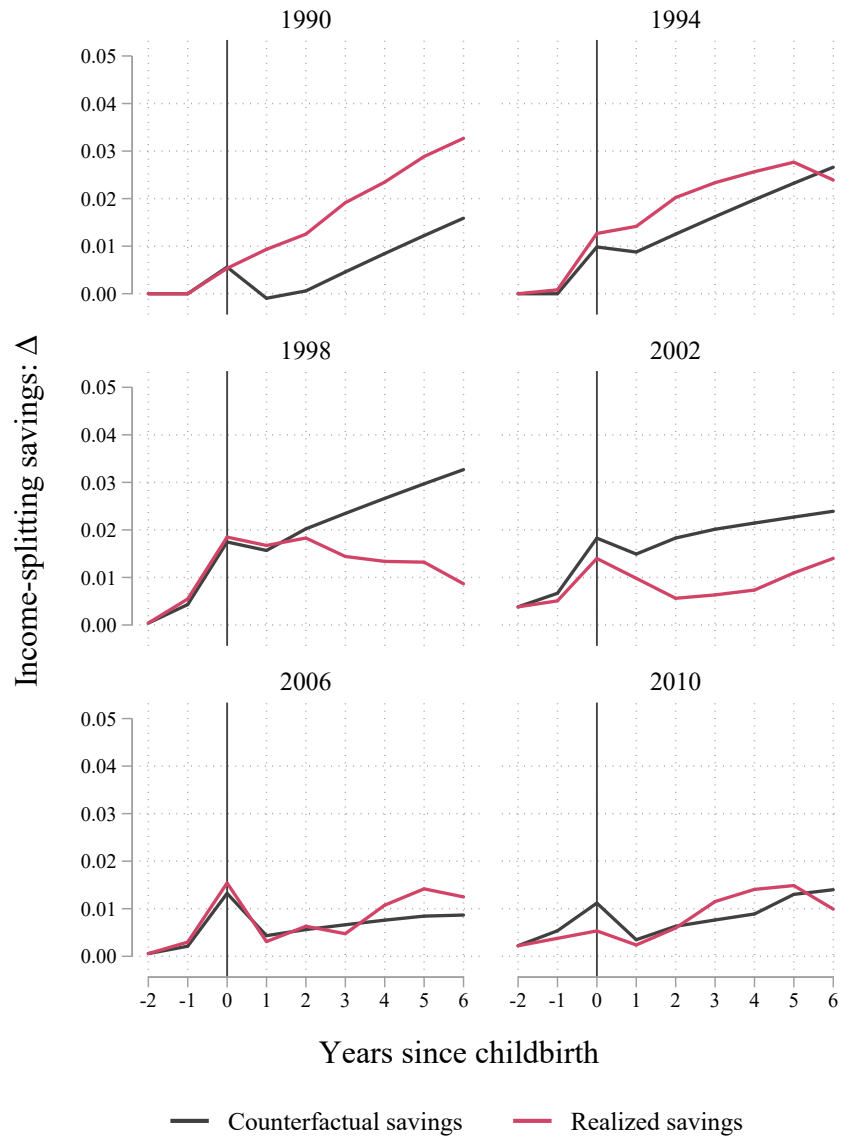


Figure 2.16: Simulated income-splitting tax savings (Δ) by birth cohort.

Note: Simulated using CTaCS simulator (Milligan, 2016). The simulation represents an Ontario couple with relative income of 45-55 and total income of \$80,000 (2016 CAD) two years prior to childbirth. The simulations assume that childbirth is associated with an asymmetric income shock of 20% in the year of childbirth, and 40% there after. Individual incomes are assumed to grow at 4% a year to account for the age-profile of earnings. The counterfactual path references the value of income splitting under the tax structure in the base period (event-time -2).

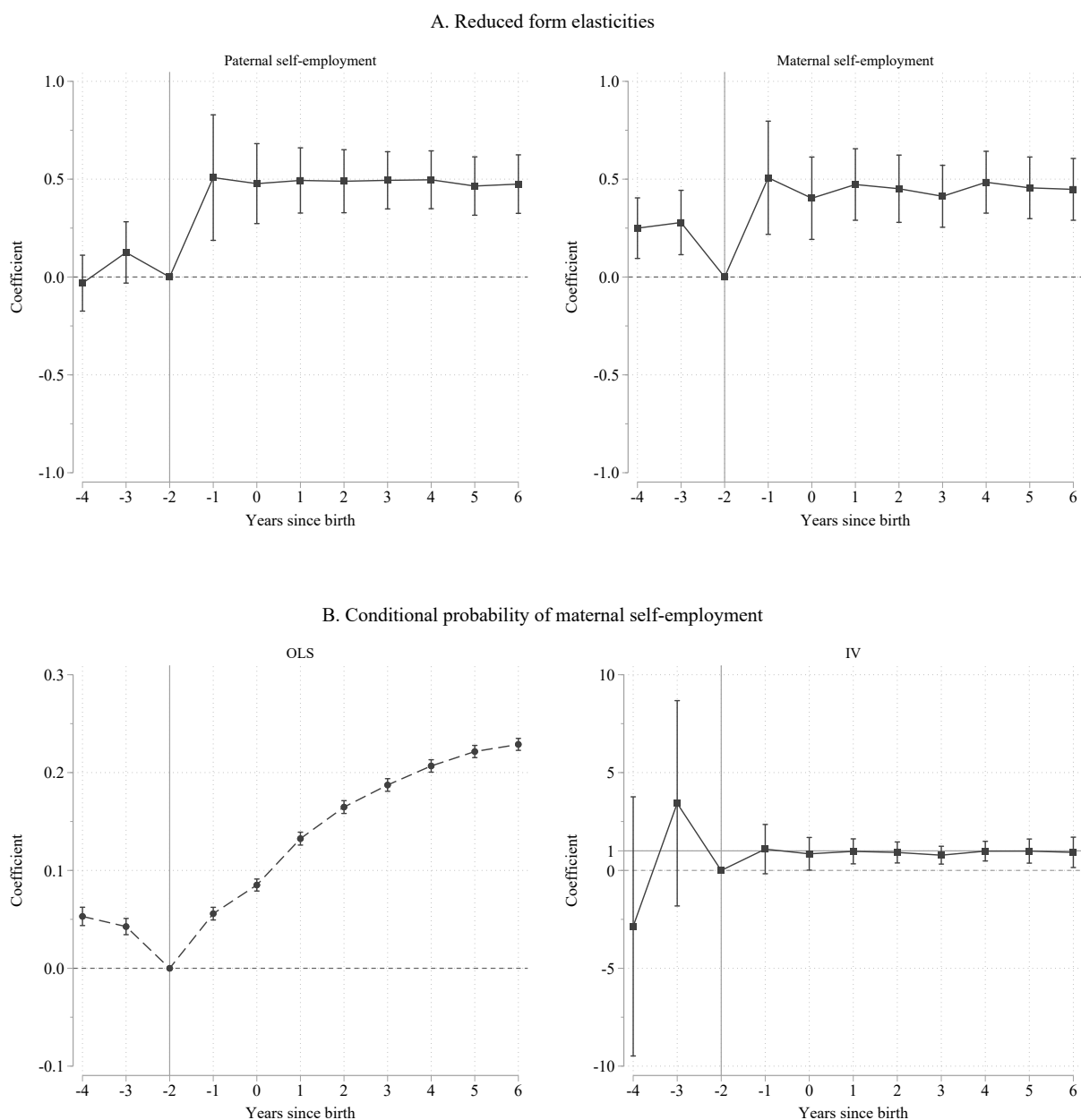


Figure 2.17: Comparison of event-time coefficients: participation in self-employment

Note: Panel A plots the coefficients corresponding to specification (2) and (5) from Table 2.3. Panel B plots the coefficients corresponding to specification (2) and (5) from Table 2.5. See Table notes for details of sample selection and model specification.

Table 2.3: Reduced form elasticities: paternal and maternal self-employment

	<i>Paternal Self-employment</i>			<i>Maternal Self-employment</i>		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Interaction of income-splitting savings (Δ) and event-time coefficients</i>						
s=-4	-0.169** (0.0725)	-0.0314 (0.0729)	0.0481 (0.0733)	0.310*** (0.0789)	0.250*** (0.0789)	0.0721 (0.0795)
s=-3	0.0208 (0.0799)	0.126 (0.0799)	0.206** (0.0804)	0.331*** (0.0837)	0.279*** (0.0839)	0.0999 (0.0842)
s=-1	0.658*** (0.164)	0.508*** (0.164)	0.484*** (0.163)	0.546*** (0.147)	0.507*** (0.148)	0.560*** (0.148)
s=0	0.540*** (0.104)	0.477*** (0.104)	0.482*** (0.104)	0.436*** (0.107)	0.403*** (0.107)	0.391*** (0.107)
s=1	0.486*** (0.0847)	0.494*** (0.0850)	0.508*** (0.0851)	0.509*** (0.0930)	0.473*** (0.0932)	0.441*** (0.0933)
s=2	0.447*** (0.0819)	0.490*** (0.0822)	0.501*** (0.0822)	0.490*** (0.0875)	0.451*** (0.0877)	0.425*** (0.0876)
s=3	0.412*** (0.0743)	0.494*** (0.0745)	0.508*** (0.0746)	0.453*** (0.0806)	0.413*** (0.0807)	0.382*** (0.0807)
s=4	0.381*** (0.0753)	0.497*** (0.0754)	0.510*** (0.0754)	0.528*** (0.0805)	0.484*** (0.0806)	0.455*** (0.0806)
s=5	0.315*** (0.0760)	0.465*** (0.0761)	0.477*** (0.0761)	0.503*** (0.0803)	0.456*** (0.0804)	0.429*** (0.0803)
s=6	0.284*** (0.0762)	0.475*** (0.0763)	0.485*** (0.0763)	0.497*** (0.0806)	0.448*** (0.0806)	0.425*** (0.0805)
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Income	No	Yes	Yes	No	Yes	Yes
Zero dummy	No	No	Yes	No	No	Yes

Note: This table contains the coefficients from the pooled model described in Section 2.5.1. The sample includes all couples where neither spouse is self-employed in event-time -2, and the male spouse is a wage-employee (with at least 5000 (2016 CAD's) in taxable income). In addition, I restrict on couples where the based period's total taxable income is less than half a million (2016 CAD's) in the event-time -2. Each coefficient corresponds to the event-time specific reduced form elasticities (β_j 's) described in formula 2.5. The estimating sample excludes period -2 as the outcome variable is zero for individuals. The estimates are weighted and standard errors are clustered at the individual level. Covariates include province fixed effects, a quadratic in age of spouse, and an indicator for common-law couples. The income controls include the log of the separate paternal and maternal predicted income values. The zero dummy is an indicator for household's where the female spouse had zero taxable income in the base period.

Table 2.4: Family formation gap in male self-employment: LFS & CPS, 1988-2015

	<i>Canada</i>		<i>United States</i>		<i>Pooled</i>	
	(1)	(2)	(3)	(4)	(5)	(6)
Children	0.0251*** (0.000431)	0.0210*** (0.000399)	0.00848*** (0.000437)	0.00658*** (0.000416)	0.00846*** (0.000320)	0.00639*** (0.000304)
Children*Canada					0.0160*** (0.000899)	0.0160*** (0.000851)
Age-group	Yes	Yes	Yes	Yes	Yes	Yes
Age-group*Canada	No	No	No	No	Yes	Yes
Year & State/Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	Yes	No	Yes	No	Yes
<i>N</i>	4371835	4371835	3744648	3744648	8116483	8116483

Note: Sample includes married, employed men, with spouses aged 20-44. The sample is further restricted to employed men aged 20-54. Age-group refers to 5-year age-group. All models include state/province and year fixed effects. Additional covariates include state-times-year fixed effects; education group (at most high school, some tertiary, bachelors or more); industry code (43-group); and state-year age-specific unemployment rate. Estimates are weighted using cross-sectional weights

Table 2.5: Relative probability of maternal self-employment conditional spousal self-employment

	<i>OLS</i>			<i>IV</i>		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Interaction of spousal self-employment (Q^1) and event-time coefficients</i>						
s=-4	0.0499*** (0.00479)	0.0530*** (0.00478)	0.0536*** (0.00476)	-5.796 (5.919)	-2.859 (3.378)	0.337 (1.241)
s=-3	0.0376*** (0.00424)	0.0426*** (0.00423)	0.0428*** (0.00422)	5.947 (5.203)	3.431 (2.676)	0.205 (0.993)
s=-1	0.0516*** (0.00331)	0.0559*** (0.00330)	0.0566*** (0.00329)	0.605 (0.506)	1.090* (0.643)	1.252* (0.700)
s=0	0.0824*** (0.00313)	0.0851*** (0.00313)	0.0855*** (0.00312)	0.549 (0.397)	0.850** (0.426)	0.792* (0.410)
s=1	0.131*** (0.00337)	0.132*** (0.00337)	0.133*** (0.00336)	0.915*** (0.321)	0.975*** (0.326)	0.877*** (0.309)
s=2	0.163*** (0.00334)	0.165*** (0.00334)	0.165*** (0.00333)	0.990*** (0.287)	0.919*** (0.273)	0.842*** (0.262)
s=3	0.186*** (0.00329)	0.187*** (0.00329)	0.188*** (0.00328)	0.978*** (0.260)	0.781*** (0.232)	0.696*** (0.223)
s=4	0.206*** (0.00324)	0.207*** (0.00323)	0.207*** (0.00322)	1.311*** (0.301)	0.986*** (0.256)	0.913*** (0.243)
s=5	0.220*** (0.00319)	0.222*** (0.00318)	0.222*** (0.00318)	1.528*** (0.400)	0.990*** (0.315)	0.939*** (0.296)
s=6	0.228*** (0.00312)	0.229*** (0.00312)	0.229*** (0.00311)	1.804*** (0.561)	0.924** (0.395)	0.909** (0.364)
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Income	No	Yes	Yes	No	Yes	Yes
Zero dummy	No	No	Yes	No	No	Yes

Note: This table contains the coefficients from the pooled model described in Section 2.5.2. The sample includes all couples where neither spouse is self-employed in event-time -2, and the male spouse is a wage-employee (with at least 5000 (2016 CAD's) in taxable income). In addition, I restrict on couples where the based period's total taxable income is less than half a million (2016 CAD's) in the event-time -2. Each coefficient corresponds to the interaction of the male spouse's self-employment status and event-time dummy (η_j 's) described in formula 2.6. The estimating sample excludes period -2 as the outcome variable is zero for individuals. The estimates are weighted and standard errors are clustered at the individual level. Covariates include province fixed effects, a quadratic in age of spouse, and an indicator for common-law couples. The income controls include the log of the separate paternal and maternal predicted income values. The zero dummy is an indicator for household's where the female spouse had zero taxable income in the base period.

Chapter 3

Does incorporation matter? Causal evidence from professional corporation reforms in Canada

3.1 Introduction

There are significant tax implications associated with the incorporation of a business, but do any of these matter for the labour market decisions of the business owner? In particular, their own labour supply and willingness to take on employees? This paper examines this question within the context of self-employed professionals. Regulated professionals typically have a high earning potential, relatively stable life-cycle income profile, and have some of the highest self-employment rates by occupation group. As such, incorporation has significant wealth and (personal) marginal tax rate implications, as demonstrated recently by [Wolfson and Legree \(2015\)](#) and [Wolfson et al. \(2016\)](#).⁸⁵⁸⁶ However, for reasons concerning the limitations of administrative data, these analyses ignore the possibility of a real individual or business response to the new legal structure.

Given their distinct tax treatment, we expect there to be a relationship between the legal structure of a business and the wedge between the corporate and individual income tax rates. Indeed, a number of studies find a significant relationship between the corporate tax rate or individual-corporate tax rate wedge and the legal structure of firms (among others [Egger et al., 2009](#); [Gordon and Slemrod, 1998](#); [Liu, 2014](#)). However, variation in the tax structure does not provide a valid instrument with which to examine the relationship

⁸⁵In their discussion of the income-tax-planning benefits of private corporations, [Wolfson et al. \(2016\)](#) find that the share of income earned by individuals at the top of the distribution increases substantially when closely held corporate income is correctly accounted for. These results are inline with [Smith et al. \(2019\)](#), who also document the popularity of corporate ownership (in this instance US S-corporations and partnerships) at the top of the income distribution. [Smith et al. \(2019\)](#) argue that these “pass-through” business profits represent a return to human capital.

⁸⁶Income splitting is an important feature of [Wolfson and Legree \(2015\)](#)’s discussion on incorporation. This places the discussion within a broader literature of tax evasion. For a good review of this literature see ([Slemrod, 2019](#)). See also [Alstadsæter et al. \(2019\)](#)’s analysis of the inequality implications of tax evasion, as well as [Scheuer and Slemrod \(2020\)](#)’s discussion on the taxation of the superrich.

between the legal structure of a business and labour market outcomes, as there can be a more direct link between these outcomes and the tax structure. In this chapter, I exploit Canadian regulatory reforms at the province-profession level that were introduced to permit the registration of regulated *professional corporations* as a direct instrument for the legal (and tax) status of self-employed individuals. Given the uniqueness of these reforms, this paper provides the first set of causal evidence on the relationship between legal business structure, owner labour supply, and hiring decisions.⁸⁷

Given its federal structure, Canadian professionals are regulated at the provincial level. Each profession is therefore represented by ten distinct regulatory bodies with their own governing Acts. The aforementioned reforms were introduced at the province and province-profession level to permit the registration of Canadian Controlled Private Corporations (CCPCs) with professional bodies (for additional discussion of these reforms see [Wolfson and Legree, 2015](#)). These professional corporations do not have limited liability but are treated as corporations for tax purposes. Using the staggered adoption of the reforms at the province-profession level, I model the take up of corporate structures within an event-study design framework. I find that the reforms were significant for health and legal professionals, but not for business and finance professionals (including accountants) who are mentioned in the legislation. Incorporation increased by 28% among health professionals and 12% among legal professionals by 10 years after the reform, holding constant the small business corporate tax rate. Notably, the increase in incorporation was relatively gender neutral. This provides me with a robust first stage with which to examine the relationship between incorporation, labour supply, and hiring; as well as heterogeneity by gender.

Using a pooled sample of the Canadian Labour Force Survey (LFS), I focus on the outcomes of weekly hours of work, part-time status, and hiring status (whether the firm has at least one employee). Over the period of 1987-2019, hours of work declined significantly across all four non-education professional occupation groups identified in the LFS: business and finance, natural and applied sciences, health, and legal. A decline that is mirrored across both paid employee and self-employed professionals. The decline in health professional hours is explored in detail by [Crossley et al. \(2009\)](#), who argue the increase in female representation partly accounts for the reduction in patient hours. Indeed, over this period female representation has increased from a third to a half of all employed, non-education professionals. The change for self-employed professionals is of a similar magnitude, although starting from a lower base. I find that neither the self-employment rate nor the gender composition of the health and legal professions was affected by these reforms.

⁸⁷As far as I am aware there are no existing studies on this topic. While many papers note difference in the characteristics of incorporated and unincorporated businesses and their owners, these largely reflect correlations with the natural issues of self-selection.

As with own labour supply the hiring practices of self-employed professionals has changed dramatically over this period. In the 1990s close to 80% of all self-employed professionals had at least one employee. Three decades later, this has declined to 40%. However, the relationship between hiring and legal structure may have a different mechanism to that of hours worked. Corporations generate tax savings for the professionals through retained earnings and income splitting. Retained earnings, play a similar role to a tax deferred pension scheme. However, through retained earnings, the corporate structure also increases the after-tax cash flow of the business, which can be used to ensure against uncertain revenue (Baron, 2013). In this way, the corporate structure can help cover regular expenses in the midst of irregular revenue, including labour costs.

On average, incorporated professionals work 12% more hours, are 8% less likely to work part-time, and 22% more likely to have employees (own calculation; see results in Section 3.7). Applying the ESD framework within an instrumental variable (IV) setting, I find no evidence of a causal relationship between legal structure and individual labour supply: either weekly hours of work or part-time status. However, I do find a significant result for hiring. Incorporation increases the likelihood that a self-employed professional has employees by 12%. This result is consistent with cash flow benefits of retained corporate earnings. This employment affect is also explained entirely by female professionals. Incorporation does not appear to change the hiring decisions of male professionals but increases the likelihood that a female self-employed professional has at least one employee by 33%. This marked gender difference is consistent with the existing evidence on gender differences in risk-taking (Bertrand, 2011; Charness and Gneezy, 2012; Eckel and Grossman, 2008) and entrepreneurial outcomes (Cowling and Taylor, 2001).

The paper proceeds as follows. I begin by providing an overview of the reasons for incorporation, specifically for professional practices. I then discuss the history of institutional reforms enacted to permit the registration of professional corporations and how these may be modelled in an ESD framework. Here, I provide a detailed discussion and decomposition of the IV estimator with an event-study first-stage. Thereafter, I provide an overview of the Labour Force Survey data and begin the results section with a summary of the rise of incorporated professionals and relationship to the reforms across various occupation groups. A discussion of the results relating to own labour supply and hiring practice follows.

3.2 Why professional businesses incorporate

There are several reasons a business may choose to incorporate, but for professional practices - which tend to be smaller owner-operated firms - the primary benefit lies in retained

earnings. In this section I discuss the primary differences between incorporated and unincorporated firms, highlighting their relevance for professional practices. I explain why factors including limited liability, credit, management, and income splitting are not the primary reasons that professionals choose to incorporate. Instead, retained earnings are shown to have two primary benefits: they can be used to increase the lifetime wealth of a professional or assist in stabilizing the cash of a business.

All differences between incorporated and unincorporated businesses stem from the creation of a separate legal entity: the corporation itself.⁸⁸ For all forms of unincorporated business structures – from the freelance writer to the large legal partnership – there is no legal distinction between the firm and its owner(s).⁸⁹ Incorporation changes this, by creating a legal barrier between the firm and its owner(s) with consequences for liability, agency, and taxation.⁹⁰ Here, I briefly explore each of these themes in the context of professional practices that have three important characteristics: the vast majority are small firms (even own account workers), they bare professional liability, and on the whole professionals have relatively high incomes.

Limited liability is a foundational institution in the theory of the firm, and is commonly understood to improve risk allocation and reduce monitoring costs for investors ([Eastbrook and Fischeltt, 1985](#)). Incorporation should be associated with greater levels of risk-taking, investment, and business growth; however, [Freedman \(2000\)](#) argues that this efficiency logic does not extend to smaller firms. For one, incorporation does not necessarily preclude small business owners from personal liability. [Devereux and Liu \(2016\)](#) find that “more than 70 percent of newly incorporated small and medium sized firms (SMEs) in the UK are required to provide personal security for their loans and mortgages”. For professionals, there is also the matter of professional liability. This will be discussed further in Section 3.3, but limited liability conflicts with professional liability, and is the reason that

⁸⁸In the Canadian context, privately owned corporations are commonly referred to as Canadian Controlled Private Corporations (CCPCs).

⁸⁹The caveat here is limited partnerships and limited liability partnerships (LLP), where the latter is typically reserved for groups of professionals in Canada. In a limited partnership a shareholder’s liability will depend on their role within the firm. This allows shareholders who hold a financial stake in the company but are not active in the running of the business, to limit their liability. Limited liability companies (LLCs) do not exist in Canada, as they do the US, UK, Switzerland, Chile, Columbia, Italy, Japan, and India. LLCs are a hybrid legal structure that provides limited liability protections to businesses, while retaining the same tax treatment as an unincorporated business. They also preserve the internal governance structure of a closely held firm. LLCs were first promoted in the US as means of avoiding double taxation: at both the corporate and shareholder level ([Freedman, 2000](#); [Ribstein, 1995](#)). The LLC is a separate legal entity but is taxed in the same way as a closely held firm. This is particularly attractive in cases where business owners (or partners) wish to personally benefit from business losses. LLC status should not be confused with S corporation status (or S subchapter), which reflects the tax filing status of a C corporation or LLC.

⁹⁰[Devereux and Liu \(2016\)](#) argue that formality is another important consideration. They argue that accounting requirements placed on incorporated firms provides a form of government back legitimacy, which lowers the information barrier in credit markets.

until recent reforms in Canada many self-employed professionals (especially those in health and legal professions) were unable to incorporate.⁹¹

A second motivation for incorporation is the *potential* separation of management and control [Fama and Jensen \(1983\)](#). With incorporation, one can limit the control that shareholders have over management decisions. While this is particularly advantageous in the context of large corporations - particularly, publicly traded companies - it is less so in the context of professional practices which typically remain closely held (or owner-managed) companies after incorporation. Limited liability and control cannot therefore be the primary motivation for professional incorporation.⁹² This leaves taxation.

As separate legal entities, corporations are taxed separately to their individual owners. The corporation pays corporate taxes on profits, which exclude any wages paid to shareholders. After corporate taxes are paid, shareholders may choose to issue remaining earnings as dividends (in accordance with their share allocations) or retain them within the corporate structure. The shareholders then pay personal income tax on any wages and dividends received from the corporation. In Canada, dividend income receives a tax credit to account for corporate taxes already paid on these earnings. This credit ensures that wage and dividend income enjoy the same marginal tax rate.

In contrast, unincorporated business income receives the same tax treatment as any other of employment income, with the exception of payroll tax liabilities. For this reason, the wedge between individual and corporate tax structures, which tend to be less progressive and have lower rates, plays an important role in the decision to incorporate. There is now a vast literature showing a robust positive relationship between the individual-corporate tax wedge and ownership structure (or organization form; [De Mooij and Nicodème, 2008](#); [Demirguc-Kunt et al., 2004](#); [Egger et al., 2009](#); [Goolsbee, 1998](#); [Gordon and MacKie-Mason, 1994](#); [Gordon and Slemrod, 1998](#); [Liu, 2014](#); [MacKie-Mason and Gordon, 1997](#)).⁹³ Notably, [Gordon and Slemrod \(1998\)](#) argue that this amounts to a form of income-shifting and not *real* economic change: the growing wedge between individual and corporate tax rates explains the shift from individual to corporate tax revenues in government accounts.⁹⁴

⁹¹Similar reforms were introduced in the US in the 1960s. In part, the Canadian reforms beginning in 1970 can be seen as a response to the US reforms ([Arnold, 1981](#); [Cader and Weinrib, 1965](#)).

⁹²Some practitioners promote the formal separation of personal and business accounts as an important advantage of incorporation. I do not address this concern here, but certainly accede that this may be preferential. However, it is not clear that

⁹³This wedge also has important implications for capital structures ([Faccio and Xu, 2015](#)).

⁹⁴It should be noted that business owners do not always prefer corporate structures for tax purposes. Indeed, in the US, the double taxation of business profits at both the corporate and individual level (when paid out in the form of dividends) was a primary motivation for the creation of LLC structures ([Ribstein, 1995](#)). Firms wanted the limited liability of corporations, while retaining the individual tax treatment and simpler management structures of unincorporated firms. However, as corporate tax rates have continued to decline globally the trend favours incorporation.

3.2. Why professional businesses incorporate

For professionals, whose incomes often place them in the top personal income tax brackets, this wedge is even more significant (see also discussion in [Wolfson and Legree, 2015](#)). Similarly, in jurisdictions with small business tax credits - a reduction in corporate tax rate applied to small businesses based on size, revenue, or profit - the wedge between the top individual income tax rate (which typically is lower than the small business threshold) and corporate tax rate can be large. For example, British Columbia's combined federal and provincial top individual tax rate in 2019 was 49.80% (for income above \$210,371), while the corporate tax rate at this level is 11% given the small business income threshold of \$500,000.⁹⁵ However, this wedge can be misleading. In Canada, the tax code is intentionally designed to eliminate any advantage from passing income through a CCPC. Similarly, it does not penalize (or double tax) income passed through a CCPC.⁹⁶ On net, dividends should be taxed at the same marginal tax rate as employment income once corporate taxes are taken into account.⁹⁷ For this reason, there are no significant tax advantages to being incorporated if all business profits are allocated to owner-managers through dividends and wages.⁹⁸ Moreover, when you take into account the administrative cost of incorporation (e.g. registration and obligatory accounting/bookkeeping requirements), passing professional earnings through a CCPC for is arguably costly to the business owner ([Baron, 2013](#)).

However, tax advantages do exist and they are well known ([Baron, 2013](#); [Wolfson and Legree, 2015](#); [Wolfson et al., 2016](#)). Indeed, there are even articles in the *Canadian Medical Association Journal* explaining in detail the various tax benefits of incorporation to its members ([Capen, 1994](#)). The first, is in the form of income splitting. Incorporation provides additional avenues through which income can be split with other household members

⁹⁵Canada's small business tax credit is based on a taxable corporate income threshold. Given the country's federal structure there are both provincial and federal tax credits, with their own corresponding thresholds. In 2019, Manitoba had the lowest small business tax rate of 0%, applying a full 100% credit of the provincial corporate tax rate to businesses below a \$500,000 taxable income threshold.

⁹⁶This in contrast to the US's double taxation of dividends which motivated the creation of LLCs ([Ribstein, 1995](#)). In contrast, Canada's Revenue Agency applies a dividend tax credit to a grossed-up portion of the dividends. This credit accounts for any corporate taxes paid.

⁹⁷The caveat here is the different payroll tax implications. Dividend income is exempt from any form of payroll tax. In Canada, this includes employment insurance (EI), and Canadian Pension Plan (CPP) or Quebec Pension Plan (QPP) contributions. EI is not a factor, as wage earnings paid to a corporate shareholder (who holds more than 40% of the shares) are by default uninsured. Similarly, are unincorporated earnings. A federal reform in 2010 allows incorporated and unincorporated business owners to voluntarily contribute to EI, but the uptake is very low. The exception here is Quebec, whose EI structure differs from the rest of Canada and does require EI contributions on self-employment income. Regardless, the primary difference is in the treatment of CPP/QPP. As dividend income is not included in pensionable income, individuals may choose to move money through dividends to avoid such contributions. However, there are no CPP savings when comparing unincorporated self-employment income and incorporated wage earnings. Moreover, Canadians tend to favour making CPP contributions given the investment track record of the CPPIB and long-term benefits of these contributions.

⁹⁸Small slippages may occur when dividend gross-up rates are not adjusted to account for changes in provincial corporate tax rates: however, these are typically minor.

to lower the overall liability of the household.⁹⁹ For married households and those with dependents this may provide a valuable source of tax savings. However, these savings have declined with changes to the income tax structure and measures have been put in place to limit the abuse of these avenues.¹⁰⁰ For example, the introduction of a ‘kiddie tax’ in January 2000: this increased the lowest marginal tax rate on dividends issued to child-dependents (under 18 years) to the highest marginal income tax rate (Bauer et al., 2015; Donnelly et al., 2000). The reform was immediately effective and resulted in a decline of 89% in dividend income reported by individuals under the age of 19 (Bauer et al., 2015). Finally, income splitting is also not exclusively available to incorporated households and should not therefore be considered the primary tax benefit. Instead, retained earnings represent the primary tax advantage.

There are *real economy* reasons for why businesses choose to retain profits and not issue dividends to their shareholders, including self-funded investment opportunities.¹⁰¹ But, for high income professionals a CCPC can be a valuable personal financial tool. As previously discussed, passing money through a CCPC is not beneficial, but income retained in a corporate structure is only taxed at the corporate rate. Given the personal-corporate tax differential, this leaves a greater share of business profits to be invested, allowing business owners to accumulate capital at a faster rate. For those concerned more with financial investments, as opposed to productivity focused capital investments, this provides a similar advantage to a tax-deferred pension scheme; in this instance, deferring taxes on dividends.¹⁰² As many smaller professional practices qualify for the small business tax credit (available to CCPCs), incorporation can dramatically increase the lifetime wealth of high-income professionals who save a large proportion of their after-tax income.¹⁰³ In addition, retained earnings within a CCPC reduce the marginal tax rate of income allocated

⁹⁹ Acts governing the registration of corporations with professional bodies vary in content too. Some require all shareholders of a registered corporation to be registered members of the profession. Such a requirement was amended by the regulated health profession body of Ontario in 2005, allowing the issuance of non-voting shares to family members. Similar laws were already in place in British Columbia. Crucially, this allows Ontario physicians to issue dividends to family members (i.e., income splitting). No such law was passed in favour of legal professionals, although legal professionals gained other tax benefits. See discussion in (Wolfson and Legree, 2015).

¹⁰⁰ See the discussion in Chapter 2.

¹⁰¹ A number of recent studies find strong causal evidence that businesses face external funding constraints and therefore rely on internal capital for investment (Chaney et al., 2012; Rauh, 2006; Zwick and Mahon, 2017).

¹⁰² For example, the Registered Retirement Savings Plan (RRSP) in Canada; Traditional IRA or 401(k) in the US; and workplace or personal pension (e.g., Self-Invested Personal Pension) in the UK.

¹⁰³ In recent years the Canadian government has acted to limit the value of this tax shelter. By distinguishing between active business income and investment income (i.e., rentals, interest, dividends and royalties) earned by a CCPC the government is able to tax passive income streams within a CCPC. In 2016, the Canadian government increased the marginal federal tax rate on CCPC investment income from 34.7% to 38.7% and in 2018 introduced a reform to reduce the small business tax credit limit (which then stood at \$500,000) based on the level of investment income. Any investment income over \$50,000 will see a reduction in the limit.

to the individual through wages and dividends. Thus, there is the potential for both a positive wealth and negative substitution effect, with an unknown impact on professional labour supply.

However, when it comes to retained earnings, *real economy* factors should not be overlooked. In addition to funding physical capital investments, retained earnings increase the cash flow available to small businesses that would otherwise be liable in the form of individual income taxes. This may help to stabilize business income and assist in meeting regular costs, such as employee wages. Note, this is not because incorporation lowers the cost of hiring, net of taxes. Employee wages are tax deductible in both incorporated and unincorporated settings. The mechanism described here is one of ‘precautionary’ cash-flow savings. Given the lower corporate tax rate, a larger share of after-tax earnings remains available for future expenses if left inside the corporate structure. Corporate structures may therefore facilitate hiring among professionals in settings where business owners are risk averse, face uncertain revenue streams (for demand or own labour supply reasons), or have credit constraints. For this reason, we can expect to find heterogeneity by occupation, gender, and age along this margin.

In summary, matters of limited liability and control are less likely to play a role in the incorporation of professional practices, as there already exist unincorporated structures that can provide inactive business partners with limited liability (while maintaining professional liability) and most professional corporations retain an owner-manager structure. Instead, the value of incorporation lies in retained earnings. These have the potential to increase the lifetime wealth of professionals and decrease the personal marginal tax rate of the professional. In addition, retained earnings increase the cash flow of professional practices ensuring that it is easier to meet operating costs such as wages/salaries. We can therefore expect to see a direct relationship between incorporation status, labour supply and hiring.

3.3 History of institutional reforms

As in the US, Canadian professions are regulated at the provincial (or state) level, in keeping with the country’s federal structure. Each professional body acts to ensure that its members abide by their given set of regulations, thus maintaining professional accountability and insuring the value of the profession. Registration therefore creates liability: those who act outside of these regulations can be held accountable, both legally and financially. It is for this reason, that practising a profession within a limited liability corporation is problematic. As a separate legal entity, a limited liability corporation could shield professional earnings, thereby limiting accountability, or as [Graschuk \(1977\)](#) puts it,

“...since the professional service is intangible and an assessment of its quality

relies wholly on judging the conduct of the practitioner, rather than on determining whether a piece of physical property is faulty, the professional must supply a guarantee of the good conduct of his service by offering up his possessions and professional reputation as a lifetime warranty.”

Canadian professionals – in particular, health and legal professionals – have historically been restricted from incorporating their businesses. This explains the historical significance of unincorporated sole-proprietorships and partnerships, as well as the disproportionate share of unincorporated professional earnings at the top of the income distribution. As this precludes professionals from accessing the aforementioned tax benefits of incorporation, professional groups have lobbied for reforms that allow for the registration of professional corporations. Following a sequence of reforms in the US through the 1960s, similar reforms became popular in Canada (Smeltzer, 1970). The result has been a series of legislative reforms at both the province and province-profession level to allow for the registration of a *professional corporation*.¹⁰⁴ While professional corporations are treated as CCPCs for tax purposes, they have unlimited liability.¹⁰⁵

The dates of these various reforms are collected in Table 3.1 which shows the date of provincial reform, followed by a set of profession-specific dates. The table is incomplete, with the exception of accountants, physicians, and lawyers. This is, in part, because the number of regulated professions has grown over time and the various reforms are difficult to keep track of as regulatory bodies change their structure. For example, all health professionals are now regulated under a single Act in Ontario, while in other provinces this is not the case.¹⁰⁶

British Columbia was the first to adopt a province-level reform in 1970; however, after deliberations beginning in 1972, it was finally repealed in 1976 (Cooper, 1978; MacIntyre, 1972).¹⁰⁷ Professional corporations did return to BC, with The College of Physicians and Surgeons of British Columbia registering medical corporations from 1979 and lawyers from 1992.¹⁰⁸¹⁰⁹ Alberta was next, adopting legislation permitting the establishment of professional corporations in 1975 (see discussion in Grashuk, 1977).¹¹⁰ This covered

¹⁰⁴In many instances, there is a lag between reforms at the professional level following a provincial reform. An additional source of variation that I exploit in this paper.

¹⁰⁵The legal structure typically allows for “limited liability with respect to non-professional obligations and activities” (Grashuk, 1977).

¹⁰⁶Similarly, there has been a unification of regulated accounting professions across Canada during this time, which has made it difficult to track these reforms. Where possible I have done my best to find the earliest date relevant to a profession.

¹⁰⁷The *Professional Corporations Act*, R.S.B.C. 1970, c. 37. See MacIntyre (1972) for a discussion on the act’s chequered start.

¹⁰⁸Under the *Medical Practitioners Act*, since replaced by *Health Professions Act* (2009).

¹⁰⁹*Legal Professions Act* (September 15, 1992)

¹¹⁰An amendment to the *The Companies Act*, as well as the Acts governing the four ‘senior’ professions, was

3.3. History of institutional reforms

Table 3.1: Years of regulation change

Province	NFL	PEI	NS	NB	QUE	ONT	MAN	SAS	ALB	BC
	1990 ¹	1988 ²	1989 ³	1981 ⁴	2001 ⁵	2001 ⁶	1999 ⁷	2001 ⁸	1975 ⁹	1979 ¹⁰
<i>By occupation</i>										
Accountants	2009	2005	1989	1986	2003/5/11	2004	1999	2001	1975	1979
Lawyers	2000	1999	2005	1996	2004	2007	1999	2001	1975	1992
Engineers			1989	1974		1660			1955	
Architects		1990			2012	1990			1969	
Physicians	2001	1987	1996	1983	2007	2001	1999	2000	1975	1979
Dentists						2001	1999	2001	1975	
Social Workers						2001				
Veterinarians						2001		2002		

¹ *Corporations Act* (R.S.N.L., 1990, c. C-36)

² *Business Corporations Act* (R.S.P.E.I., 1988, c. B-6.01) or the *Companies Act* (R.S.P.E.I. 1988, c. C-14)

³ *Companies Act* (R.S.N.S., 1989, c. 81)

⁴ *Business Corporations Act* (S.N.B., 1981, c. B-9.1)

⁵ *Professional Code*, CQLR c C-26. See *An Act to amend the Professional Code and other legislative provisions as regards the carrying on of professional activities within a partnership or company* (SQ 2001, c 34). Assented to 21 June 2001. See also Bill 46, *An Act to amend the Professional Code and other legislative provisions*. Assented to 19 June 2009

⁶ See Bill 152, *Balanced Budgets for Brighter Futures Act*, which received Royal Assent on December 21, 2000. See also Bill 45, *Responsible Choices for Growth and Accountability Act*, which received Royal Assent on June 29, 2001.

⁷ *The Professional Corporations (Various Acts Amendment) Act* (S.M. 1999, c. 41). Assented to July 14, 1999

⁸ *Professional Corporations Act* (S.S. 2001, c P-27.1). See also *Professional Corporations Regulations* (2002, RRS c P-27.1 Reg 2)

⁹ *The Attorney General Statutes Amendment Act* (R.S.A. 1975). Royal Assent given on December 15, 1975

¹⁰ *Professional Corporations Act* (R.S.B.C. 1970, c. 37) repealed April 1, 1976. See Part 2 of *Company Act* (R.S.B.C. 1979, c. 59)

members of Institute of Chartered Accountants, the Dental Association, the Legal Profession, and the Medical Profession; while other professions - including engineers, geologists, and architects - already had permission under their governing acts (see discussion in [Graschuk, 1977](#); [Rodney, 1969](#)). Similar provincial legislation appears in the Atlantic provinces during the 1980s, but the subsequent adoption of changes at the professional level within these provinces is much slower.

Canada's two most populous provinces – Ontario and Quebec – introduced provincial-level reforms in 2001. The 2001 reforms introduced the provincial legislation for the registration of professional corporations and listed the set of impacted professional Acts.¹¹¹ In Ontario, the *Regulated Health Professions Act* was amended on 1 November 2001, but the *Law*

given Royal Assent on December 15, 1975 ([Graschuk, 1977](#)). *The Attorney General Statutes Amendment Act*, R.S.A. 1975

¹¹¹This list included *Chartered Accountants Act*; *Certified General Accountants Association of Ontario Act*; *Law Society Act*; *Regulated Health Professions Act*; *Social Work and Social Service Work Act*; and *Veterinarians Act*.

Society Act was amended on 1 May 2007. In the case of Quebec, the provincial amendment does not include physicians, but a subsequent reform to the *Medical Act* on 21 February 2007 does. Thus, in Quebec lawyers adopted these reforms before physicians, while Ontario it was the other way around. Manitoba and Saskatchewan adopted professional corporation legislation in 1999 and 2001, respectively. In both instances, these reforms were met with simultaneous legislative changes at the professional level. In later sections I refer to the four neighbouring provinces of Quebec, Ontario, Manitoba, and Saskatchewan, as the ‘late adopters’.

The set of professions mentioned in each set of provincial reforms is different. For example, Ontario does not mention insurance professionals or architects, while Quebec does. Moreover, it is not always clear that the reforms were binding. Quebec’s legislation directly implicates engineers, but it appears that the existing professional act permitted incorporation. Indeed, engineers are one group which appear to have allowed incorporation at the professional level from as early as 1960 across various provinces. Similarly, accountants are mentioned in many provincial level reforms, yet in my own direct communications with professional accounting bodies, it does not appear that the reforms mattered. That is, many accountants practised through a corporation prior to the introduction of these reforms. This is consistent with the findings of Section 3.6.

As noted by [Wolfson et al. \(2016\)](#), additional distinctions can be made between some of these profession-specific reforms. For example, in a later 2005 reform, health professionals in Ontario secured the option of allocating non-voting shares of a medical corporation to an individual not registered as a health professional. This is widely interpreted as an income splitting opportunity; allowing non-professional spouses to be shareholders.¹¹² These benefits were not extended to legal professionals in Ontario. However, they in turn were given permission to organize practices as a collection of sub-contractors, each with their own corporate structure and small business income threshold. As discussed in Section 3.6, I find that incorporation increases faster among medical professionals than legal professionals. This may also reflect the appeal of alternative legal structures, such as limited liability partnerships which are popular among law and accounting firms (see discussion in [Hamilton, 1994](#)).¹¹³

¹¹²While this reform can be interpreted as legitimizing income splitting, it does not introduce them. Unincorporated businesses can also find ways to split income. As argued in Section 3.2, I do not perceive income splitting to be the primary reason for the rapid take-up of incorporation among Ontario medical professionals. Supporting this claim, I find that the take up incorporation in Quebec is equally steep.

¹¹³An important advantage of an unincorporated structure is the ability to “flow through” losses to the individual owner/partners.

3.4 Research design

The staggered adoption of these reforms at both the province and province-profession level can be modelled using an event-study design (ESD). As with a standard difference-in-difference (DID) methodology, causal inference depends on the parallel-trends assumption. A unique feature of the ESD is that this parallel-trend assumption can be applied between multiple treated groups so long as the timing of treatment varies across group.¹¹⁴ For this reason, the methodology does not necessitate the presence of a ‘never treated’ control group, but does require the assumption that responses do not pre-empt treatment. In this application all broad occupations groups are treated within the observed sample period and the regulatory framework ensures that businesses may only incorporate after the reform.

The methodology is implemented using a measure of event-time that varies independently from calendar time because of variation in the timing of treatment. Let S_{pn} denote the reform year in each province-profession pair; denoted by p and n , respectively. Then event-time, $s_{it} = t - S_{pn}$, is the number of years before/after the reform. If treatment effects are stationary, then one can specify the treatment using a dummy indicator: $D_{it} = 1\{s_{it} \geq 0\}$. This yields the standard two-way fixed effects (FEs) regression addressed in [de Chaisemartin and D’Haultfœuille \(2020\)](#). However, [Borusyak and Jaravel \(2018\)](#) show that if there are indeed dynamic treatment effects, the regression coefficient on $D_{it} = 1\{s_{it} \geq 0\}$ is a weighted average of dynamic treatment effects in which the weights may be negative and less weight is placed on treatment units that are treated early (see also [Abraham and Sun, 2018](#); [Goodman-Bacon, 2018](#)). The authors therefore suggest estimating a flexible, dynamic treatment effects model. This is done by estimating the standard ESD specification that includes an indicator for each period before and after the year of treatment; like that applied in Chapter 2 for years since birth of a first child. Practically, a dynamic specification fits the gradual adoption of incorporation with the reforms, thereby exploiting more of the variation than a static specification.

Let $Z_{it} = 1\{\text{incorporated}\}$ be an indicator of the incorporation status of self-employment professional i in year t . Then the event-time coefficient β_j represents the expected change in outcome variable j -periods from treatment. Under certain assumptions, β_j is the average treatment effect on the treated (ATT) j periods from treatment.

$$Z_{it} = \sum_{j \neq -1} \beta_j 1\{s_{it} = j\} + \delta_{pn} + \delta_t + X'_{it}\gamma + \varepsilon_{it}$$

One such assumption is an absence of pre-emptive behaviour, which ensures that future treated cohorts provide a valid counterfactual for earlier treated groups. However, a

¹¹⁴This contrasts the standard DID setting that makes use of an untreated control group. This control group may be a ‘never treated’ group, or a group treated outside of the period.

second important assumption is that of treatment effect homogeneity across treated units (or cohorts). In this instance, this implies homogeneity in the take up of incorporation across province-occupation pairs or province when the model is specified separately by occupation. As demonstrated by (Abraham and Sun, 2018), if there are heterogeneous treatment effects across cohorts, then the β_j regression coefficients are not only an average of cohort-specific ATT's for event-time j , but all event-times. Moreover, as demonstrated by Borusyak and Jaravel (2018) in their discussion of stationary versus dynamic specifications, these weights may be negative.¹¹⁵ This homogeneity assumption is also required for the pre-treatment coefficients, $\{\beta_j; j < -1\}$, to provide a valid test of parallel trends.

I provide occupation-specific and pooled estimates of β_j . As treatment effect homogeneity is more likely to hold within an occupation more weight should be put on the former occupation-specific estimates. Within occupation, homogeneity implies that the response to regulatory reforms do not vary by province. This would be problematic if provinces treated much earlier had a different take up of incorporation. One reason this may be the case is that the value of incorporating changes with the tax structure. For this reason, I control for the province-specific corporate tax rate in all estimated models.

In this paper, I apply the ESD methodology within an instrumental variable framework to study how incorporation affects the labour supply and hiring practices of self-employed professionals. While the take-up of incorporation is assumed to be dynamic, accommodating a more gradual response to the reform, the relationship between incorporation status and other labour market outcomes is modelled as a static relationship. The relationship between labour market outcome Y_{it}^k and incorporation status Z_{it} may be dynamic with respect to the timing of incorporation but is assumed to be constant with respect to event-time.¹¹⁶ Below, I specify k structural equations underlying the static relationship between incorporation status and the three labour market outcomes of hours worked ($k = 1$), part-time status ($k = 2$), and hiring ($k = 3$).

$$Y_{it}^k = \alpha^k Z_{it} + \mu_{pn}^k + \mu_t^k + X'_{it} \rho^k + v_{it}^k \quad k = 1, 2, 3$$

However, this does not imply that there isn't a relationship between the IV estimate of α^k and event-time. Indeed, I show that the ESD first-stage yields an IV estimator that is a weighted average of each reduced form and first-stage event-time coefficient, and the pooled

¹¹⁵Both Borusyak and Jaravel (2018) and (Abraham and Sun, 2018)'s arguments regarding the irregular weights applied in this setting are accounted for in the broader discussion of de Chaisemartin and D'Haultfœuille (2020) on two-way FEs models.

¹¹⁶Under the IV exclusion restriction the reform has no direct bearing on each of the assessed labour market outcomes. Nevertheless, the relationship between Z_{it} and Y_{it}^k may vary with the event-time if there is heterogeneity in the underlying structural equation that is correlated with the take up of incorporation. Such heterogeneity is not accounted for in this specification.

event-time first-stage (analogous to a static first-stage). Here I provide the decomposition for the case of a single endogenous variable, multiple event-time instruments, and no additional covariates (including year FEs). See Appendix B.2 for the case with additional exogenous covariates. Let J denote the set of excluded event-time indicators, then the standard 2SLS IV estimator is given by,

$$\hat{\alpha}^{IV} = \sum_{j \in J} \psi_j \hat{\alpha}_j^{IV} - \phi \hat{\alpha}_J^{IV}$$

where $\hat{\alpha}_j^{IV}$ is obtained either by excluding only the j -th event-time or the ratio of the j -th reduced form and first-stage coefficient. The coefficient $\hat{\alpha}_J$ is the IV estimate obtained by pooling (adding) all excluded event-times into a single indicator. The weights are given by,

$$\psi_j = \frac{(N_{jc} + N_j)\Omega_j}{\sum_{j \in J} (N_{jc} + N_j)\Omega_j - N_J\Omega_J}$$

$$\phi = \frac{N_J\Omega_J}{\sum_{j \in J} (N_{jc} + N_j)\Omega_j - N_J\Omega_J}$$

where N_j is the number of observations in event-time j , N_J is the number of observations in the excluded set J , and N_{jc} the number of observations outside of the excluded set of event-times (including the base period). Ω_j describes the increase in the explained sum squares from the inclusion of the j -th excluded event-time indicator in the first stage (or the pooled indicator in the case of Ω_J). The proof of this decomposition is provided in Appendix B.1.

It is evident that such an estimator will be sensitive to imprecisely estimated first-stage and reduced form event-time coefficients. It is for this reason that it is problematic to include pre-event coefficients in the excluded set of instruments. In a well design event-study-design these should be close to zero, which can lead to a very large IV estimate if the reduced form is non-zero: thereby biasing the above weighted estimator. Likewise, any post-event event-time coefficients that are small (in the first stage) can lead to a noisy IV estimate. Indeed, in this setting the delay between legislative and regulatory changes leads to a small first-stage coefficient in the immediate wake of the reform. I therefore focus on event-time periods $s = 2$ through 10 years after the event for my excluded set of instruments. I also provide comparable IV estimates based on a static first-stage and a parametric break-in-trend research design (see Appendix B.3).

As discussed in Chapter 2, ESDs are underidentified ([Abraham and Sun, 2018](#); [Borusyak and Jaravel, 2018](#)). Suggested solutions for this problem include excluding unit/cohort FEs from the estimation, including a control group, or to specify a semi-dynamic model

that assumes away pre-trends. As in Chapter 2, my primary specifications will follow the approach of excluding unit FEs. However, in this instance I also provide specifications that attempt to make use of a control group or exclude pre-trends to include unit FEs. As a control group, I make use of self-employed workers in the business and finance, and natural and applied sciences professions. The former group, which includes accountants, demonstrate no response to their reform dates (see Section 3.6); in part, because incorporation was permitted and relatively common prior to the reform. The same can be said for natural and applied science professionals, many of whom have been permitted to incorporate practices since the 1960s. Technically, this makes the two groups an ‘always treated’ control, as opposed to ‘never treated’ control. Regardless, the inclusion of a control group helps to pin down the year FEs, which in turn enables one to include unit FEs without a concern of underidentification. I do not put a lot of weight on estimates that make use of this control, as parallel trends assumption is likely to be violated. Consider that the trend in incorporation leading up to the reform should be close to 0, while in the control group of ‘always treated’ occupations, firms can incorporate freely in response to province specific policies, such as the corporate tax rate.¹¹⁷ It for this reason that a full or semi-dynamic ESD, without a control group, provides the best feasible estimator.

3.5 Data

The following analysis is based on the Canadian Labour Force Survey (LFS, 1987-2019).¹¹⁸ The primary benefit of using this survey data is its joint identification of professional occupations, self-employment and incorporation status. In addition, it is the primary source of Canadian data on hours worked, part-time status, and for self-employed workers identifies whether or not the business has employees.¹¹⁹ While Canadian administrative data bases can provide a clearer picture of earnings at both the individual and corporate level, they typically lack any records of labour supply.¹²⁰ Moreover, the existing administrative records, such as the LAD used in Chapter 2, do not include sufficient information on occupation

¹¹⁷I do include a control for the small business corporate tax rate in all of my specifications.

¹¹⁸The analysis is based on a public use version of the Labour Force Survey. Unfortunately, the 2016 files exclude crucial occupation codes needed. Any results using matched spousal details will exclude the three most recent years, as spousal variables are not released in the public files.

¹¹⁹Crossley et al. (2009) make use of a profession specific survey in their analysis of Canadian physician labour supply. While this survey, and other similar profession-specific surveys, provide more detailed information on labour supply (time spent directly with patients, time in administrative activities, etc.) it does not identify the incorporation status of the professional practice. The LFS is therefore the ideal tool to study this specific research question.

¹²⁰Employer-employee matched records, that include both T1 individual and T2 corporate tax records, would be needed to study this topic and these would not share any information on labour supply. The Record of Employment (ROE) has information on weeks of work, but this record does not cover unincorporated self-employed workers as they typically do not contribute to employment insurance.

(or industry) needed to match records to a specific profession.¹²¹ With the exception of the Census, the pooled LFS is the only micro-level dataset to provide a large enough sample to select on professional occupations, while retaining a sample large enough for provincial comparisons.¹²²

Professional occupations are identified using the publicly available occupation codes. These separately identify professional occupations in (1) business and finance, (2) natural and applied sciences, (3) health, and (4) social and legal professions.¹²³ I specifically exclude education professionals as they are unlikely to take up self-employment. In addition, I exclude semi-professional occupations (e.g., registered nurses and paralegals), and technical assistants. It should be noted that these occupation categories are relatively broad and, as such, may include individuals who are not treated by specific reforms. However, a benefit of these broader groupings is that it smooths over any occupation code changes over this three-decade period.

3.6 Rising incorporation among professionals

Over the three-decade period, from 1989-2019, the incorporation rate of self-employed professionals increased by 50%, from approximately 40% to 60% of all professional practices. Figure 3.1 demonstrates that this shift is marked by a break in trend beginning in the early 2000s. In the decade prior, the incorporation rate remained relatively stable. Notably, the shift towards incorporation has been remarkably gender neutral. This is significant, as the gender profile of professional workers has shifted dramatically over this period. The female share of non-education professionals has risen in a linear fashion from a third to a half over this three-decade period (see Figure 3.2). Although, there is heterogeneity across professions as women remain under-represented in the natural and applied sciences.

Among self-employed professionals, the rise in female representation has been of a similar magnitude; although, starting from a lower level ($\sim 20\%$) in early 1990s. This is in part because the rate at which professionals enter self-employment has remained fairly constant over this period and the gender difference has declined only marginally (see Figure

¹²¹Industry is available for wage income from a CCPC after 2000, but not unincorporated income. It is therefore hard to cleanly identify the incorporation of an existing business. Moreover, this research design requires a sample covering the 1990s as a number of reforms take place in the early 2000s. The Canadian Employer Employer Dynamics Database (CEEDD) is the ideal candidate to study how incorporation affects personal and household income structure, and possibly hiring. As with all Canadian tax-based administrative databases, the CEEDD does not contain information on labour supply. Unfortunately, the CEEDD has two major limitations. First, it starts in 2001 and, second, it does not have occupation/industry information for unincorporated workers; for the same reason as the LAD, which is that it is not a requirement of self-reported income.

¹²²Unfortunately, the Census (1991-2016) has not consistently recorded information on hours of labour supply.

¹²³Including, Judges, Lawyers, Psychologists, Social Workers, Ministers of Religion, and Policy and Programs

3.3).¹²⁴ The changing demographics of professionals is unlikely to explain the increase in incorporation; although it may be important for changes in outcomes such as hours of labour supply (see discussion in [Crossley et al., 2009](#)).

The precipitous rise in incorporation aligns with the collection of provincial level reforms, but also occurs at point in time when a number of provinces reduced their small-business corporate tax rate (see Figure 3.5). Corporate tax rates fell dramatically with the recession in 2008, but also fell at the provincial level in the early 2000s. Except for Quebec, Nova Scotia, Prince Edward Island, Newfoundland & Labrador, Canadian provinces lowered their small business corporate tax rates by as much as 10 %-points in the early 2000s. In 1989 the lowest small business rate ranged from 16.2%-22.84%, by 2019 this range had declined to 9%-15%; with Manitoba placing a zero marginal tax rate (9% with federal rate) on eligible small businesses from 2011.

However, the declining corporate tax rate is unlikely to explain the shift towards incorporation after 2000. Figure 3.6 shows the incorporation rate of professionals by province and occupation group, where the provinces are grouped by whether the province was an early or late adopter of professional corporation reforms. Recall, that the 'late adopters' include Quebec, Ontario, Saskatchewan, and Manitoba. Health professionals demonstrate the starkest rise in incorporation rate. Moreover, while incorporation increased among early adopters throughout the 1990's it was flat among late adopters until after 2000. By 2019, health professionals in late adopting provinces demonstrate a similar incorporation rate to those in early adopter provinces. In contrast, natural and applied science professionals - who were largely unaffected by these reforms and could incorporate in most provinces - demonstrate little difference in incorporation rate across provincial treatment group. Their incorporation rate has steadily increased in accordance with the falling tax rate.

For business and legal professionals, the evidence is mixed. There is evidence of an increase in incorporation among late-adopting provinces after 2004, but it is not as precipitous as health professionals. Moreover, as with the applied sciences, business and legal professionals in early adopting provinces also demonstrate a shift towards incorporation in the early 2000s; once again highlighting the potential contribution of the corporate tax rate.¹²⁵ However, the level difference between legal professionals in the late and early adopting provinces prior to the reform period suggests that there may be a significant

¹²⁴The self-employment rate of professionals rose during the 1990s but has since stabilized around 20% and for men in particular has declined. Notably, this decline in male self-employment is too small to explain the rapid change in incorporation rate. Instead, the decline in male self-employment is consistent with the analysis of Chapter 2 which shows that male self-employment in Canada increases with the value of income splitting, which began to decline around 1997 in Canada. In Figure 3.4 I show that the self-employment rate of professionals does not change with the reform. This further, demonstrates that the corporation reforms were not responsible for the decline in male self-employment.

¹²⁵If you refer to Table 3.1 you will see that at the professional level certain Atlantic provinces also adopters reforms during this period.

response. In contrast, this difference was much smaller among business professionals in the pre-reform period, suggesting that the subsequent response may be related to other factors. Regardless, there is clear evidence that the response to these reforms varied considerably at the occupation level. This treatment heterogeneity may reflect variation in income and practice size, or alternative ownership structures.¹²⁶ For this reason, I will provide all estimates at the occupation-specific level.

The event-study design discussed in Section 3.3 is well suited to disentangle these various factors. Panel A of Figure 3.7 plots the first-stage event-time coefficients for a pooled sample of business (and finance), health, and legal (and social) professionals. As accounting, medical, and legal professionals fall within these occupation groups the event-study should identify any response to the reforms. The estimating model underlying each of these first-stage regressions includes controls for covariates (including gender) and the provincial small business corporate tax rate.

Figure 3.7 plots the event-time coefficients for the pooled sample (which excludes the natural and applied sciences), alongside occupation-specific models. The pooled estimates identify a precipitous rise in incorporation after the reforms, with no evidence of a pre-trend. However, this aggregate result clouds considerable variation across occupations. There is no evidence of a response to the reforms affecting accounting professionals (Panel B), but a very large response among health professionals (Panel C). This is to be expected given evidence presented in Figure 3.6. The 40% rise in incorporation among these professionals is tied to the reforms and not the corporate tax rate. For legal professionals there is a weaker response. The event-study identifies a $\sim 12\%$ increase in incorporation relative to the base period before the reform. Henceforth, business and finance professionals will be excluded from any pooled specifications and in certain specifications will play the role of a ‘always treated’ control group together with natural and applied science professionals.

Table 3.2 provides the first-stage coefficients corresponding to the forthcoming IV models. As discussed, the pooled specification includes only health and legal professionals. In each of the occupation-specific figures, there is some evidence of a non-zero pre-trend. This does not suggest a failure of the parallel trends assumption, but is rather a result of the limited variation in the timing of treatment at the occupation level (i.e., limited number of provinces). For this reason, the pooled sample, which has more variation in the timing of treatment at both the province and occupation level, demonstrates no pre-trends (see both Table 3.2 and Figure 3.7).¹²⁷

¹²⁶For example, limited liability partnerships structures. These were first introduced in Texas in 1991 and largely advocated for by lawyers and accountants to protect partners from malpractice liability claims (Hamilton, 1994). They were introduced in Canada in the late 1990s: Ontario in 1998 and Alberta in 1999. Legislation in these provinces was limited to regulated professionals. More recently, BC introduced open access to LLPs in 2004.

¹²⁷As is, this pooled specification assumes that there is no occupation-specific heterogeneity in the take up

As indicated, the models in Table 3.2 do not include treatment unit FEs. Table 3.7 provides estimates that include controls for province FE by two means: models (1)-(3) include an ‘always treated’ control group of business and natural science professionals, while models (4)-(6) exclude the pre-trend. I include separate province and occupation FEs, as opposed to province-occupation FEs. While the pooled specifications remain reasonably unchanged, the two sets of occupation-specific models estimate higher take up for medical professionals and lower take up for legal professionals. Moreover, it does not appear that including a control group removes the pre-trend in models (2) and (3). This is not surprising given the earlier discussion in Section 3.4. In this setting, an ‘always treated’ control group is problematic, as firms can incorporate in response to other incentives, while firms in untreated occupations cannot. In Appendix B.3, I include a detailed discussion of alternative research designs: a difference-in-difference and linear break-in-trend first-stage.

3.7 Labour supply and hiring among self-employed professionals

This section focuses on the labour market outcomes of own labour supply and hiring. Own labour supply is measured by (usual) weekly hours of work and part-time status, while hiring is measured by an indicator for whether the self-employed individual has employees. Both outcomes demonstrate significant trends in Canada; particularly, through the 1990s in the lead up to the provincial level reforms in the early 2000s. As will be discussed, these trends can easily lead to misidentified results.

Figure 3.8 depicts the trend in each of these labour market outcomes by the group of early and late adopter provinces for health and legal professionals. In the early 1990s around 80% of these professional practices had employees. That number has declined to the point where only 40% of practices do. This trend extends to the business and natural science professions. Among health and legal professionals, the level of hiring was lower (at the extensive margin) among late adopter provinces in the pre-period, while that difference has shrunk since 2000.

A similar pattern is seen with labour supply. Average hours of self-employed professionals has fallen quickly since the early 1990s from peak of 50 hours a week to around 42 hours a week since the recession. See Figure 3.9 for a comparison of hours worked across paid and self-employed professionals, in the legal and health professions. The figure shows the decline in self-employed hours is comparable to that of paid employees, and that gender gap is larger among self-employed professionals.¹²⁸ The level of decline in hours depicted

incorporation. This is obviously a problematic assumption given the existing evidence. In the future, I intend to explore alternative specifications that allow for post-treatment heterogeneity by occupation.

¹²⁸This is consistent with the gender differences in the self-employed labour market discussed in Section 2.2 of Chapter 2.

by the LFS is directly comparable that found by [Crossley et al. \(2009\)](#) in their analysis of the Canadian Medical Association (CMA) Physician Resource Questionnaire (PRQ) Survey. The authors measure a decline in hours of direct patient care of 5hrs between 1981 and 2001. Likewise, the share of self-employed professionals who work part-time has risen from 6% to approximately 18%. As with hiring, there is some evidence that hours of work have equalized across the early and late adopting provinces since 2000, but less so for part-time status.

Table 3.3 provides OLS estimates of the impact of incorporation on labour supply (log of hours and part-time status) as well as hiring. The pattern is strikingly similar across both health and legal professionals. On average, incorporated business owners work 12% more hours, are around 9% less likely to work part-time, and 22% more likely to have employees. These results, are conditional on observed covariates, including the corporate tax rate. However, these results likely demonstrate the impact that being an employer or high income professional has on the decision to incorporate.

The IV results presented alongside these OLS estimates, reflect the local average treatment affect of those professionals who incorporated because of the reform. As the reforms had no direct bearing on a professional's labour supply or decision to hire employees, these estimates identify the causal affect that corporate structures have on labour market outcomes; either through a tax channel or a business operations channel (i.e., increased cash flow through retained earnings). Recall, that these responses reflect the labour supply response to a take-up of incorporation during event-times 2-10 and changes in labour market outcomes relative to the pre-period of event-time -1. The labour supply results are largely insignificant, suggesting that in the context of professionals, incorporation has little to no affect on labour supply. Thus, the considerable tax benefits that accrue to these workers through incorporation neither increases nor decreases labour supply.¹²⁹

Where there is evidence, it is in the area of hiring. Among compliers, incorporation is associated with a 12% increase in the probability of having at least one employee. The result is significantly larger for legal professionals. This hiring result appears to be driven by female professionals. Tables 3.4 and 3.5 repeat the same set of results for men and women separately. For men, the IV results are insignificant, despite the OLS results suggesting that incorporated male professionals are 17% more likely to have employees. For women, incorporation has a much stronger correlation with hiring. On average, incorporated female professionals are 31% more likely to take on employees. The IV results suggest that this is a partially a causal relationship between business structure and hiring. Hiring increased by 33% among compliers. The local average treatment effect is larger for female, legal

¹²⁹For legal professionals the coefficient on hours of work is large, negative, and significant at the 10% level. However, has the same sign to the part-time status coefficient. This result appears to be driven by specific outlier results and is not a robust finding.

professionals. However, the difference between the marginally significant male coefficient for legal professionals in Table 3.4 and the female coefficient in Table 3.5 is of a similar magnitude to the medical professional result.

Figure 3.10 depicts the trend in hiring and hours worked by early and late adopter provinces, separately by gender. A comparison of panels A and B clearly demonstrates that there was a larger hiring differential (between early and late adopters) for women compared to men. The figure also demonstrates why IV results do not find a significant impact of incorporation on hours worked. For women, average hours converged prior to the onset of these reforms. For male professionals one might argue that there has been an increase in hours worked. Some alternative specifications discussed in Appendix B.3 support this claim.

Table 3.6 provides some additional evidence that incorporation matters for the hiring practices of women, but not men. In place of the IV model, I estimate a simple OLS regression of each labour market outcome on incorporation status interacted with gender and control for province FEs. I estimate this model separately for the periods 1995-1999 and 2008-2012. These two periods bookend the period of health and legal reforms in Quebec, Ontario, and a few additional smaller provinces. For men, the relationship between incorporation and each labour market outcome changes before and after the treatment period. This suggests a change in selection: with the reform high earning professionals who work longer hours and are more likely to have employees select into incorporation. Hence, the increase in male incorporation coefficients for hours of work, part-time status and hiring.

For women, there is a similar increase in the size of the coefficient for hours of work and part-time status. Suggesting that women working longer hours (i.e., higher earners) were more likely to incorporate after the reform. However, for hiring the coefficient remains stable (with the exception legal professionals, for whom the coefficient declines). In the pre-period incorporated female professionals were 34% more likely to have employees relative to unincorporated women, while after the reforms the number is 33%. This could be driven by a change in the composition of female professionals, driving selection in the opposite direction and counteracting the above selection channel. However, this seems unlikely when the hours of work coefficients change for women in the same way as for men. Instead, the stability of the hiring result suggests that this correlation reflects, in part, a causal relationship between incorporation status and hiring.¹³⁰ This result is consistent

¹³⁰Unfortunately, the sample is too small to examine sub-populations (e.g., age groups) more closely by gender. Future research on this topic using administrative records that match individual and corporate tax files could identify this relationship more clearly. For example, the Canadian Employer Employee Dynamics Dataset would provide a closer examination of this topic. One would not be able to replicate this event-study using the CEEDD, because it begins in 2000. However, one could use the delayed reforms to health care professionals in

with the gender differences in risk taking (Bertrand, 2011; Charness and Gneezy, 2012; Eckel and Grossman, 2008). Hiring staff increases the operating costs of a business in the midst of uncertain revenue. Incorporation can assist to ensure against this risk through retained earnings. The lower corporate tax rate increases the stock of retained earnings, which can then be set aside to ensure against fluctuations in revenue.

Despite the graphical evidence that hiring increased in treated areas among women, alternative specifications do not all corroborate this finding. Table 3.8 presents IV results based on the first-stage specifications in Table 3.7 that include province FEs. The hiring result is negative in models (1)-(3) which combines the sample of men and women; although for women alone it is significant and positive ($\sim 22\%$).¹³¹ Models (4)-(6) estimate insignificant results at the occupation-specific level, but a positive result in the pooled sample.¹³²

The results presented in Appendix B.3 are also mixed. In general, these specifications are less flexible. The difference-in-difference specification incorporates the same ‘always treated’ control group as models (1)-(3) in Table 3.7 and 3.8. Here too, the hiring effect is negative. This may be because of differences in the occupation-specific trends of these outcomes. All occupations demonstrate a decline in both hours of work and hiring over this period, but there is significant variation in this trend across occupation groups that does not match within occupation variation based on event-time (see Figure 3.11 for hiring status). For example, it is evident from Figure 3.11 that trend in hiring among the ‘always treated’ professions is different to that among the ‘early adopters’ who are also treated throughout the sample period. Indeed, this figure demonstrates the within occupation variation used to estimate the positive relationship reported in Table 3.3. Year FEs in specifications that include a control group may be mis-specified. As argued in Appendix B.3, more work is needed to find a robust research design for this setting. In future research I intend to try and account for the introduction of Limited Liability Partnerships, which may have played a role in the lower estimated response among legal and, even, accounting professionals.

3.8 Conclusion

Limited liability corporate structures are commonly seen as an essential institution in the promotion of optimal entrepreneurial risk taking. Similarly, small business tax credits are deemed essential to ensure for the health of SMEs. However, there is a growing understanding that these tax-preferred legal structures further exacerbate the growing inequalities in

Quebec and legal professionals in Ontario to estimate a more restricted model.

¹³¹Results not shown here. For men, the coefficient is of a similar magnitude, but negative.

¹³²Separately by gender, the coefficient on hiring is larger and significant in the pooled sample alone. Again, the difference between men and women is around 0.3; as in Tables 3.4 and 3.5. The own labour supply coefficients in Table 3.8 are significant in the pooled specification and suggest an increase in labour supply.

our society, as they disproportionately benefit those in the top end of the income distribution who are more likely to own such businesses (Alstadsæter et al., 2019; Smith et al., 2019; Wolfson et al., 2016). From this perspective, the establishment of *professional corporations* is simply tax expenditure favouring wealthy professionals (Wolfson and Legree, 2015). While this certainly might be the case, we should not overlook the possibility that there are potential labour market gains - and public health gains in the context of health professionals - in the form of increased hours worked. Afterall, rising top marginal tax rates during the early 1990s may be a factor that partially explains the decline in professional labour supply during this period (Crossley et al., 2009). The results of this analysis suggest not. Providing high income professionals with the significant tax benefits of incorporation does not change their labour supply.

In addition, there are real business planning benefits to incorporation that should not be overlooked. While the tax-deferred benefits of retained corporate earnings can be used to generate additional passive investment income, they can also be a vital source of funding for real capital investment and increase cash flow needed to hire staff. This paper explores whether the adoption of corporate structures has had any impact on the labour market behaviour of self-employed professionals: their own labour supply and the decision to hire staff. I find no evidence that incorporation has any impact on the labour supply of health and legal professionals, suggesting that any measured increase in after-tax income can be attributed to tax savings (Wolfson and Legree, 2015). However, for female professionals I do find evidence that incorporation increases the likelihood that self-employed professional will employ at least one individual. This is consistent with the reduction in risk associated with increased retained earnings. Moreover, the gendered nature of this result is consistent with views of risk-taking behaviour in the labour market and entrepreneurship (Bertrand, 2011; Cowling and Taylor, 2001). Further evidence of such a result should be sought in more detailed administrative records.

3.9 Tables and Figures

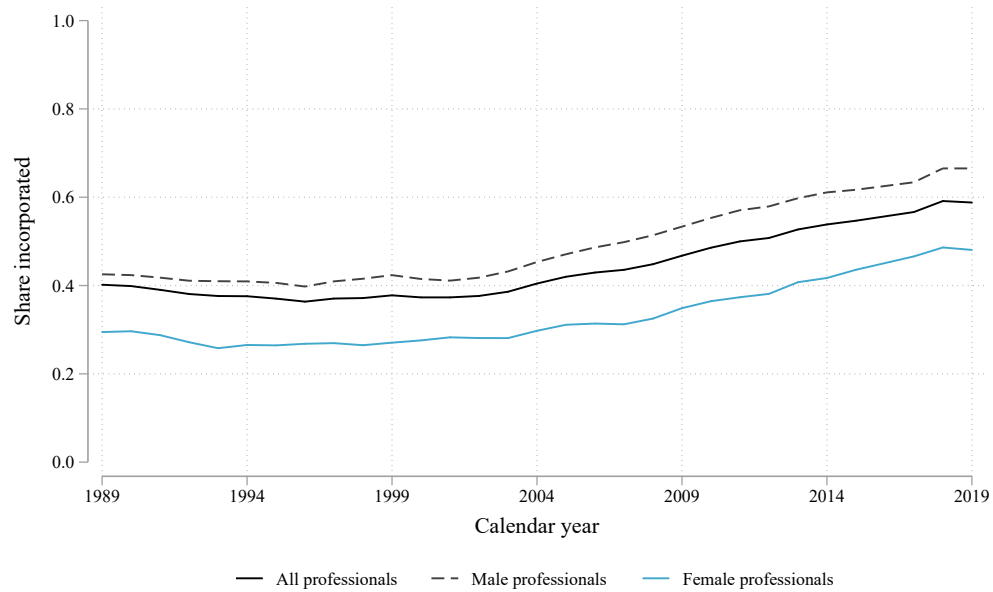


Figure 3.1: Trend: Incorporation rate of self-employed professionals, by gender.

Note: Constructed using Canadian Labour Force Survey (1989-2019). Sample includes all non-education self-employed professionals, aged 25-59. This trend displays the 3-year moving average and is constructed without sampling weights.

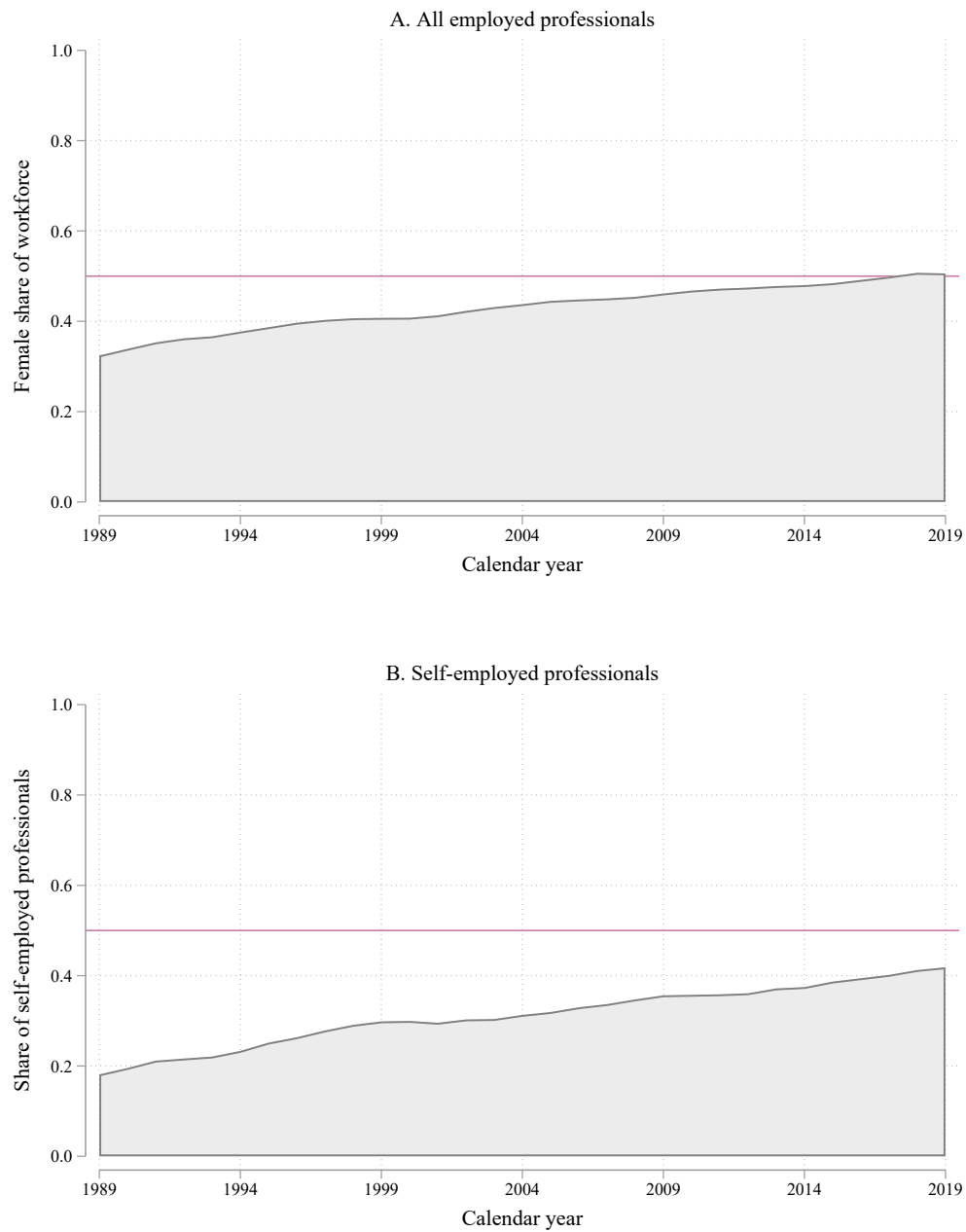


Figure 3.2: Female share of non-education professional employees.

Note: Constructed using Canadian Labour Force Survey (1990-2015). Sample includes all non-education professionals, aged 25-59. Trend constructed without sampling weights. Panel A plots the self-employment rate (share of employed professionals), while Panel B the incorporation rate (share of self-employed professionals).

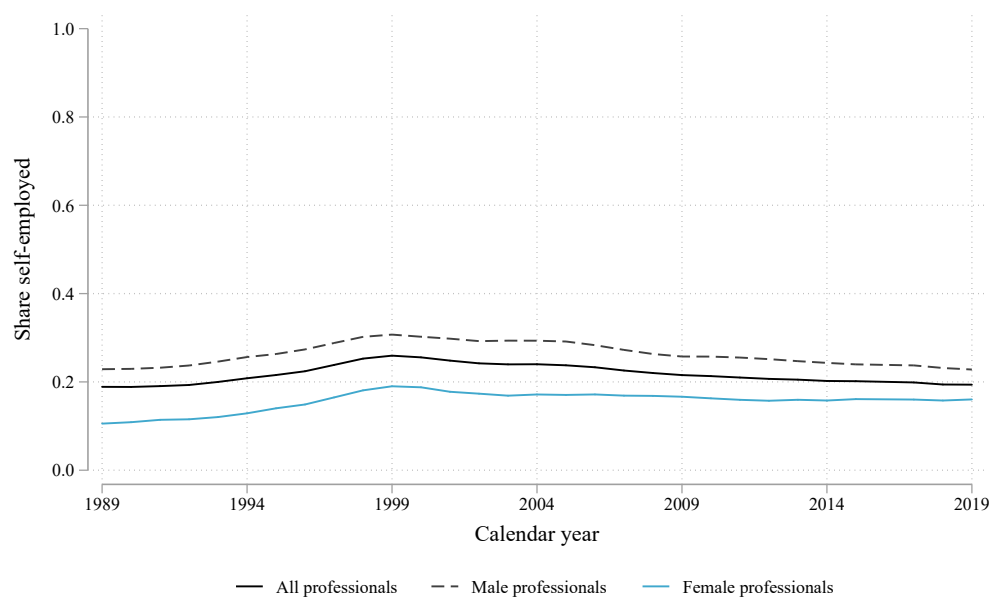


Figure 3.3: Trend: Self-employment rate of professionals, by gender.

Note: Constructed using Canadian Labour Force Survey (1989-2019). Sample includes all employed professionals, aged 25-59. This trend displays the 3-year moving average and is constructed without sampling weights.

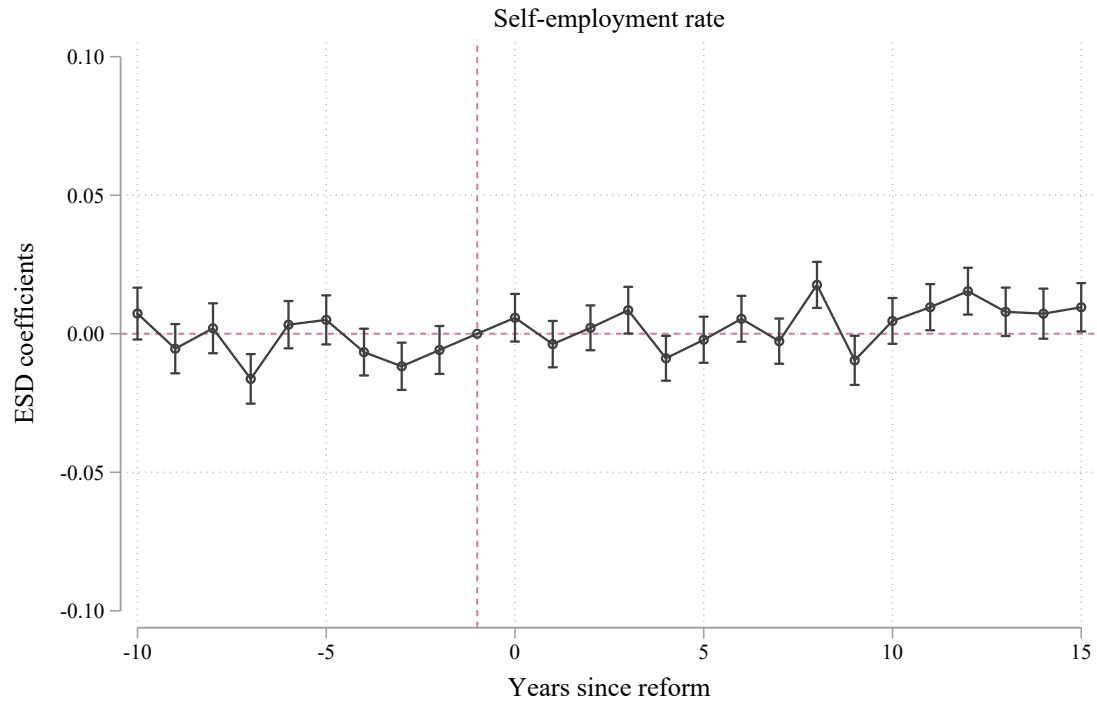


Figure 3.4: Event-study coefficients from model examining selection into self-employment.

Note: The figure displays the event-study coefficients associated with the first stage estimated using Canadian Labour Force Survey (1989-2019). Sample includes all employed professionals, aged 25-59, in the business, health, and legal professions. The model is estimated without sampling weights and includes covariates: sex, age, education, marital status, age of youngest child, industry, and corporate tax rate. In addition to year FEs, it includes occupation FEs, but not province FEs or province-occupation FEs.

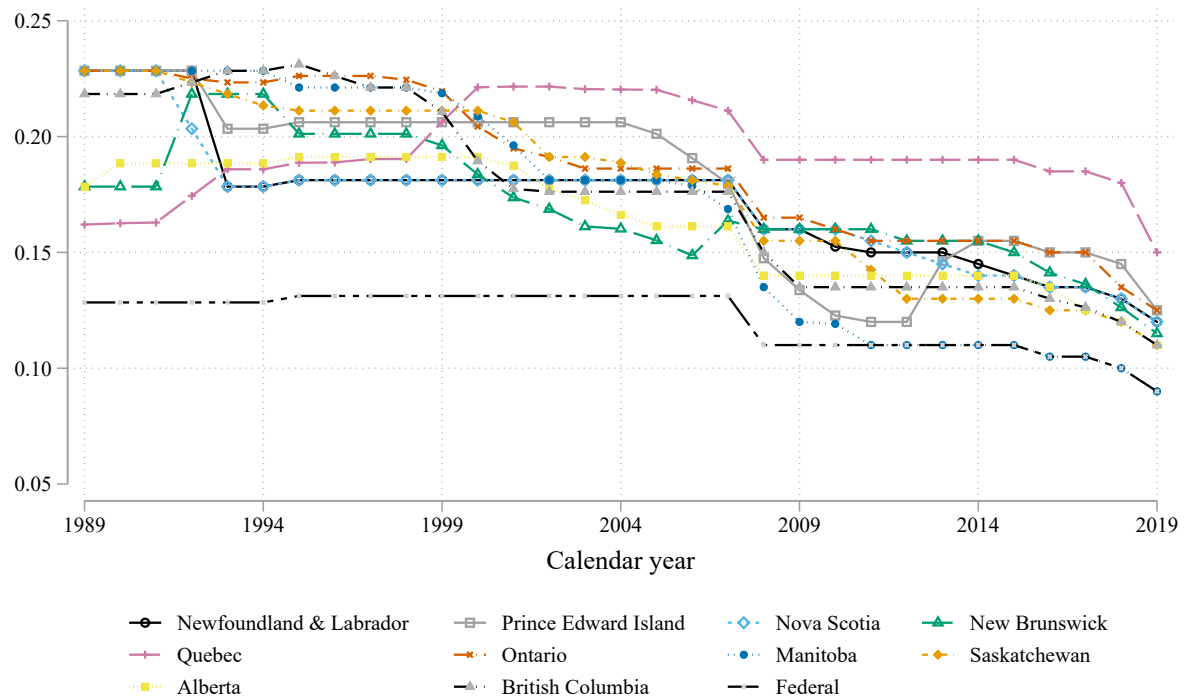


Figure 3.5: Canadian corporate tax rate, including small business tax credit.

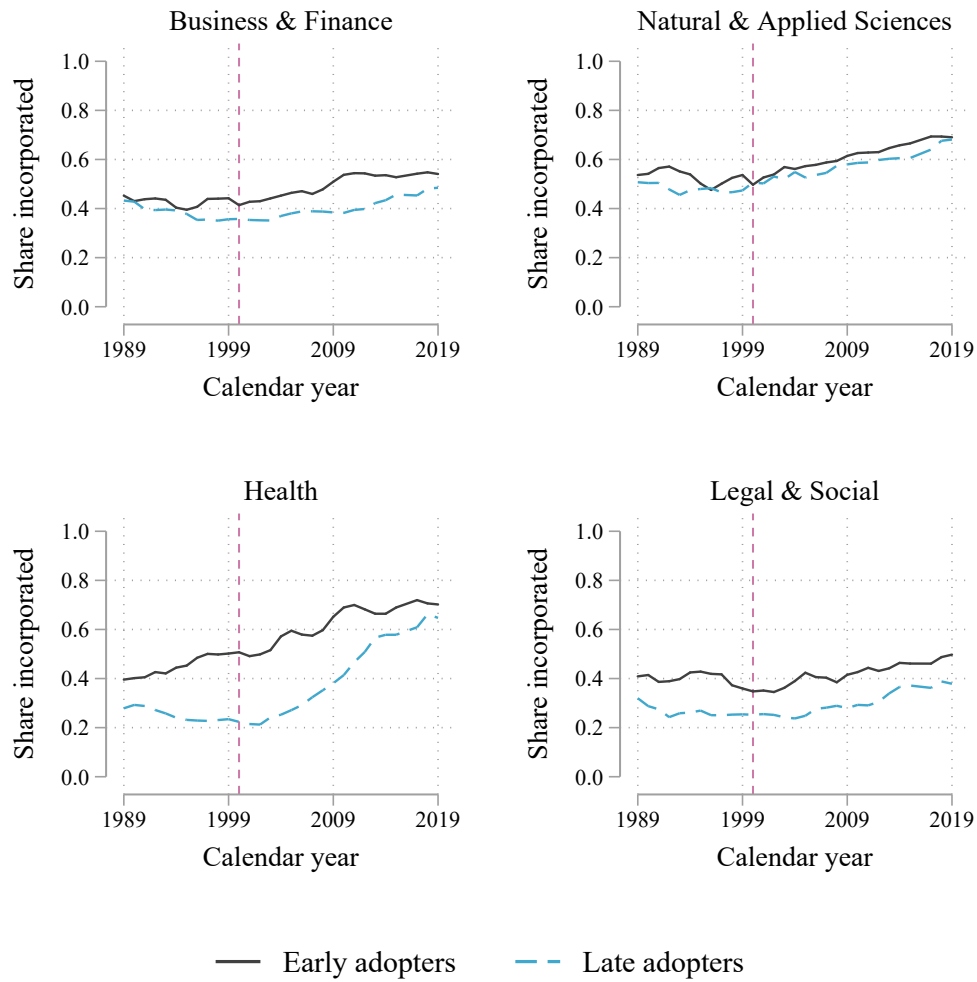


Figure 3.6: Trend: Incorporation by professional group and provincial treatment group.

Note: Constructed using Canadian Labour Force Survey (1989-2019). Sample includes all employed professionals, aged 25-59. This trend displays the 3-year moving average and is constructed without sampling weights. Late adopter provinces include Manitoba, Saskatchewan, Ontario, and Quebec. The vertical line denotes the year 2000, the year before the 2001 reform period.

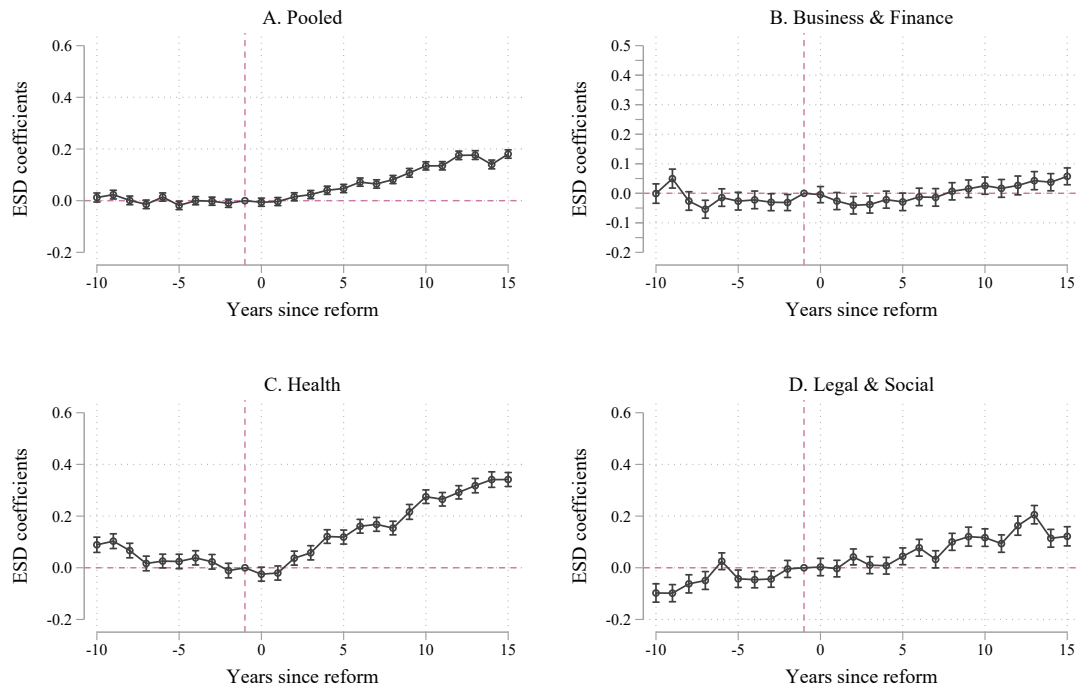


Figure 3.7: Event-study coefficients from first stage: incorporation.

Note: The figure displays the event-study coefficients associated with the first stage estimated using Canadian Labour Force Survey (1989-2019). Sample includes all employed professionals, aged 25-59, in the business, health, and legal professions. The model is estimated without sampling weights and includes covariates: sex, age, education, marital status, age of youngest child, industry, and corporate tax rate. In addition to year FEs, it includes occupation FEs, but not province FEs or province-occupation FEs.

Table 3.2: First stage by occupation based on event-study research design

	Without covariates			With covariates		
	(1) Pooled	(2) Health	(3) Legal	(4) Pooled	(5) Health	(6) Legal
Event-time:						
s=-5	-0.0151 (0.0109)	0.0245* (0.0148)	-0.0543*** (0.0173)	-0.0201* (0.0106)	0.0244* (0.0143)	-0.0425** (0.0171)
s=-4	-0.00202 (0.0106)	0.0435*** (0.0146)	-0.0634*** (0.0164)	-0.00168 (0.0102)	0.0384*** (0.0140)	-0.0460*** (0.0161)
s=-3	0.00419 (0.0107)	0.0232 (0.0150)	-0.0485*** (0.0168)	0.00284 (0.0103)	0.0232 (0.0144)	-0.0434*** (0.0163)
s=-2	-0.0300*** (0.0110)	-0.0287* (0.0150)	-0.0200 (0.0173)	-0.0141 (0.0106)	-0.0109 (0.0144)	-0.00450 (0.0168)
s=0	-0.0142 (0.0108)	-0.0243* (0.0143)	-0.00435 (0.0176)	-0.0119 (0.0104)	-0.0247* (0.0138)	0.00295 (0.0171)
s=1	0.00525 (0.0106)	-0.00559 (0.0145)	0.00146 (0.0168)	-0.00350 (0.0102)	-0.0202 (0.0139)	-0.00312 (0.0164)
s=2	0.0411*** (0.0104)	0.0521*** (0.0144)	0.0320** (0.0162)	0.0390*** (0.0100)	0.0372*** (0.0138)	0.0420*** (0.0157)
s=3	0.0271** (0.0108)	0.0520*** (0.0146)	-0.0118 (0.0173)	0.0394*** (0.0104)	0.0580*** (0.0140)	0.0102 (0.0168)
s=4	0.0664*** (0.0103)	0.115*** (0.0139)	0.00446 (0.0170)	0.0722*** (0.00998)	0.121*** (0.0134)	0.00801 (0.0165)
s=5	0.0850*** (0.0106)	0.124*** (0.0143)	0.0313* (0.0171)	0.0888*** (0.0103)	0.119*** (0.0138)	0.0444*** (0.0166)
s=6	0.120*** (0.0105)	0.167*** (0.0142)	0.0674*** (0.0170)	0.121*** (0.0101)	0.161*** (0.0136)	0.0776*** (0.0166)
s=7	0.107*** (0.0106)	0.173*** (0.0142)	0.0175 (0.0173)	0.111*** (0.0102)	0.168*** (0.0136)	0.0327* (0.0169)
s=8	0.126*** (0.0104)	0.155*** (0.0139)	0.0804*** (0.0172)	0.136*** (0.0101)	0.154*** (0.0135)	0.100*** (0.0168)
s=9	0.169*** (0.0112)	0.210*** (0.0149)	0.0985*** (0.0193)	0.184*** (0.0110)	0.216*** (0.0147)	0.120*** (0.0188)
s=10	0.194*** (0.0105)	0.273*** (0.0139)	0.0992*** (0.0176)	0.201*** (0.0102)	0.275*** (0.0134)	0.117*** (0.0173)
Corporate rate				-0.0585 (0.0642)	-0.323*** (0.0893)	0.450*** (0.110)
Province FE	No	No	No	No	No	No
Covariates	No	No	No	Yes	Yes	Yes
N	176433	111892	64541	172922	109497	63425

Note: Estimated using Canadian Labour Force Survey (1987-2019). Sample includes either health or legal self-employed professionals, aged 25-59. The table reports the event-time coefficients and coefficient on the corporate tax rate. As standard, all models include year FEs. Models are estimated without sampling weights and when indicated include covariates: sex, age, education, marital status, age of youngest child, industry, and corporate tax rate. As standard, all models include year FEs. Estimates that pool professionals also include occupation FEs.

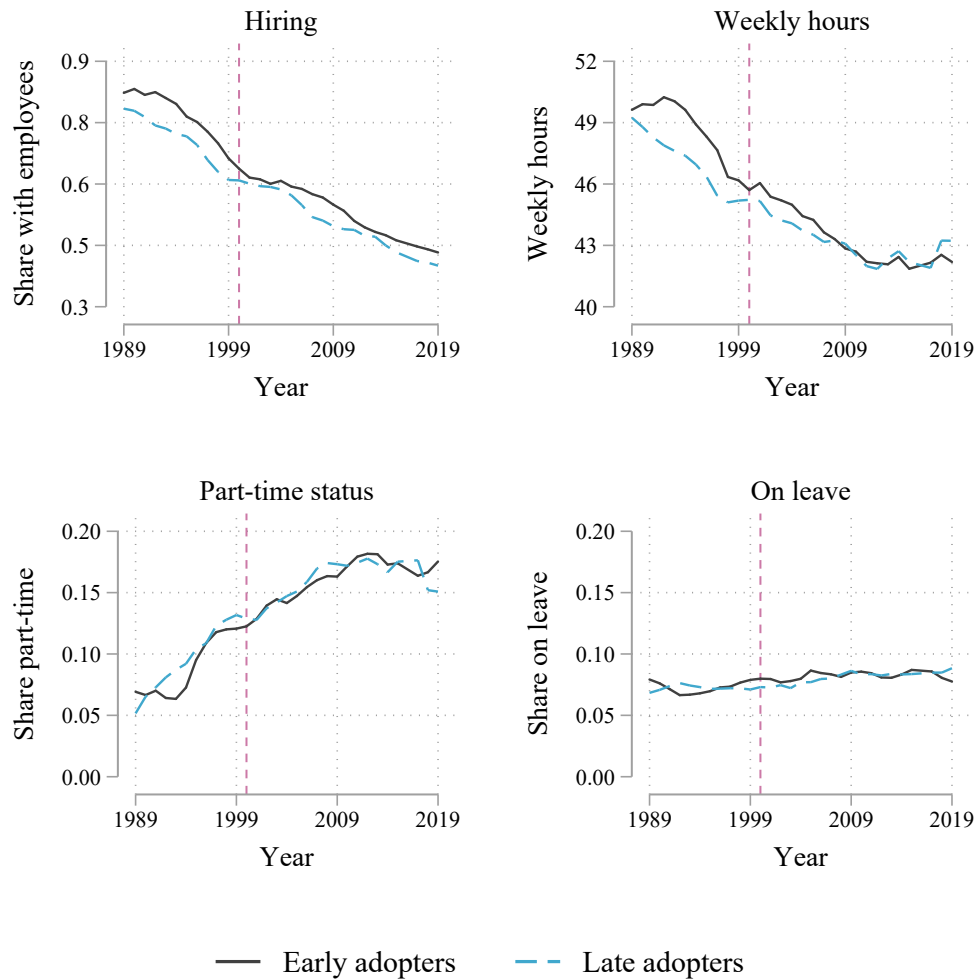


Figure 3.8: Trend: Outcome variables by provincial treatment group.

Note: Estimated using Canadian Labour Force Survey (1987-2019). Sample includes all health and legal professionals, aged 25-59. This trend displays the 3-year moving average and is constructed without sampling weights. Late adopter provinces include Manitoba, Saskatchewan, Ontario, and Quebec. The vertical line denotes the year 2000, the year before the 2001 reform period.

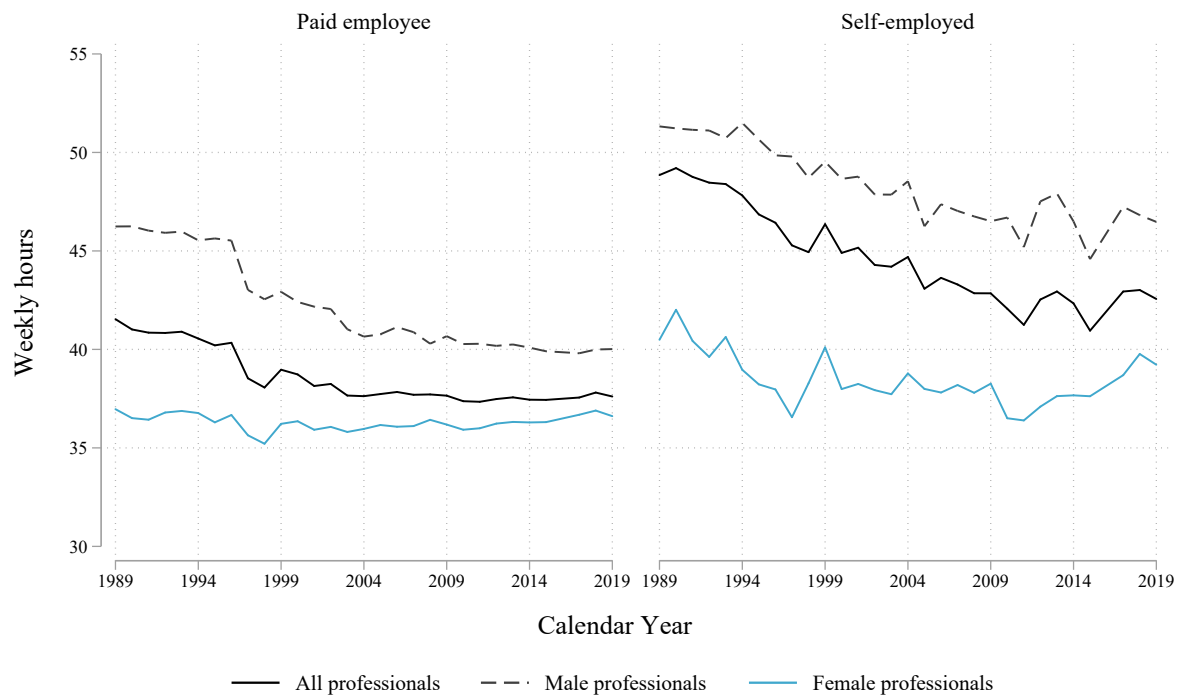


Figure 3.9: Trend: Average hours worked by gender: paid employees compared self-employed professionals.

Note: Constructed using Canadian Labour Force Survey (1987-2019). Sample includes all self-employed health and legal professionals, aged 25-59. This trend displays the 3-year moving average and is constructed without sampling weights.

Table 3.3: Labour supply and hiring: OLS and IV results based on event-study first stage

	OLS			IV		
	(1) Pooled	(2) Health	(3) Legal	(4) Pooled	(5) Health	(6) Legal
Panel A. Log of weekly hours						
Incorporated	0.119*** (0.00235)	0.111*** (0.00272)	0.136*** (0.00435)	-0.0431 (0.0388)	-0.0132 (0.0362)	-0.203* (0.107)
Province FE	No	No	No	No	No	No
<i>N</i>	172922	109497	63425	172922	109497	63425
C-D Wald F-stat				75.26	74.19	12.95
Sargan P-value				0.099	0.047	0.040
Panel B. Part-time status						
Incorporated	-0.0872*** (0.00169)	-0.0882*** (0.00204)	-0.0863*** (0.00296)	-0.00870 (0.0277)	0.0204 (0.0273)	-0.111 (0.0698)
Province FE	No	No	No	No	No	No
<i>N</i>	172922	109497	63425	172922	109497	63425
C-D Wald F-stat				75.26	74.19	12.95
Sargan P-value				0.000	0.000	0.000
Panel C. Hiring status						
Incorporated	0.222*** (0.00224)	0.240*** (0.00282)	0.196*** (0.00369)	0.116*** (0.0367)	0.0682* (0.0377)	0.309*** (0.0875)
Province FE	No	No	No	No	No	No
<i>N</i>	172922	109497	63425	172922	109497	63425
C-D Wald F-stat				75.26	74.19	12.95
Sargan P-value				0.000	0.000	0.006

Note: Constructed using Canadian Labour Force Survey (1989-2019). Sample includes all health and legal self-employed professionals, aged 25-59. Models are estimated without sampling weights and include covariates listed in Table 3.2.

Table 3.4: Men: Labour supply and hiring: OLS and IV results based on event-study first stage

	OLS			IV		
	(1) Pooled	(2) Health	(3) Legal	(4) Pooled	(5) Health	(6) Legal
Panel A. Log of weekly hours						
Incorporated	0.0630*** (0.00229)	0.0500*** (0.00274)	0.0886*** (0.00406)	0.0451 (0.0358)	0.0387 (0.0324)	-0.192* (0.0983)
Province FE	No	No	No	No	No	No
<i>N</i>	106606	67746	38860	106606	67746	38860
C-D Wald F-stat				51.04	57.90	8.462
Sargan P-value				0.020	0.491	0.263
Panel B. Part-time status						
Incorporated	-0.0418*** (0.00152)	-0.0356*** (0.00181)	-0.0524*** (0.00272)	-0.0166 (0.0239)	-0.00172 (0.0215)	-0.0209 (0.0623)
Province FE	No	No	No	No	No	No
<i>N</i>	106606	67746	38860	106606	67746	38860
C-D Wald F-stat				51.04	57.90	8.462
Sargan P-value				0.062	0.039	0.263
Panel C. Hiring status						
Incorporated	0.174*** (0.00274)	0.181*** (0.00341)	0.162*** (0.00458)	0.0445 (0.0433)	0.0102 (0.0410)	0.217** (0.105)
Province FE	No	No	No	No	No	No
<i>N</i>	106606	67746	38860	106606	67746	38860
C-D Wald F-stat				51.04	57.90	8.462
Sargan P-value				0.000	0.000	0.000

Note: Constructed using Canadian Labour Force Survey (1989-2019). Sample includes all health and legal self-employed professionals, aged 25-59. Models are estimated without sampling weights and include covariates listed in Table 3.2.

Table 3.5: Women: Labour supply and hiring: OLS and IV results based on event-study first stage

	OLS			IV		
	(1) Pooled	(2) Health	(3) Legal	(4) Pooled	(5) Health	(6) Legal
Panel A. Log of weekly hours						
Incorporated	0.214*** (0.00492)	0.209*** (0.00552)	0.221*** (0.00962)	-0.0935 (0.0875)	-0.0467 (0.0833)	-0.131 (0.203)
Province FE	No	No	No	No	No	No
<i>N</i>	66316	41751	24565	66316	41751	24565
C-D Wald F-stat				25.23	22.06	6.545
Sargan P-value				0.000	0.000	0.000
Panel B. Part-time status						
Incorporated	-0.164*** (0.00371)	-0.171*** (0.00444)	-0.151*** (0.00665)	-0.00430 (0.0649)	0.0429 (0.0671)	-0.182 (0.137)
Province FE	No	No	No	No	No	No
<i>N</i>	66316	41751	24565	66316	41751	24565
C-D Wald F-stat				25.23	22.06	6.545
Sargan P-value				0.000	0.000	0.000
Panel C. Hiring status						
Incorporated	0.310*** (0.00383)	0.336*** (0.00487)	0.267*** (0.00618)	0.326*** (0.0660)	0.273*** (0.0717)	0.542*** (0.132)
Province FE	No	No	No	No	No	No
<i>N</i>	66316	41751	24565	66316	41751	24565
C-D Wald F-stat				25.23	22.06	6.545
Sargan P-value				0.000	0.000	0.004

Note: Constructed using Canadian Labour Force Survey (1989-2019). Sample includes all health and legal self-employed professionals, aged 25-59. Models are estimated without sampling weights and include covariates listed in Table 3.2.

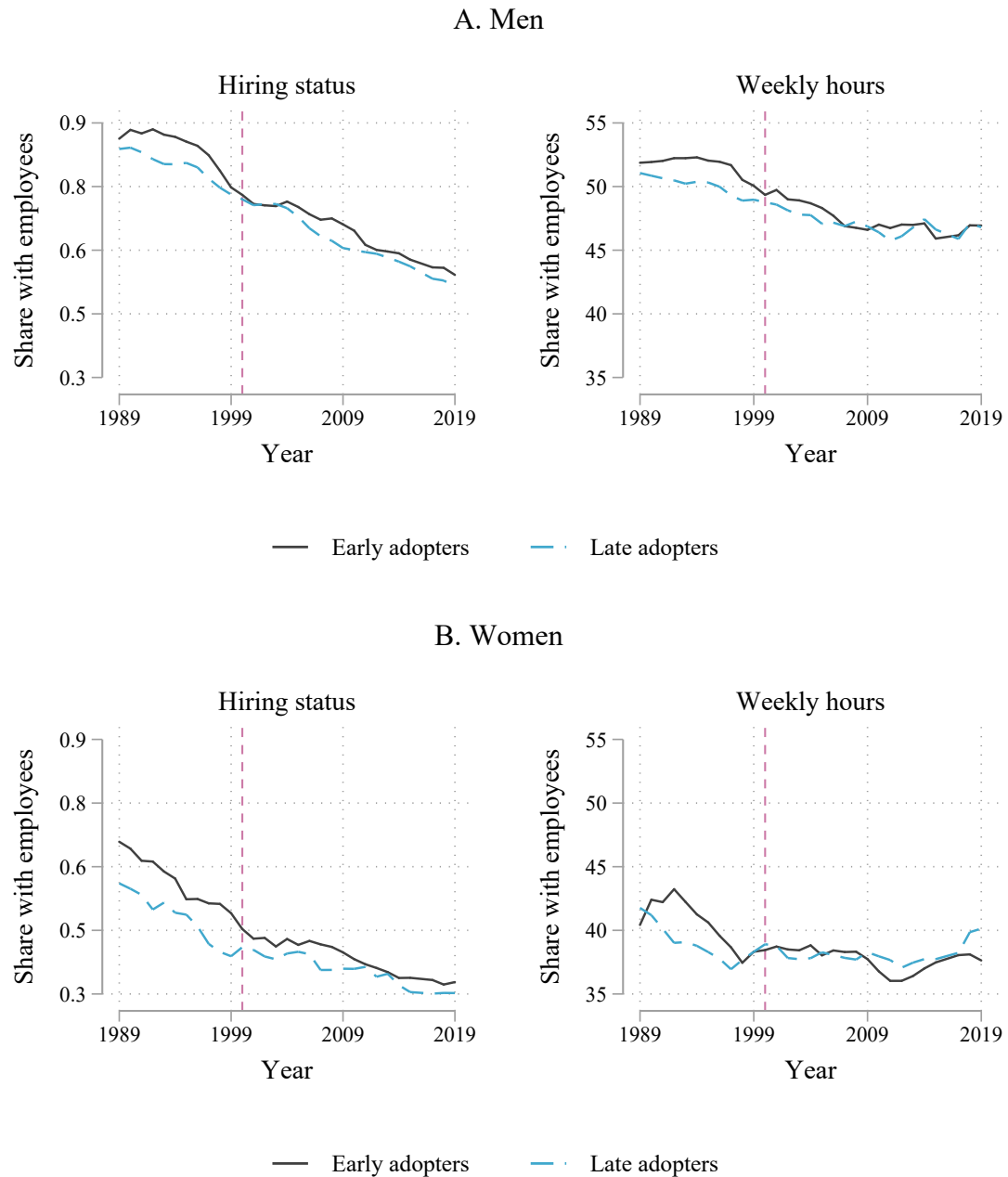


Figure 3.10: Trend: Outcome variables by professional group and provincial treatment group.

Note: Constructed using Canadian Labour Force Survey (1989-2019). Sample includes all health and legal self-employed professionals, aged 25-59. This trend displays the 3-year moving average and is constructed without sampling weights. Late adopter provinces include Manitoba, Saskatchewan, Ontario, and Quebec. The vertical line denotes the year 2000, the year before the 2001 reform period.

Table 3.6: Labour supply and hiring: gender differences in OLS before and after treatment spells

	1995-1999			2008-2012		
	(1) Pooled	(2) Health	(3) Legal	(4) Pooled	(5) Health	(6) Legal
Panel A. Log of weekly hours						
Incorporated (Male)	0.0484*** (0.00696)	0.0212** (0.00851)	0.0973*** (0.0118)	0.129*** (0.00776)	0.124*** (0.00888)	0.136*** (0.0148)
Incorporated (Female)	0.170*** (0.0110)	0.154*** (0.0133)	0.192*** (0.0189)	0.237*** (0.00875)	0.223*** (0.00950)	0.257*** (0.0182)
Female	-0.337*** (0.00676)	-0.347*** (0.00822)	-0.276*** (0.0117)	-0.287*** (0.00778)	-0.272*** (0.00924)	-0.269*** (0.0138)
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	28233	17051	11182	29110	18597	10513
Panel B. Part-time status						
Incorporated (Male)	-0.0307*** (0.00481)	-0.0311*** (0.00606)	-0.0353*** (0.00778)	-0.0754*** (0.00581)	-0.0724*** (0.00710)	-0.0772*** (0.0102)
Incorporated (Female)	-0.131*** (0.00760)	-0.132*** (0.00947)	-0.129*** (0.0125)	-0.186*** (0.00656)	-0.182*** (0.00760)	-0.190*** (0.0126)
Female	0.228*** (0.00467)	0.235*** (0.00586)	0.180*** (0.00773)	0.218*** (0.00583)	0.218*** (0.00739)	0.186*** (0.00955)
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	28233	17051	11182	29110	18597	10513
Panel C. Hiring status						
Incorporated (Male)	0.172*** (0.00654)	0.181*** (0.00826)	0.165*** (0.0105)	0.221*** (0.00728)	0.218*** (0.00938)	0.190*** (0.0118)
Incorporated (Female)	0.344*** (0.0103)	0.362*** (0.0129)	0.327*** (0.0169)	0.331*** (0.00822)	0.364*** (0.0100)	0.274*** (0.0145)
Female	-0.253*** (0.00635)	-0.267*** (0.00799)	-0.195*** (0.0105)	-0.193*** (0.00730)	-0.234*** (0.00976)	-0.127*** (0.0110)
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	28233	17051	11182	29110	18597	10513

Note: Constructed using Canadian Labour Force Survey (1989-2019). Sample includes all health and legal self-employed professionals, aged 25-59. Models are estimated without sampling weights and include covariates listed in Table 3.2.

Table 3.7: First stage by occupation based on event-study research design with treatment unit FEs

	With a control group			Without a pre-trend		
	(1) Pooled	(2) Health	(3) Legal	(4) Pooled	(5) Health	(6) Legal
Event-time:						
s=-5	-0.00752 (0.0103)	0.0265** (0.0135)	-0.0366** (0.0159)			
s=-4	0.0165 (0.0102)	0.0542*** (0.0135)	-0.0254 (0.0157)			
s=-3	0.00813 (0.0102)	0.0339** (0.0139)	-0.0350** (0.0153)			
s=-2	-0.00171 (0.0103)	-0.00193 (0.0138)	0.00322 (0.0156)			
s=0	-0.0174* (0.0102)	-0.0144 (0.0132)	-0.00927 (0.0160)	-0.00745 (0.00795)	-0.00823 (0.0112)	0.0201 (0.0136)
s=1	0.0000479 (0.0101)	0.00464 (0.0134)	-0.00115 (0.0157)	-0.000954 (0.00804)	0.00487 (0.0118)	-0.00319 (0.0135)
s=2	0.0304*** (0.00995)	0.0423*** (0.0132)	0.0206 (0.0155)	0.0416*** (0.00798)	0.0711*** (0.0121)	0.0367*** (0.0137)
s=3	0.0387*** (0.0102)	0.0812*** (0.0134)	-0.0129 (0.0158)	0.0451*** (0.00829)	0.103*** (0.0127)	0.00657 (0.0145)
s=4	0.0687*** (0.0100)	0.139*** (0.0133)	-0.0188 (0.0157)	0.0765*** (0.00842)	0.180*** (0.0132)	-0.00857 (0.0150)
s=5	0.0769*** (0.0100)	0.128*** (0.0133)	0.0134 (0.0154)	0.0920*** (0.00855)	0.179*** (0.0139)	0.0222 (0.0150)
s=6	0.114*** (0.00993)	0.179*** (0.0130)	0.0360** (0.0156)	0.126*** (0.00856)	0.232*** (0.0144)	0.0409*** (0.0156)
s=7	0.103*** (0.0101)	0.190*** (0.0132)	-0.00833 (0.0161)	0.115*** (0.00890)	0.252*** (0.0151)	-0.00702 (0.0166)
s=8	0.117*** (0.00984)	0.171*** (0.0128)	0.0485*** (0.0155)	0.141*** (0.00879)	0.249*** (0.0157)	0.0562*** (0.0167)
s=9	0.174*** (0.0109)	0.253*** (0.0139)	0.0591*** (0.0180)	0.191*** (0.00993)	0.327*** (0.0174)	0.0395** (0.0191)
s=10	0.183*** (0.00987)	0.287*** (0.0127)	0.0405** (0.0158)	0.205*** (0.00927)	0.377*** (0.0170)	0.0523*** (0.0182)
Corporate rate	-0.303*** (0.0493)	-0.460*** (0.0575)	-0.305*** (0.0606)	-0.138** (0.0657)	-0.0381 (0.0937)	0.238** (0.110)
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
N	365078	301653	255581	172922	109497	63425

Note: Estimated using Canadian Labour Force Survey (1987-2019). Sample includes non-education, self-employed professionals, aged 25-59. Models (1)-(3) include a control group of business and natural science professionals, while models (4)-(6) exclude pre-trends. The table reports the event-time coefficients and coefficient on the corporate tax rate. Models are estimated without sampling weights and include covariates listed in Table 3.2.

Table 3.8: Labour supply and hiring: IV results based on event-study first stage with treatment unit FEs

	With a control group			Without a pre-trend		
	(1) Pooled	(2) Health	(3) Legal	(4) Pooled	(5) Health	(6) Legal
Panel A. Log of weekly hours						
Incorporated	-0.0507 (0.0418)	-0.0368 (0.0349)	-0.0518 (0.154)	0.0865** (0.0419)	0.0156 (0.0416)	-0.0757 (0.231)
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	365078	301653	255581	172922	109497	63425
C-D Wald F-stat	68.08	93.78	5.315	113.5	100.4	4.777
Sargan P-value	0.125	0.132	0.105	0.343	0.356	0.251
Panel B. Part-time status						
Incorporated	-0.0197 (0.0292)	0.0177 (0.0249)	-0.191* (0.106)	-0.116*** (0.0301)	-0.0901*** (0.0310)	-0.173 (0.155)
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	365078	301653	255581	172922	109497	63425
C-D Wald F-stat	68.08	93.78	5.315	113.5	100.4	4.777
Sargan P-value	0.000	0.000	0.000	0.015	0.004	0.006
Panel C. Hiring status						
Incorporated	-0.118*** (0.0401)	-0.129*** (0.0343)	-0.0993 (0.140)	0.436*** (0.0407)	-0.000248 (0.0440)	0.00735 (0.196)
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	365078	301653	255581	172922	109497	63425
C-D Wald F-stat	68.08	93.78	5.315	113.5	100.4	4.777
Sargan P-value	0.000	0.000	0.000	0.018	0.000	0.007

Note: Estimated using Canadian Labour Force Survey (1987-2019). Sample includes non-education, self-employed professionals, aged 25-59. Models (1)-(3) include a control group of business and natural science professionals, while models (4)-(6) exclude pre-trends in the first stage (see Table 3.7). Models are estimated without sampling weights and include covariates listed in Table 3.2.

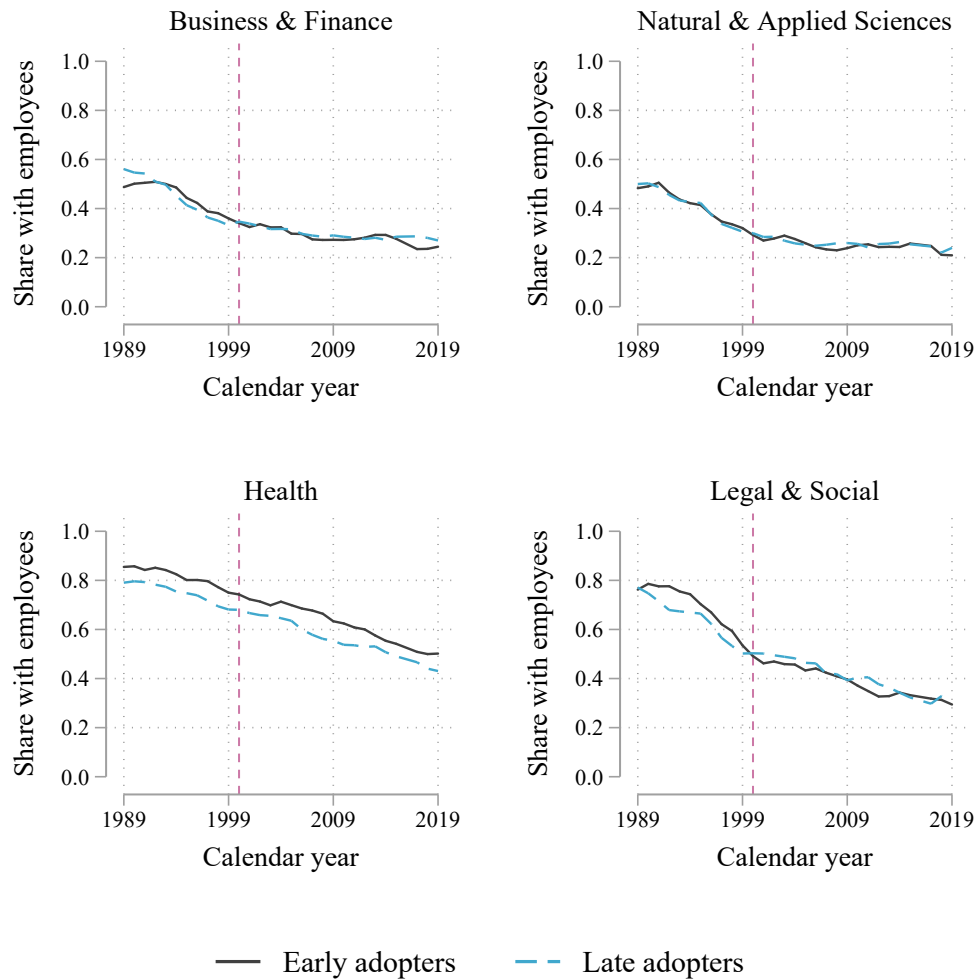


Figure 3.11: Trend: Hiring status by professional group and provincial treatment group.

Note: Constructed using Canadian Labour Force Survey (1989-2019). Sample includes all non-education self-employed professionals, aged 25-59. This trend displays the 3-year moving average and is constructed without sampling weights. Late adopter provinces include Manitoba, Saskatchewan, Ontario, and Quebec. The vertical line denotes the year 2000, the year before the 2001 reform period.

Chapter 4

Labour market institutions and the distribution of wages: The role of spillover effects

4.1 Introduction

A vast literature has investigated the causes of substantial and continuing growth in wage and earnings inequality in the United States. Although most studies suggest that various forms of technological change are a leading explanation for these changes (see, e.g., [Acemoglu and Autor, 2011](#)), other explanations such as changes in labour market institutions have been implicated. For instance, [DiNardo et al. \(1996, DFL from hereinafter\)](#) show that the decline in the real value of the minimum wage during the 1980s helps accounts for a significant fraction of the growth in wage inequality at the bottom of the distribution during this period. [Card \(1996\)](#), [Freeman \(1993\)](#) and DFL show that the decline in unionization contributed to the rise in male wage inequality over the same period. [Card et al. \(2020, 2004\)](#) and [Firpo et al. \(2018\)](#) find that the continuing decline in unionization after the late 1980s accounts for some of the continuing growth in inequality, while [Farber et al. \(2018\)](#) reach a similar conclusion using data going back to the 1940s.

A critical limitation of the earlier literature is that it typically ignored potential spillover effects of institutional changes. These could magnify the impact of such changes on the wage distribution. In an influential study, [Lee \(1999\)](#) shows that accounting for spillover or "ripple" effects of the minimum wage on the wage of workers earning slightly above the minimum substantially increases the impact of the minimum wage on the wage distribution. [Lee \(1999\)](#) finds that the decline in the minimum wage can explain half of the increase in the standard deviation of log wages, and almost all of the increase in the 50-10 differential between 1979 and 1989 once spillover effects are taken into account. Lee's estimates of the contribution of the minimum wage to inequality growth are substantially larger than those of DFL, who ignore spillover effects, although they have been recently challenged by [Autor et al. \(2016\)](#). DFL find that declining minimum wages explain about a quarter (25% for men and 30% for women) of the increase in the standard deviation of log wages between 1979

and 1988, and about 60% of the increase in the 50-10 differential.

With a few exceptions, existing studies of the impact of de-unionization on wage inequality ignore possible spillover effects of unionization. The existing decompositions typically assume that the observed non-union wage structure provides a valid counterfactual for how union workers would be paid in the absence of unionization. However, it has long been recognized that union power, as measured by the unionization rate (or related indicators), may also influence wage setting in the non-union sector (e.g., [Lewis, 1963](#)). In particular, non-union employers may seek to emulate the union wage structure to discourage workers from supporting unionization. This “threat effect” ([Rosen, 1969](#)) likely increases the equalizing effects of unionization by making non-union wages more similar to the more equally distributed ones observed in the union sector. Based on cross-country evidence, [Freeman \(1996\)](#) conjectures that failing to incorporate threat effects biases down existing estimates of the effect of de-unionization. [Taschereau-Dumouchel \(2020\)](#) reaches a similar conclusion by calibrating a search model of the U.S. economy.

Empirical evidence on the distributional impact of threat effects is limited by the challenge of finding exogenous sources of variation in the rate of unionization (the conventional measure of threat effects) across labour markets. Older studies such as [Freeman and Medoff \(1981\)](#), [Moore et al. \(1985\)](#), and [Podgursky \(1986\)](#) estimate threat effects by including the unionization rate in the relevant market (defined by industry, occupation, and geography) in a standard wage regression, but only make limited attempts at controlling for possible confounding factors. One exception is [Farber \(2005\)](#) who uses the passage of “right-to-work” (RTW) laws in the states of Idaho (1985) and Oklahoma (2001) as an arguably exogenous source of variation in union power. Unfortunately, Farber’s results based on Current Population Survey (CPS) data are inconclusive because of a lack of statistical power linked to the small samples available in these two states. More recently, [Farber et al. \(2020\)](#) exploit cross-state variation linked to the National War Labor Board and the introduction of the Wagner Act to show that the unionization rate reduced wage dispersion in the middle of the 20th century. We provide additional evidence of the threat effects of unions using the more recent adoption of RTW laws in Indiana, Michigan, Wisconsin, West Virginia, and Kentucky.¹³³ As in [Farber \(2005\)](#), we lacked statistical power to draw clear implications for wage inequality; in part, because union activity increases with the passing of these laws. Thus, while there is a significant effect on union coverage, the results for non-union wages are not robust.

The contribution of this paper is twofold. First, we update DFL’s analysis until 2017 to see whether changes in labour market institutions have remained an important source of

¹³³Some of these results appear in an earlier version of the paper ([Fortin et al., 2019](#)). However, the results are absent in the forthcoming published paper.

inequality change over the last 25 years. Second, we extend DFL by taking into account spillover effects of the minimum wage and unionization. In the case of the minimum wage, we depart from [Lee \(1999\)](#) and [Autor et al. \(2016\)](#) by estimating a rich model of the wage distribution using distribution regressions ([Chernozhukov et al., 2013](#); [Foresi and Peracchi, 1995](#); [Fortin and Lemieux, 1998](#)). The model is analogous to a distributional difference-in-differences approach that yields estimates of spillover effects regardless of whether the minimum wage varies at the state or federal level.

We extend this framework to consider the case of union threat effects, which are estimated jointly with the effect of the minimum wage, allowing us to construct counterfactual wage distributions with and without minimum wage and union spillovers. Given the challenges of finding a suitable instrument for unionization, we proxy the threat effect using the unionization rate at the state-industry level, and include it as an additional regressor in the distribution regressions. A rich set of controls, including industry- and state-specific trends, are also included to control for other factors that could confound the relationship between wages and the rate of unionization. As in DFL, we also estimate the direct impact of unions using a “shift-share” analysis.

Our key findings are as follows. First, we estimate minimum wage spillover effects that are roughly as large as those found by [Lee \(1999\)](#) for the 1980s, though the magnitude of spillover effects is smaller in subsequent years. These differences partly reconcile the difference in results between [Lee \(1999\)](#) and [Autor et al. \(2016\)](#), who found smaller spillover effects using data for more recent years. Second, we find that minimum wage changes account for most of the substantial growth in lower tail inequality (50-10) in the 1980s, and its relative stability since then. Our main finding concerning unions is that spillover effects of unionization on non-union wages are similar in magnitude and distributional impact (shape) to the direct, or “shift-share”, impact of unionization. The effects are largest in the lower middle of the distribution but negative at the top. Adding spillover effects roughly doubles the contribution of de-unionization to the growth in wage inequality. For instance, in the case of men, the contribution of unions to the steady growth in the 90-50 gap over the entire 1979-2017 period goes from 20% to 40% when spillover effects are taken into account. Overall, we find that changes in labour market institutions account for 53% and 28% of the 1979-2017 growth in the standard deviation of log wages for men and women, respectively.

The remainder of the paper is organized as follows. In Section 4.2, we propose a distribution regression approach to estimate the minimum wage spillover effects. In Section 4.3, we discuss our strategy for measuring threat effects in a distributional context. We present the data and estimation results in Section 4.4 and use decompositions to compute the contribution of changing institutions to changes in the wage distribution in Section 4.5. We conclude in Section 4.6.

4.2 Estimating spillover effects of the minimum wage

A key contribution of DFL was to present visual evidence based on kernel density estimates to illustrate the role of the decline in the real value of the minimum wage in the growth of wage inequality between 1979 and 1988. DFL then made two main assumptions to quantify the contribution of the minimum wage to inequality growth. First, they assumed that the changes in the minimum wage did not affect employment. [Card and Krueger \(1995\)](#)'s contemporary work was used in support of the assumption of no employment effect. DFL also showed that allowing for modest employment effects had little impact on the findings. On the other hand, recent work by [Brochu et al. \(2018\)](#) based on Canadian data shows substantial spillover effects even after controlling for employment effects using a hazard rate estimation approach. [Cengiz et al. \(2019\)](#) find evidence that spillover effects easily counterbalance dis-employment effects using a "bunching" estimator implemented in a distributional event study approach. In light of this recent evidence, we ignore the minimum wage's possible employment effects in this study.

More importantly, DFL assumed that minimum wages had no spillover effects. This assumption allowed them to use a simple "tail pasting" approach where the bottom end of the distribution in a low minimum wage year (1988) is replaced by the corresponding bottom end of the distribution in a high minimum wage year (1979).

[Lee \(1999\)](#) relaxed the assumption of no spillover effects by exploiting the fact that a prevailing federal minimum wage is relatively higher in low-wage than high-wage states. His basic estimation approach consists of running flexible regressions of selected wages percentiles relative to the median on the relative value of the minimum wage by state and year. This involves running regressions of $w_{st}^q - w_{st}^5$ on a polynomial function in $mw_{st} - w_{st}^5$, where w_{st}^q is the q th percentile of log wages in state s at time t , while mw_{st} is the corresponding value of the minimum wage. The term $mw_{st} - w_{st}^5$ measures the relative "bite" of the minimum wage in different states. The minimum wage "bites" more in low-wage states where $mw_{st} - w_{st}^5$ is larger than in high-wage states where it is lower. Using this approach, Lee finds that the minimum wage impacted wage percentiles above and beyond the corresponding value of the minimum wage. He concludes that changes in the minimum wage can explain most of the change in inequality in the lower tail of the distribution between 1979 and 1989, once spillover effects are taken into account.

This finding has been challenged by [Autor et al. \(2016\)](#) who point out that sampling error in the estimated median wage w_{st}^5 can positively bias estimates of Lee-type regressions as the noisily measured median is included on both sides of the regression. They suggest correcting for this problem by instrumenting the right-hand side variable $mw_{st} - w_{st}^5$ with the value of the minimum wage mw_{st} . As Lee-type regressions also include year dummies,

this strategy can only work in periods where there is substantial variation in state minimum wages, given that time dummies fully absorb the variation in the federal minimum wage. [Autor et al. \(2016\)](#) take advantage of the substantial variation in state minimum wages after the 1980s (see Figure 4.1) to revisit Lee's estimates and find substantially smaller spillover effects.

One alternative interpretation of these findings is that Lee's estimates of spillover effects are not substantially biased, but they have become smaller over time. It is indeed unclear that the more frequent and smaller changes in state minimum wages of the post-1980s period have a comparable impact to the massive (over 30%, see Figure 4.1) and permanent decline in the real value of the federal minimum wage that took place during the 1980s. Indeed, a large and permanent change in the minimum wage may affect the composition of firms at the lower end of the wage distribution. [Butcher et al. \(2012\)](#) argue that when firms have monopsony power, spillover effects can arise as unproductive firms shut down when the minimum wage increases and their workers move to more productive and higher-paying firms.¹³⁴ Such a reallocation channel is unlikely to occur for smaller and more transitory changes in the minimum wages. Spillover effects may still arise because of internal wage considerations ([Dube et al., 2019](#); [Grossman, 1983](#)), but the magnitude of these spillover effects may be smaller than when longer-term labour reallocation effects are involved.¹³⁵

In what follows, we propose a new estimation approach based on distribution regressions that makes it possible to estimate minimum wage spillover effects regardless of whether the minimum wage varies at the state or federal level. Intuitively, [Lee \(1999\)](#) uses a two-step procedure by estimating distributional statistics like the median in a first step and plugging it in a regression model for wage percentiles in a second step. [Autor et al. \(2016\)](#) then propose an IV procedure to correct the bias from the noisy measure plugged into the second step estimation. By contrast, in our approach, we jointly estimate the wage distribution and the impact of the minimum wage in a single step. As a result, our approach does not yield biased results because of the estimated regressor problem.

4.2.1 Distribution regressions

Following [Foresi and Peracchi \(1995\)](#) and [Chernozhukov et al. \(2013\)](#), we use a distribution regression approach to model the whole wage distribution and the effects of the minimum wage at different points of the distribution. The logic is straightforward. The probability of

¹³⁴See also [Haanwinckel \(2020\)](#), who highlights a similar channel in a model where, as in [Teulings \(2000\)](#), firms differ in their task requirements but also have some monopsony power.

¹³⁵See [Brochu et al. \(2018\)](#) for a more thorough discussion of possible economic explanations for minimum wage spillover effects, and [Dustmann et al. \(2020\)](#) for evidence of reallocation effects following the introduction of a nation-wide minimum wage in Germany.

an outcome variable Y being above (or below) a given cut-point y_k is modeled as a flexible function of covariates X , and estimated using a probit, logit, or linear probability model. For example, in the case of a probit model we have:

$$\text{Prob}(Y \geq y_k) = \Phi(X\beta_k) \quad \text{for } k = 1, 2, \dots, K \quad (4.1)$$

The y_k cutoffs can either be chosen using a fine grid or as percentiles ($k=1, 2, \dots, 99$) of the unconditional wage distribution. The method is quite flexible as rich functions of the covariates, including state and year dummies, can be included as regressors, and no restrictions are imposed on how β_k varies across cutoff values. Once the series of distribution regressions have been estimated, various counterfactual scenarios can be computed by either changing the distribution of the covariates or some of the β_k coefficients.

However, the flexibility of these distribution regressions comes at a cost as there is no guarantee to get positive counterfactual probabilities, especially when the set of covariates is large. Furthermore, and as we discuss below, minimum wage effects are modeled by adding to the list of covariates a set of dummy variables indicating where the minimum wage stands (at, below, or above) relative to a given cutoff point y_k . Allowing for different minimum wage effects at each cutoff would be an overly flexible approach yielding identification challenges (see Section 4.2.3). For the same reason, [Firpo et al. \(2009\)](#)'s RIF-regressions are not ideally suited to the estimation of minimum wage spillover effects. We instead propose a more parametric, though still flexible, approach where the impact of the minimum wage is modelled in a relatively parsimonious way.¹³⁶ As we show in Section 4.4.2 below, doing so provides a better connection to the bunching analysis of [Cengiz et al. \(2019\)](#), and helps highlight the channels through which the minimum wage reshapes the wage distribution.

We impose restrictions on the β_k coefficients by letting them evolve in a smooth way over the wage distribution. Doing so also helps provide an economic interpretation to the distribution regressions. For example, when the β_k 's are fixed across the distribution, the model can be represented using a standard latent variable framework. Consider a latent log wage or skill index $Y^* = X\beta + \varepsilon$, where $\varepsilon \sim N(0, 1)$. The observed wage is assumed to be a monotonic transformation $Y = g(X\beta + \varepsilon)$ of the skill index. [Fortin and Lemieux \(1998\)](#) call this model a “rank regression” as the main restriction imposed is that each observation's rank is the same in both the wage and skill distributions.

The model is flexibly estimated by dividing the wage range into a fine grid. [Fortin and Lemieux \(1998\)](#) use about 200 cutoff points y_k . The corresponding cutoff points in the skill

¹³⁶[Brochu et al. \(2018\)](#) use a closely related method based on hazard functions instead of distribution regressions. For reasons explained below, a useful feature of our approach is that it is more directly connected to a latent skill index model of wage setting latent variables. That said, the two methods yield similar estimates of minimum wage spillover effects, which is re-assuring given the differences between the two methods.

distribution, c_k , are defined as $c_k = g^{-1}(y_k)$. It follows that:

$$Prob(Y \geq y_k) = \Phi(X\beta - c_k)$$

This corresponds to a standard ordered probit model where the probability of observing wages in a wage category $[y_k, y_{k+1}]$ is given by:

$$Prob(y_k \leq Y \leq y_{k+1}) = \Phi(X\beta - c_{k+1}) - \Phi(X\beta - c_k)$$

When the transformation function $g(\cdot)$ is linear, it follows that $Y = \sigma \cdot (X\beta + \varepsilon) = X\beta' + u$, where $\beta' = \sigma\beta$ and $u = \sigma\varepsilon$ is a homoskedastic normal error term with a standard deviation of σ . It also follows that the cutoff points in the ordered probit model, c_k , are a linear function $c_k = y_k/\sigma$ of the wage cutoffs y_k .

While log normality may not be a bad approximation of the conditional wage distribution, the homoskedasticity assumption is quite strong and clearly violated in wage data (see, e.g. [Lemieux, 2006](#)). For the rank regression model to fit the data reasonably well, it is essential to allow for heteroscedasticity in the error term ε . To see how this changes the probability model, consider a simple case where individuals belong to two education groups: high school ($X = 0$) and college ($X = 1$) graduates. Assume that log wages are normally distributed with a different mean and variance for each of the two groups:

$$\begin{aligned} Y &= \beta_0 + \varepsilon && \text{with } \varepsilon \sim N(0, \sigma_0) \text{ for } X = 0, \text{ and} \\ Y &= \beta_1 + \varepsilon && \text{with } \varepsilon \sim N(0, \sigma_1) \text{ for } X = 1 \end{aligned}$$

It follows that

$$\begin{aligned} Prob(Y \geq y_k | X) &= \begin{cases} \Phi\left(\frac{\beta_0}{\sigma_0} - \frac{y_k}{\sigma_0}\right) & \text{if } X = 0 \\ \Phi\left(\frac{\beta_1}{\sigma_1} - \frac{y_k}{\sigma_1}\right) & \text{if } X = 1 \end{cases} \\ &= \Phi\left[\beta'_0 + X\beta + c_k + X\left(\frac{1}{\sigma_1} - \frac{1}{\sigma_0}\right)y_k\right] \end{aligned} \quad (4.2)$$

where $c_k = y_k/\sigma_0$, $\beta'_j = \beta_j/\sigma_j$, $\beta = \beta'_1 - \beta'_0$ is the main effect of education, and $\left(\frac{1}{\sigma_1} - \frac{1}{\sigma_0}\right)$ is the coefficient on the interaction between X and y_k . In other words, introducing heteroskedasticity leads to a specification where the effect of education varies in a smooth (linear) way over the wage distribution.¹³⁷

¹³⁷An alternative interpretation is that the cutoff points in the ordered probit model are now $c_k + X\left(\frac{1}{\sigma_1} - \frac{1}{\sigma_0}\right)y_k$, and depend on the value of the covariate X .

The heteroskedastic model provides a middle ground between distribution regressions where β_k vary in a completely unrestricted way and a rank regression model where β_k is constrained to be the same (except for the intercept) at each cutoff point y_k . While we use linear interactions in the empirical applications presented here, a more flexible set of interactions between X and polynomial functions in y_k can be used.

4.2.2 Empirical implementation

We empirically implement the distribution regression model by dividing the log wage distribution into 58 intervals of width 0.05 log points.¹³⁸ As we are constraining the coefficients to change smoothly across wage cutoffs, the model is estimated by jointly fitting 57 “stacked” probit regressions.¹³⁹ The covariates (Z_{ist}) include a set of state (θ_s), year (γ_t), and quarter (η_q) fixed effects, as well as state-specific trends ($t \cdot \pi_s$) and a rich set of individual characteristics (X_{ist}) similar to those used by DFL. These covariates include years of education, a quartic in potential experience, experience-education interactions (16 categories plus experience times education), 11 industry categories 4 occupation categories, and dummy variables for race, marital status, public sector, part-time, and SMSA.¹⁴⁰ These variables are included in the probit models as:

$$Z_{ist}\beta = X_{ist}\beta_x + \theta_s + \gamma_t + \eta_q + t \cdot \pi_s$$

In light of the above discussion, we also include interactions $y_k Z_{ist} \lambda$ between the covariates and the cutoff points y_k .¹⁴¹ We model how the minimum wage mw distorts the wage distribution by creating a large spike at mw and changing the wage distribution above and below the minimum, using event-study type parameters Φ_m , $m \in \{-M_1, \dots, M_2\}$ in the wage space. The parameter Φ_0 captures the magnitude of the spike “right at” mw . The parameters Φ_m for $m > 0$ are used to model spillover effects, while Φ_m for $m < 0$ capture the decline in the wage density below the minimum wage. Although these minimum wage parameters’ role is intuitively simple, the implementation in a distributional regression

¹³⁸For over 99% of observations from 1979 to 2017, the log wage falls in the range going from 1.6 (4.95) to 4.4 (81.50). All wages are converted into 2017 dollars. There are 56 intervals of width 0.05 going from 1.6 to 4.4, plus two intervals for log wages below 1.6 or above 4.4. While a finer grid could be used in the estimation, doing so would increase the computational burden with limited gains in fitting detailed distributional features. The issue is that most of the wage variation within narrow wage bins is dominated by idiosyncratic variation linked to how wages are reported in the CPS data (e.g., rounding off at 10 cents, 25 cents, or dollar values of nominal wages).

¹³⁹As with an ordered probit model, this estimator is equivalent to a stacked probit model, where the right-hand-side variables are repeated K times and paired with an indicator variable, $\mathbf{1}[y_{its} > y_k]$, $k = 1, \dots, K$.

¹⁴⁰In models including state-industry union coverage rate, we include industry-specific trends and an interaction between industry and the linear wage cut-off term.

¹⁴¹Note that not all covariates are interacted with the linear wage cut-off term. For computational reasons, we restrict the set of interactions to the state and year effects, as well as the experience-interaction dummies.

model is complicated because we consider minimum wage effects on the cumulative distribution. For illustration, consider a simple case where spillover effects are limited to one wage bin above the minimum wage, $m \in \{0, 1\}$, and where we ignore covariates. The probability that $Y \geq y_k$ in absence of the minimum wage is $Prob(Y \geq y_k) = \Phi(\beta_0 - c_k)$. Introducing a minimum wage does not change these probabilities in the upper part of the wage distribution above the range of spillover effects. As the minimum wage gets closer to the wage cutoff y_k , $Prob(Y \geq y_k)$ increases due to spillover effects even if the minimum wage is still below y_k . $Prob(Y \geq y_k)$ increases even more when the minimum wage crosses above y_k since, in that case, $Prob(Y \geq y_k)$ includes the probability of being at the mass point exactly at mw . Using the parameters Φ_0 and Φ_1 , we can write the probabilities in these three cases - mw much below, a bit below, and just above y_k - as follows:

$$Prob(Y \geq y_k) = \begin{cases} \Phi(\beta_0 - c_k) & \text{if } mw < y_{k-1} \\ \Phi(\beta_0 - c_k + \varphi_1) & \text{if } y_{k-1} \leq mw < y_k \\ \Phi(\beta_0 - c_k + \varphi_0 + \varphi_1) & \text{if } y_k \leq mw < y_{k+1} \end{cases} \quad (4.3)$$

Using equation (4.3) to compute the probability of being in the wage bin just above the minimum wage, we have:

$$Prob(y_k \leq Y < y_{k+1}) = \Phi(\beta_0 - c_k + \varphi_1) - \Phi(\beta_0 - c_{k+1}) \quad \text{if } y_{k-1} \leq mw < y_k$$

Increasing the value of Φ_1 increases $\Phi(\beta_0 - c_k + \varphi_1)$ and the probability mass linked to spillover effects in the wage bin just above the minimum wage. Likewise, the probability of being in the wage bin “right at” the minimum wage depends on the parameter Φ_0 since:

$$Prob(y_k \leq Y < y_{k+1}) = \Phi(\beta_0 - c_k + \varphi_0 + \varphi_1) - \Phi(\beta_0 - c_{k+1}) \quad \text{if } y_k \leq mw < y_{k+1}$$

More generally, these minimum wage effects can be captured using a set of dummy variables, D_{kst}^m , and writing: $Prob(Y_{ist} \geq y_k) = \Phi(\beta_0 + \sum_m D_{kst}^m \varphi_m - c_k)$, where $D_{kst}^m = \mathbf{1}[y_{k-m} \leq mw_{st}]$, $m \in \{-M_1, \dots, M_2\}$.¹⁴² In practice, we allow for up to six spillover effect parameters, and use three parameters to model how the minimum wage reduces the probability distribution below minimum: $m \in \{-3, 6\}$.

This method is reminiscent of the bunching/event study design of Cengiz et al. (2019), who estimate changes in the fraction of observations in different dollar wage bins following increases in the minimum wage. Our parameters φ_m implicitly capture the same type of

¹⁴²This general model allows for spillover effects higher up in the distribution (e.g. $D_{kst}^2 \varphi_2$ for two wage bins above the minimum wage), and for negative effects to the part of the distribution below the minimum wage (e.g. $D_{kst}^{-1} \varphi_{-1}$).

changes along the wage distribution, but are imbedded in the overall estimation of the wage distribution in the presence of covariates.

We also include another set of dummy variables to account for the substantial heaping at integer values of hourly wages in the CPS data, especially at \$5 and \$10 (see Appendix Figure C.2). Heaping can have an important impact on estimated probabilities depending on whether a given cutoff point y_k is just below or above an integer value. It can also affect the estimated effect of the minimum wage if workers earning the minimum wage round off their wage report to the nearest integer. This measurement error could create spurious spillover effects when the minimum wage is slightly below an integer value.¹⁴³ This is a critical issue in the literature as Autor et al. (2016) present calculations suggesting that minimum wage spillovers effects may be a spurious consequence of measurement error.

Since we are working with log real wages Y_{ist} , nominal wages in levels, W_{ist} , can be written as $W_{ist} = P_t \cdot \exp(Y_{ist})$, where P_t is the price level in year t relative to the base (year 2017 in the empirical analysis). Likewise, the wage cutoffs can be written in nominal terms as $w_{kt} = P_t \cdot \exp(y_k)$. We want the interval probability, $Prob(y_k \leq Y_{ist} < y_{k+1}) = Prob(w_{kt} \leq W_{ist} < w_{k+1,t})$, to be larger when, for example, \$10 is included in the $[w_{kt}, w_{k+1,t}]$ interval. In this case, the increased probability can be modeled using a dummy variable $L_{kst}^{10} = \mathbf{1}[w_{k,t} \leq 10]$ and the corresponding parameter γ_{10} , so that $Prob(Y_{ist} \geq y_k) = \Phi(\beta_0 + L_{kst}^1 \gamma_{10} - c_k)$.¹⁴⁴ It follows that $Prob(w_{kt} \leq W_{ist} < w_{k+1,t}) = \Phi(\beta_0 - c_k + \gamma_{10}) - \Phi(\beta_0 - c_{k+1})$, where the parameter γ_{10} determines how much extra mass there is in the wage bin containing \$10. As wage heaping is most pronounced for values of wages up to \$10, we create dummies for heaping at \$5, \$10, and any other integer value up to \$10. These dummies, L_{kst}^p , are included as additional covariates in all estimated models. The resulting probit models being estimated for $k = 1, \dots, 57$ are:

$$Prob(Y_{ist} \geq y_k) = \Phi \left(Z_{ist}\beta + y_k Z_{ist}\lambda + \sum_{m=-3}^6 D_{kst}^m \varphi_m + \sum_{p=1}^{10} L_{kst}^p \gamma_p - c_k \right) \quad (4.4)$$

Note that the model nests the case of no spillover effects ($\varphi_m = 0$ for $m > 1$) considered by DFL. Standard errors are clustered at the state level to allow for correlation across the 57 probit models and for autocorrelation over time.

¹⁴³For instance, if workers earning a \$9.80 minimum wage report a \$10 wage in the CPS, this will increase the mass just above the minimum wage and give a false impression about the importance of spillover effects.

¹⁴⁴To use a concrete example, consider the case where two successive cutoff points w_{kt} and $w_{k+1,t}$ are just below and above \$10 (for example, \$9.75 and \$10.25). The dummy variable L_{kst}^{10} is equal to 1 at w_{kt} ($L_{kst}^{10} = \mathbf{1}[w_{k,t} \leq 10] = \mathbf{1}[9.75 \leq 10] = 1$) but turns to 0 at $w_{k+1,t}$ ($L_{k+1,st}^{10} = \mathbf{1}[w_{k+1,t} \leq 10] = \mathbf{1}[10.25 \leq 10] = 0$).

4.2.3 Identification

As mentioned earlier, the distribution regression model is identified regardless of whether the prevailing minimum wage is set at the federal or state level. This may be surprising at first glance since the model in equation (4.4) includes a full set of state and time dummies, where the latter would absorb all the variation in the federal minimum wage in a traditional difference-in-differences setting. As it turns out, only allowing for a smooth change in the probit coefficient across wage cutoffs plays an essential role in the identification when the minimum wage only varies at the federal level.

To see this, note that allowing for an unrestricted set of time effects γ_t^k for each cutoff point y_k would make it impossible to identify the federal minimum wage's distributional effect. Such an approach would be overly flexible in light of the above discussion on the economic interpretation of the coefficients in the distribution regression. Going back to the example in equation (4.2), if X was a time instead of an education dummy, the main effect β would capture a shift in mean wages over time, while the coefficient $\left(\frac{1}{\sigma_1} - \frac{1}{\sigma_0}\right)$ on the interaction between X and y_k would capture changes in the variance over time. One could also go further by including interaction terms between X and polynomial functions of y_k that would capture changes in moments of the wage distribution besides the mean and the variance. The implication remains that time effects should only vary smoothly across the various cut points y_k of the distribution.

Identification of minimum wage effects is now possible as the minimum wage “bites” at different points of the distribution at different times, a feature of the wage distribution that cannot be captured by smoothly varying time effects. Intuitively, the minimum wage creates a sharp discontinuity in the probability of being just above and just below its value. As in [Doyle \(2006\)](#) and [Jales \(2018\)](#), identification can be achieved as in a regression discontinuity design provided that the underlying latent wage distribution is smooth around the value of the minimum wage. Constraining the distribution regression's coefficients to change smoothly across the various cut points y_k implies that the latent distribution is also smooth.¹⁴⁵ We further discuss these identification issues using a series of graphs in Appendix C.3.

¹⁴⁵As in [Fortin and Lemieux \(1998\)](#), the underlying latent wage distribution is quite flexible despite the normality assumption used to estimate the probit models. The source of additional flexibility is the $g(\cdot)$ function in [Fortin and Lemieux \(1998\)](#); here, it is implemented empirically by estimating a separate coefficient c_k at each cutoff.

4.3 Union threat effects

We use two approaches to assess the importance of union threat effects. First, we examine the reduced evidence for a union threat effect using the recent adoption of “right-to-work” laws (Ellwood and Fine, 1987; Farber, 2005).¹⁴⁶ We focus on the case of three relatively large Midwestern states, Indiana, Michigan, and Wisconsin, that did so in 2011, 2013, and 2015, respectively.¹⁴⁷¹⁴⁸ The second and main approach uses the unionization rate at the state-industry-year level as a measure of the (declining) threat of unionization. An important advantage of this approach is that it can easily be integrated in the distribution regression approach proposed in Section 4.2 by adding the unionization rate at the state-industry-year level to the list of covariates included in the model.

4.3.1 Event study of the introduction of right-to-work laws

Under the 1935 National Labor Relations Act, all U.S. workers covered by collective bargaining receive the same benefits from unionization including compensation, benefits, and access to grievance procedures regardless of whether they are members of the union. In most states, workers covered by a collective agreement have to pay union dues (typically withheld from paychecks by employers) regardless of whether they decide to become members of their union.

However, following the passage of the Taft-Hartley Act in 1947, it became possible for States to introduce so-called “right-to-work” (RTW) laws making it no longer compulsory for workers covered under a collective bargaining agreement to pay union dues. As shown in Figure 4.2, several (mostly Southern) states quickly adopted RTW around that time. A few states then adopted RTW laws in the 1950s, 1960s, and 1970s. The impact of these RTW adoptions cannot be studied using micro data on union status and wages that only became available (with a full set of state indicators) in the late 1970s.

The next two RTW adopters, Idaho (1985) and Oklahoma (2001), were studied by Farber (2005) who could not draw informative conclusions because of the statistical imprecision

¹⁴⁶Right-to-work laws typically prohibit union security agreements, or agreements between labour unions and employers, that govern the extent to which an established union can require employees’ membership, payment of union dues, or fees as a condition of employment, either before or after hiring.

¹⁴⁷For public workers in the state of Wisconsin we use 2011 as the date of the introduction of right-to-work laws. Under Governor Scott Walker, the State introduced a law (Bill 10) in June 2011 that suspended collective bargaining and made union dues contribution voluntary in the public sector. However, it took several years for the law to have a full impact as provisions only started binding upon expiration of existing collective bargaining agreements.

¹⁴⁸We estimate models for Indiana, Michigan, and Wisconsin only, using other Rust Belt states as controls, as well as more general specifications for all states where right-to-work policy changes were introduced. In the latter case, Oklahoma (2001), West Virginia (2016), and Kentucky (2017) are also used in the estimation.

linked to the small CPS sample sizes in these two small states.¹⁴⁹ In this paper, we take advantage of the introduction of RTW laws in the three relatively larger Midwestern states of Indiana, Michigan, and Wisconsin. Two other states, Kentucky (2017) and West Virginia (2016), have also adopted RTW laws very recently. As we will see below, these two states don't play much of a role in our analysis due to the very short time span available after the adoption of RTW laws. Furthermore, it is not yet possible to study the impact of a recent Supreme Court decision (Janus case, June 2018) that has imposed RTW to the entire U.S. public sector in January 2019.

RTW laws weaken unions by allowing free riding by workers covered under a union contract. For instance, recent work by [Feigenbaum et al. \(2018\)](#) shows that the passage of RTW laws had an adverse impact on union finances and campaign contributions.¹⁵⁰ As in [Farber \(2005\)](#), we expect that by reducing union power, RTW laws should have a negative impact on unionization rate and non-union wages due to declining threat effects. As differences across RTW and non-RTW states could reflect cross-states differences in confounding factors, we adopt a difference-in-difference (DiD) and event-study approach to isolate the impact of RTW laws on state unionization rates and the wages of union and non-union workers. Table 4.2 presents the difference-in-difference results, while Figure 4.3 plots the coefficients from the corresponding event-study design examining the adoption of RTW in Oklahoma, Indiana, Michigan, Wisconsin, West Virginia, and Kentucky.¹⁵¹ As discussed in Section 2.4 of Chapter 2, there various ways of implementing an event-study-design. In this application, we include the sample of 'never-treated' (non-RTW) states as an additional control.¹⁵² For this reason, both the DiD and event-study methodologies include state (treatment unit) FEs.

The difference-in-difference and event-study results both suggest a decline of union coverage in the order of 2 %-points.¹⁵³ The difference-in-difference results for non-union wages are significant and suggest a decline in non-union wages larger than 2%. However, the corresponding event-study results are noisy and suggest that if there is an effect it is

¹⁴⁹Some studies include a change to the RTW laws in Texas in 1993 as an additional source of policy variation ([Taschereau-Dumouchel, 2020](#)). Since Texas's original RTW legislation was introduced in 1947, we group the state with earlier adopters.

¹⁵⁰[Ellwood and Fine \(1987\)](#) show that RTW laws have an adverse impact on union organizing activities.

¹⁵¹A difficulty with an event-study of RTW is that often it takes several years for the law to have a full impact as provisions only start binding upon expiration of existing collective bargaining agreements. [Biasi and Sarsons \(2020\)](#) who study the specific case of Wisconsin find that this is indeed the case using data on the expiration dates of public school teachers' collective bargaining agreements.

¹⁵²In practice, this is done by setting event-time to -1 for 'never-treated' states. This enables us to include state FEs in the estimation. We find that using a control group improves the estimation of longer lags. This is, in part, because the state FEs better control for the changing composition of states observed 3-5 years after treatment.

¹⁵³We find that in Wisconsin the decline is considerably larger as a result of public sector cuts that resulted in a loss of unionized public sector employment. This is evident in the larger - although insignificant - coefficient in column (4) of Table 4.2.

smaller than 2%.¹⁵⁴ One reason that the wage results may not be robust is that union activity need not decline with the passing of these laws. Indeed, we find some evidence to support the idea that union activity, as measured by union election activity, increased with the passing of the RTW laws in the three Midwestern states. Figure C.1 plots aggregate union election outcomes by event-time for the three Midwestern switching states. There is a clear dip in the number of elections held in the year the law passes, but the level remains stable thereafter. There is also evidence of an increase in union election success after the reform, corresponding with an increase in certified workers.¹⁵⁵ Together, these results suggest that the passing of RTW resulted in an increase in union activity, which could explain the lack a strong negative result for non-union wages.

In principle, RTW laws could also be used as an instrumental variable in a regression of non-union (and union) wage on state unionization rate. Resulting estimates could then be used to compute the contribution of declining unionization rates to change in the distribution of wages of non-union workers.¹⁵⁶ This approach could potentially provide a way of quantifying the role of declining threat effects on the wage distribution.

As discussed above, the statistical imprecision makes it challenging to use the event-study estimates to compute the contribution of threat effects to changes in wage inequality. The main purpose of the analysis of RTW laws is, thus, to provide some evidence supporting the view that threat effects are a significant factor in wage setting, as opposed to a spurious consequence of the fact unionization rates at the state-year level may be correlated with omitted factors.

4.3.2 Measuring threat effects in a distributional context using the unionization rates

Older studies based on cross-sectional data or short repeated cross-sections have generally found that the unionization rate was positively correlated with non-union workers' wages.¹⁵⁷ We generalize this approach by looking at how U_{jst} , the rate of unionization in industry j and state s at time t , affects the whole distribution of wages. This is achieved by estimating separate distribution regressions for union and non-union workers, and

¹⁵⁴We are actively exploring new synthetic control methodologies as these new methods can provide more precise estimates in both difference-in-difference and event-study settings.

¹⁵⁵In results not shown here, this rise increase in certification contrasts a continued decline in other Midwestern states over this period. Unfortunately, we lack the statistical power to provide robust regression analysis results using this data.

¹⁵⁶One would actually need to go beyond simple regressions to look at distributional impacts. This could be done, for instance, by adapting the distribution regression approach to the case where there is an endogenous regressor.

¹⁵⁷See, for instance, [Freeman and Medoff \(1981\)](#) and [Podgursky \(1986\)](#). A similar approach has been adopted in recent studies like [Rosenfeld et al. \(2016\)](#) and [Denice and Rosenfeld \(2018\)](#) that use data for a much longer time horizon.

allowing threat effects to vary at different points of the distribution by interacting U_{jst} with a quartic function in the wage cutoff points y_k . The resulting probit model being estimated separately for union and non-union workers are:

$$Prob(Y_{ist} \geq y_k) = \Phi \left(A_{ist}^k + \sum_{q=0}^4 U_{jst}^q \kappa_q + t \cdot \pi_j \right) \quad (4.5)$$

where κ_q are the parameters associated to the quartic function in U_{jst} , and $A_{ist}^k = Z_{ist}\beta + y_k Z_{ist}\lambda + \sum_{m=-3}^6 D_{kst}^m \varphi_m + \sum_{p=1}^{10} L_{kst}^p \gamma_p - c_k$ is the set of other covariates defined in equation (4.4).

In addition to the state effects, state trends, industry effects, and time effects included in $Z_{ist}\beta$, we also control for industry trends ($t \cdot \pi_j$) in equation (4.5). These covariates are included to control for common shocks that may be correlated with wages and the rate of unionization at the state or industry level. For instance, states with more profitable (“high rent”) industries may pay higher wages and have higher unionization rates. After controlling for state and industry effects and trends, the primary source of identifying information left is state-industry specific trends in unionization rates and wages.

For example, consider the case of two industries (manufacturing and services) in two states (Michigan and South Carolina). Including state and industry trends and fixed effects controls for the fact that, for instance, wages and unionization rates may be declining faster in Michigan than in South Carolina because of adverse shocks in the manufacturing sector that account for a larger share of employment in Michigan. Thus, our empirical strategy leverages variation linked to the faster relative decline in unionization in Michigan’s manufacturing sector relative to South Carolina. We then look at whether this faster decline in the unionization rate is linked to a faster decline in non-union (or union) workers’ wages in the Michigan manufacturing sector.

Figure 4.4 illustrates these trends by Census Bureau Region and high- vs. low-unionization industries. Panel A shows that, in some industries (e.g., services and trade), the unionization rate is uniformly low across all regions. By contrast, there is much more regional (state) variation among high-unionization rate industries, shown in Panel B, like manufacturing, construction, transportation, education, and public administration. Notably, unionization rates have fallen fastest among high unionization industries in the Midwest relative to other regions.

Of course, there are possibly state-industry specific shocks that affect both wages and unionization rates. However, there is no particular reason to believe the impact of these shocks would follow the pattern expected from union wage compression effects. For example, [Chetverikov et al. \(2016\)](#) extend the [Autor et al. \(2013\)](#) analysis of the “China shock” to quantiles of the wage distribution. They found that among men, the wage impacts of

commuting zone-level trade exposure are uniform across most of the distribution, whereas among women, these wages impacts are monotonically increasing.¹⁵⁸ By contrast, the union wage effects literature (e.g. [Card, 1996](#); [Firpo et al., 2009](#)) indicates that unions have a relatively larger impact on workers' wages in the middle (or bottom) of the distribution, but little or even a negative impact on workers at the top of the distribution. Based on this evidence, it is natural to expect that union threat effects should be much more significant in the middle or bottom of the distribution than at the top. Finding such a pattern would be more supportive of a story based on unions' threat effects than on unmodelled state-industry shocks.

An additional way of probing the validity of the results is to examine whether the rate of unionization in other industries in the same state affects the wage distribution. For example, suppose the unionization rate declines in the construction sector but remains constant in manufacturing. In that case, we should not observe a decline in non-union manufacturing wages manufacturing in response to declining threat effects (a lower unionization rate) in the construction sector. Looking at the impact of changing unionization in other sectors can be viewed as a falsification test of our central hypothesis that threat effects are captured by unionization at the narrower state and industry level.¹⁵⁹ We explore these issues in Section 4.4.4 by including both the unionization rate in narrower and broader sets of industries.

As an alternative to union coverage, we also explored the possibility of using data on union organizing as a measure of the threat effect.¹⁶⁰ Unfortunately, information about industry is unavailable for recent election records (2010-2017). Furthermore, results based on union elections (for years that include industry codes) were fairly similar to those obtained when measuring the threat effect using the unionization rate at the industry-state level; in particular for men. This makes sense if union organizing explains part of the variation in U_{jst} conditional on state- and industry-specific trends. Figure 4.5 plots the average (unweighted) 2-year change in U_{jst} by year, alongside the net-certification rate for that two year period.¹⁶¹ The flow of newly unionized workers over a two year period,

¹⁵⁸Figure 2 of [Chetverikov et al. \(2016\)](#) shows that the point estimates of negative effects of the shock are larger than OLS estimates in the three bottom vingtiles and smaller in the four upper vingtiles of the female wage distribution. Figure 4.6 shows that negative effects larger than OLS estimates for male workers are found only in the three bottom vingtiles. However, in neither case are these differences statistically significant.

¹⁵⁹An alternative explanation could be that in economically close sectors—in the sense that workers often move between these sectors—the threat of unionization may also depend on what happens in these related sectors. But even in these circumstances, we should expect the unionization rate in these related industries to have a smaller impact on wages than the unionization rate in the industry where the worker is employed. By contrast, if unionization rates in other sectors capture spurious shocks hitting a local area, unionization rates in other sectors should be strongly correlated with wages, suggesting that our research design is invalid.

¹⁶⁰A special thank you to Henry Farber and John Ferguson for making their NLRB files publicly available. These files were updated to include post 2010 data that does not include industry codes.

¹⁶¹Net-certification is measured as the total number of eligible workers in certification elections won by a union minus the total number of eligible worker in decertification elections lost by a union, in a given

should correspond to the change in union coverage over that same period. Indeed, it is evident from Figure 4.5, that periods of increased union activity correspond to years in which the average union coverage (at the state-industry level) declines at a slower rate.

4.4 Data and estimation results

4.4.1 Data

Data from the 1979-2017 MORG CPS are used to estimate the distribution regressions. The sample selection criteria and variable definitions are similar to those used in DFL. Note that the union status of workers is only available from 1983 on. As in DFL, we use union status information from the 1979 May CPS matched with the May-August MORG to extend the analysis back to 1979. One difference relative to DFL is that we impute top-coded wages using a stochastic Pareto distribution (see [Firpo et al., 2009](#)). This imputation helps obtain a smoother wage density at the upper end of the distribution. In the case of workers paid by the hour, our wage measure is the hourly wage directly reported by the worker. The wage measure is average hourly earnings (usual earnings divided by usual hours of work) for workers not paid by the hour. Wages are deflated into constant dollars of 2017 using the CPI-U. See [Lemieux \(2006\)](#) for more information about data processing.

We use union coverage as our measure of unionization throughout. We focus only on observations with unallocated wages to avoid the large attenuation bias linked to the fact union status is omitted in the CPS wage imputation ([Hirsch and Schumacher, 2004](#)). The value of the minimum wage used in the estimation is the maximum of the federal and state minimum computed at the quarterly level.

Summary statistics are reported in Table 4.1. These statistics, as well as distribution regression models, are all weighted using CPS sample weights. As is well known, measures of overall inequality (the 90-10 gap, the standard deviation of log wages, and the Gini coefficient) and top-end inequality (the 90-50 gap) increase steadily over time. By contrast, low-end inequality (50-10) only increases between 1979 and 1988 when the minimum wage's real value was rapidly declining. Table 4.1 also shows that the rate of unionization declined much faster for men than women, and that the four years used to divide the sample (1979, 1988, 2000, and 2017) were at similar points in the economic cycle (comparable unemployment rates, especially for men).

state-industry pair. The 2-year net-certification rate is the total number of certified workers (on net) divided by total employment in that state-industry pair. This flow variable should explain some of the variation in U_{jst} over time.

4.4.2 Minimum wage effects

We separately estimate the distribution regression models for men and women over the 1979-88, 1988-2000, and 2000-17 periods. After some experimentation, we settled on specifications that allow for spillover effects up to 30 log points above the minimum wage in 1979-88 and 20 log points above the minimum wage in subsequent periods.¹⁶² Besides the minimum wage variables, other variables included in the models consist of state and year effects, state-specific trends, and the other covariates mentioned when discussing the model in equation (4.4).

Table 4.3 reports the estimated coefficients for the set of minimum wage dummies (φ_m) and heaping dummies corresponding to integer values of nominal wages (γ_p) for each of the six specifications (men and women for three time periods). The estimated coefficients for the other covariates are not reported for the sake of brevity, and standard errors are clustered at the state level. The estimated coefficient for being “right at” the minimum wage, φ_0 , is large and significant in all specifications, though it tends to decline over time. The coefficients capturing spillover effects, φ_m for $m > 0$, are also precisely estimated and tend to decline as we move further away from the minimum wage.

Interestingly, the coefficients associated with being below the minimum wage, φ_m for $m < 0$, tend to be small and often insignificant. This finding does not imply that the minimum wage does not reduce the wage density below the minimum. Since we are working with cumulative probabilities, a large value of the “spike” parameter φ_0 means that $Prob(Y_{ist} \geq y_k)$ is much lower for all wage cutoffs y_k that are below the value of the minimum wage. Finding small values of φ_m for $m < 0$ means we do not need a sizable additional reduction in probabilities to fit the data.¹⁶³ The heaping parameters are statistically significant and substantially improve the fit of the model, consistent with the descriptive evidence reported in Appendix Figure C.2.¹⁶⁴

There is also clear evidence that minimum wage effects are substantially larger in 1979-88 than in subsequent years. Unlike Lee (1999) and Autor et al. (2016), who use different estimation methods for different years, our method yields estimates based on the same method but for different years. The results suggest that Autor et al. (2016)’s conclusion

¹⁶²Spillover effects above these levels were not found to be statistically significant.

¹⁶³In the example used in equation (4.3), the probability when the wage cutoff is one wage bin below the minimum wage is $Prob(Y_{ist} \geq y_k) = \Phi(\beta_0 - c_k + \varphi_{-1} + \varphi_0 + \varphi_1)$. Compared to a case without a minimum wage where $Prob(Y_{ist} \geq y_k) = \Phi(\beta_0 - c_k)$, the probability is much larger and the complementary probability $Prob(Y_{ist} < y_k)$ much smaller when $\varphi_0 + \varphi_1$ is large, regardless of the value of φ_{-1} . Thus, the role of φ_{-1} is to “fine-tune” the features of the wage distribution below the minimum wage, as opposed to making sure only a few workers are observed there.

¹⁶⁴To reduce the number of parameters linked to heaping at integer values of nominal wages, we constrain the parameters for \$1 to \$10 to be the same except for \$5 in 1979-88, and \$10 later on, that are allowed to exhibit a larger spike.

that Lee overstated the importance of spillover effects is at least partly due to the fact their estimates are based on more recent data.

As it is challenging to interpret the magnitude of coefficients estimated using probit models, we transform the results into marginal effects that are reported in Figure 4.6. The marginal effects are computed as the difference between the predicted probabilities with and without a minimum wage. The counterfactual probabilities without a minimum wage are obtained by setting the minimum wage coefficients, φ_m , $m \in \{-3, 6\}$, to zero in equation (4.4) and using the estimated value of the other parameters to compute:

$$\hat{p}_{ist}^{k,c} = \Phi \left(Z_{ist}\hat{\beta} + y_k Z_{ist}\hat{\lambda} + \sum_{p=1}^{10} L_{kst}^p \hat{\gamma} - \hat{c}_k \right)$$

While we could compare this counterfactual distribution to the observed wage distribution, doing so would mix the impact of different values of the minimum wage, depending on state and year. Instead, we compute the predicted wage distribution for the median minimum wage in the relevant analysis period. The predicted wage distribution with a minimum wage is, therefore, calculated as:

$$\hat{p}_{ist}^k = \Phi \left(Z_{ist}\hat{\beta} + y_k Z_{ist}\hat{\lambda} + \sum_{m=-3}^6 D_{k,med}^m \hat{\varphi}_m + \sum_{p=1}^{10} L_{kst}^p \hat{\gamma} - \hat{c}_k \right)$$

where the minimum wage dummies D_{kst}^m in equation (4.4) have been replaced by the dummies $D_{k,med}^m$ corresponding to the median minimum wage.

Since distribution regressions yield estimates of cumulative probabilities, the estimated probability of log wages Y_{ist} lying in a given interval $[y_k, y_{k+1}]$ is the difference between the predicted cumulative probabilities \hat{p}_{ist}^k and \hat{p}_{ist}^{k+1} . Call this difference $\hat{Q}_{ist}^k = \hat{p}_{ist}^k - \hat{p}_{ist}^{k+1}$. Averaging out these individual probabilities over the entire sample yields the unconditional probability \hat{Q}_t^k , where $\hat{Q}_t^k = \frac{1}{N_t} \sum_i \hat{Q}_{ist}^k$. The marginal effects are obtained by comparing \hat{Q}_t^k to the counterfactual value $\hat{Q}_t^{k,c}$ that would prevail in absence of the minimum wage.¹⁶⁵ For clarity, we show the marginal effects, ME_t^k , in percentage terms: $ME_t^k = 100 \cdot (\hat{Q}_t^{k,c} - \hat{Q}_t^k) / \hat{Q}_t^k$.

Figure 4.6 shows that the marginal effects corresponding to the minimum wage spike are quite large. Depending on year and gender, the probability of being “right at” the minimum wage increases by 150 to 300%. Spillover effects in the first interval to the right of the minimum wage are also quite large, but decline as we move further above the minimum. Visually speaking, Figure 4.6 shows that spillover effects are substantially more important in 1979-88 than in subsequent periods. The same minimum wage coefficient, φ_m , will have

¹⁶⁵ $\hat{Q}_{ist}^{k,c} = \hat{p}_{ist}^{k,c} - \hat{p}_{ist}^{k+1,c}$ and $\hat{Q}_t^{k,c}$ is defined as $\hat{Q}_t^{k,c} = \frac{1}{N_t} \sum_i \hat{Q}_{ist}^{k,c}$.

a larger effect on probabilities when the minimum wage bites more, i.e., when it is relatively higher up in the distribution. This higher relative position explains in part why, for instance, the marginal effects are larger for women than men in 1979-88.¹⁶⁶ That said, the decline in marginal effects may not solely reflect a decline in the “bite” in the minimum wage since the estimated coefficients reported in Table 4.3 are declining too.

To further explore these issues, we use our estimates to assess how much of the decline in the marginal effects over time reflect a change in how the minimum wage is reshaping the wage distribution instead of a declining “bite” in the minimum wage. We do so using a simple measurement model. First, consider the “fraction affected” FA_t , which represents the fraction of workers who would be below the minimum wage mw_t in absence of a minimum wage (we ignore state variation in the minimum wage to simplify the exposition). The fraction FA_t can be computed by summing up the predicted probabilities $\hat{Q}_t^{k,c}$ over the wage bins that are below mw_t .

Using the terminology of Cengiz et al. (2019), we can think of the minimum wage as creating some “missing” mass below the minimum that gets redistributed as “excess” mass at or above the minimum wage. These effects can be captured using the following three parameters: $\delta_{1,t}$, the fraction of affected observations (FA_t) that move up to exactly the minimum wage; $\delta_{2,t}$, the fraction of affected observations that move above the minimum wage and contribute to spillover effects; and $\delta_{0,t} = 1 - \delta_{1,t} - \delta_{2,t}$, the fraction of observations that remain below the minimum wage due to imperfect compliance, measurement error, or subminimum wage (e.g., for tip workers). In this setting, the missing mass is given by $(1 - \delta_{0,t}) \cdot FA_t$.

To illustrate how to compute these “mass changing” parameters, consider the estimated wage distributions for 1979-88 with and without the minimum wage. The two distributions are reported in Appendix Figure C.3, Panel A, where we plot the values of \hat{Q}_t^k and $\hat{Q}_t^{k,c}$ recentered around the median value of the minimum wage. The fraction affected FA_t is the sum of the bars below the minimum wage for the counterfactual distribution that would prevail in absence of a minimum wage.¹⁶⁷ The difference between the cumulative sum of the bars for the two distributions below the minimum wage corresponds to the “missing mass” in Cengiz et al. (2019).

The fraction of affected observations redistributed at the minimum wage is the difference between the two bars at the minimum wage bin: $\delta_{1,t} = (\hat{Q}_t^{k_0} - \hat{Q}_t^{k_0,c}) / FA_t$. Likewise, the

¹⁶⁶For example, the coefficient for the “at the minimum wage” dummy is 13% larger for women than men in 1979-88 (0.557 vs. 0.494 in Table 4.3), while the corresponding marginal effect is 27% larger (277% vs. 218% in Figure 4.6a).

¹⁶⁷The figures only show the distribution up to 30 log points below the minimum, but lower wage values are also used to compute FA_t .

fraction of affected observations that gets redistributed above the minimum wage is:

$$\delta_{2,t} = \sum_{k=k_0+1}^{k_0+6} (\hat{Q}_t^k - \hat{Q}_t^{k,c}) / FA_t$$

where k_0 indicates the wage bin where the minimum wage lies ($y_{k_0} \leq mw_t \leq y_{k_0+1}$).

The results of this exercise are reported in the first panel of Table 4.4. As expected, the bite of the minimum wage, as summarized by FA_t , is larger for women and declines over time. Interestingly, the fraction of affected workers who are pushed up to the minimum wage, $\delta_{1,t}$, ranges from 0.30 to 0.34, and is remarkably stable across time and gender. Thus, differences in marginal effects reported in Figure 4.6 reflect differences in the fraction of affected workers, as opposed to the fraction of affected workers who are pushed up to the minimum wage.

By contrast, the fraction of observations ($\delta_{2,t}$) that are pushed up above the minimum wage declines over time, going from 0.37 in 1979-88 to 0.26 in 2000-17 for women, and from 0.37 to 0.20 for men. Thus, the decline in marginal effects associated with spillover effects reflects a combination of how many workers are affected, and how the minimum wage transforms the distribution. Since $\delta_{2,t}$ declines while $\delta_{1,t}$ remains stable over time, the fraction of affected workers remaining below the minimum, $\delta_{0,t}$, needs to increase since $\delta_{0,t} = 1 - \delta_{1,t} - \delta_{2,t}$. One possible explanation for this finding is that a substantial share of wages observed below the minimum wage is due to measurement error. This share may have grown over time as the minimum wage has moved further down in the left tail of the wage distribution where measurement error accounts for a large share of the wage density. We explore these issues in more detail in Section 4.4.3.

In this simple measurement model, we implicitly assume that while some affected workers are pushed above the minimum, the wages of “unaffected workers” –those already earning at least the minimum wage—are unchanged. This assumption is unrealistic, as these workers’ wages would likely go up due to the presence of spillover effects. We allow for this possibility by considering an alternative model where wage ranks are preserved when the minimum wage is introduced. In practice, this means that the wage bin where the minimum wage lies should first be filled by “affected workers”. Workers who were in the minimum wage bin before the increase are pushed above the minimum as this bin gets filled by workers from lower down in the distribution. We define the fraction of these workers being pushed up as $\delta_{3,t}$.

Not surprisingly, the results reported for this “rank preservation” model in Panel B of Table 4.4 indicate that a higher fraction of affected workers is now pushed to the minimum wage bin. We also find that, in most cases, all workers previously in the minimum wage bin are now pushed above the minimum wage ($\delta_{3,t}$). Although the mechanisms underlying

these spillover effects are different from those discussed earlier, our main conclusions about why spillover effects have declined over time remain. Spillover effects get smaller over time due to a decline in the fraction of workers affected and changes in the way the minimum wage shifts the distribution, as summarized by the δ parameters.

4.4.3 Are spillover effects real or just due to measurement error?

Measurement error is another reason why Autor et al. (2016) argue that Lee (1999) may have overstated the contribution of minimum wage spillover effects on inequality growth. Measurement error may indeed result in spurious spillover effects if a fraction p of workers paid exactly the minimum wage misreport their wages, and if measurement error follows a continuous zero mean distribution (e.g., a normal distribution with mean 0 and variance σ^2). When the actual minimum wage spike is large, and many workers misreport their wages (i.e., p is large), we may expect to see an abnormal concentration of observations just above and below the minimum wage. These spurious spillover effects potentially overstate the equalizing effects on the minimum wage on the wage distribution.

We use two approaches to address this important concern. We first note that not all forms of measurement error necessarily result in spurious spillover effects. As noted above and illustrated in Appendix Figure C.2, there are large spikes in the wage distribution at integer values of wages. While some of these spikes may be real (see Dube et al., 2020), many workers likely misreport their wages by rounding them off to the nearest integer. Unlike the type of measurement error discussed above, rounding off at the nearest integer may overstate or understate spillover effects.¹⁶⁸

Whether or not measurement error linked to rounding off at integer values results in spurious spillover effects is ambiguous, and depends on the distribution of the difference between the minimum wage and the nearest integer. We can empirically evaluate the consequences of this form of measurement error by comparing our estimated spillover effects where we do control for rounding off with the set of dummies L_{kst}^p to what we would find without controlling for these dummies. The results of this comparison for 1979-88 are reported in Appendix Figure C.4. We find that the estimated spillover effects are, if anything, smaller when we do not control for the rounding off. Thus, this potentially important form of measurement error is unlikely to generate spurious spillover effects

¹⁶⁸For example, if the minimum wage is itself an integer value (e.g. \$10), the size of the spike may be overstated, and spillover effects understated, if worker paid a bit above the minimum (e.g. \$10.15) round off their reported wage to the value of the minimum wage. If the minimum wage is a bit above an integer value, the rounding off may increase the false reporting just under the value of the minimum. The case where spillover effects would be overstated is when the minimum wage is a bit below an integer value (e.g., \$9.80), and workers at the minimum falsely report the integer value, leading to spurious extra mass just above the minimum wage.

that would overstate the contribution of changes in the minimum wage to the increase in inequality.

Unlike measurement error linked to rounding off at integer values, more standard forms of measurement error like the one discussed above unambiguously overstate the magnitude of spillover effects. To assess the bias's magnitude, we follow [Autor et al. \(2016\)](#), who use the observed wage distribution below the minimum wage to construct bounds for spillover effects. One extreme assumption is that all wages observed below the minimum are due to noisy wage reports by workers earning exactly the minimum wage. Under the additional assumption that measurement error is symmetric, the lower tail of the distribution can be used to adjust the distribution above the minimum wage. The corrected wage distribution is obtained by moving up all the lower part of the distribution to the minimum wage and moving down in a symmetric way the same fraction of observations from the part of the distribution just above the minimum. We also consider an alternative assumption where measurement error only accounts for half of the wage distribution observed below the minimum wage, the remainder being due to non-compliance and subminimum wages. We refer to this case as a “partial” adjustment for measurement error.

Panels B and C of Appendix Figure C.3 illustrate the impact of the measurement error correction on the wage distribution for the 1979-88 period. As we only consider measurement error for workers paid the minimum wage, the counterfactual distribution without a minimum wage (red bars in the figures) remains unchanged when the different adjustment factors are applied.¹⁶⁹ As we move to the partial (50 percent) measurement error adjustment in Panel B and the full adjustment in Panel C, the wage distribution becomes increasingly concentrated right at the minimum wage. By the same token, the density just above and below the minimum wage declines as the measurement error corrections moves some of the density to the spike.

Spillover effects are measured as the difference between the two distributions with and without the minimum wage. Adjusting for measurement error reduces these effects, but they generally remain positive. That said, the figures suggest that even a reasonable amount of measurement error substantially understates the size of the minimum wage spike and overstates the importance of spillover effects.¹⁷⁰

¹⁶⁹It is unrealistic to assume that measurement error should affect only those paid the minimum wage rather than all workers. That said, adding measurement error to an otherwise smooth distribution —like the counterfactual distribution without a minimum wage shown in Appendix Figure C.4 — should not change the shape of the distribution very much. By contrast, when there is a massive spike at the minimum in the actual wage distribution, adding measurement error may substantially change the distribution's shape by smoothing out the spike. So although our measurement error model is highly simplified, it illustrates the consequences of measurement for estimating the relative size of the minimum wage spike and the spillover effects.

¹⁷⁰To assess how reasonable the implied measurement error is, consider the case with partial (50 percent) adjustment. In the case of women, we find that 26 percent of minimum wage workers misreport their wages, and that the variance of errors is equal to 0.006. The variance of measurement error is equal to the product of

The impact of the measurement error corrections is summarized in Panels C and D of Table 4.4 using the partial (50 percent) adjustment for measurement error. As in Appendix Figure C.4, adjusting for measurement error increases the fraction $\delta_{1,t}$ of affected workers who are pushed up to the minimum wage. Likewise, the “missing mass” increases as a smaller fraction $\delta_{0,t}$ remains below the minimum wage after controlling for measurement error. Spillover effects, captured by the parameters $\delta_{2,t}$ and $\delta_{3,t}$, also decline. That said, the main finding obtained without correcting for the minimum wage remains. The spillover effects documented in Figure 4.6 decline over time due to a combination of a smaller bite of the minimum wage, and a change in the way the minimum wage shifts the wage distribution, as summarized by the δ parameters.

It should also be noted that although spillover effects may partly reflect measurement error, this does not necessarily reduce the minimum wage’s overall impact on the wage distribution. Correcting for measurement error also increases the fraction of workers whose wages are moved up from below the minimum to the minimum wage spike, as observed when comparing Panels C and D of Table 4.4 to Panels A and B. This shift increases the equalizing effects of the minimum wage, which compensates for the more modest equalizing effects linked to spillover effects.

4.4.4 Distribution regression estimates of the effect of unionization

Before presenting our main estimates of the effect of the state-industry unionization rate on the wage distribution, we report results from simple OLS regression of the log wage on the unionization rate for union and non-union workers in Table 4.5. This provides a simple way of summarizing the average union threat effects over the whole distribution and exploring the robustness of the results to alternative definitions of the level at which threat effects operate. In our preferred specification, we use the unionization U_{jst} at the industry-state level using 11 industry categories.¹⁷¹ Using a narrower set of industries would be challenging due to sample sizes. In Table 4.5, we present the regression models for men and women pooled together to simplify the exposition. All models include the set of covariates mentioned when discussing the distribution regression model of equation (4.5). Recall that the covariates include a full set of state, industry, quarter and year dummies, as well as state- and industry-specific linear trends.

Looking first at non-union workers in columns (1) to (4), the results indicate that a one

these two numbers and represent 18 (7) percent of the total variance of log wages in 1979 (1988). The comparable numbers for men are 10 and 7 percent in 1979 and 1988, respectively. These figures are of the same order of magnitude as those found using validation surveys (see, e.g., [Bound and Krueger, 1991](#)).

¹⁷¹The 11 industries are the primary sector, construction, manufacturing, transportation and utilities, wholesale and retail trade, financial services, business and professional services, health and welfare services, personal services, education services, and public administration.

percentage point increase in unionization rate at the state-industry level increases wages by 0.11 to 0.15 percent in most specifications. The estimates are essentially unchanged when we also include the state unionization rate in column (2), or a full set of state-year interactions in column (3). Although it is challenging to identify the effect of the state-level unionization rate (the estimates are noisy and unstable) when we include state linear trends, the results suggest that spillover effects mostly take place at the narrower industry level.

A similar conclusion is reached when including the unionization rate at a broader industry level in column (4). The three broad industrial categories considered are i) traditional high-unionization private sector industries (construction, manufacturing, transportation, and utilities), ii) high-unionization public sector industries (educational services and public administration), and iii) low-unionization services, as well as primary industries. After controlling for the unionization rate in these broader industries, we only rely on the variation within these broader industries to identify the unionization rate's effects at the narrower industry level. Although the estimates are less precise than in the other columns, the overall magnitude of the effects remains relatively unchanged, ranging from 0.07 in 2000-17 to 0.18 in 1979-88. By contrast, the unionization rate's effects at the broader industry level are small and not statistically significant. These results provide strong evidence that spillover effects primarily arise at the narrower industry-state level used in the distribution regression estimates, presented below.

Interestingly, the results reported in columns (5) to (8) indicate that union spillover effects are larger for union than non-union workers, i.e., the union wage gap tends to be larger when the unionization rate is higher. For union workers, the estimated spillover effects at the narrower industry-state level are relatively stable across specifications. One exception is the model (column 8) with the unionization rate at the broader industry level in 1988-2000 and 2000-2017, where the two unionization rate variables have comparable effects. We nonetheless conclude that, on balance, most spillover effects appear to be occurring at the narrower industry level, and we focus on this measure of unionization in the remainder of the analysis.

Table 4.6 reports estimates from the distribution regression models with the state-industry-year rate of unionization added as an explanatory variable. Again, we show results separately for union and non-union workers. As noted above, we model changing impacts over the wage distribution by interacting the unionization rate with a quartic function in y_k (normalized to zero at the midpoint of the y_k range). All models include a set of industry trends in addition to the other explanatory variables listed in Section 4.2.2. The models are estimated separately for union and non-union workers for two reasons. First, we want to allow for different effects of the unionization rate (and other covariates) for these two groups. Second, and as discussed earlier, estimating separate models for union

and non-union workers is essential for computing standard counterfactual experiments illustrating the contribution of de-unionization to changes in the wage structure.¹⁷²

Panel A of Table 4.6 reports the estimated effect of the unionization rate for non-union workers. The main effect of the unionization rate is large and statistically significant in all three time periods. The unionization rate's estimated effect is substantially smaller for women, especially in the earlier periods. Panel B shows that the unionization has a larger effect on union workers, suggesting that the union wage gap increases with the unionization rate.

While most of the interactions between the unionization rate and the polynomials in y_k are statistically significant, it is difficult to infer the shape of the estimated effects from the results reported in Table 4.6. To facilitate interpretation, we translate the estimated parameters for non-union workers into wage impacts at different points of the distribution by considering a 1% increase in the unionization rate. The wage effects are obtained by first comparing the CDF computed from the distribution regressions —using the observed rates of unionization— to the counterfactual CDF that would prevail if the unionization rate was one percentage points higher. The horizontal distance between the two CDFs indicates by how much wages change at each percentile of the distribution under this counterfactual experiment.

The results of this exercise are displayed in Figure 4.7, panels A and B. The threat of unionization has the largest impact in the lower middle of the distribution and tends to be substantially larger for men than women. The effect is positive over most of the distribution before turning negative around the 80th percentile.

We also report more traditional direct or “shift-share” effects of unionization in Panels C and D. These effects are computed by contrasting the observed wage distribution with the counterfactual distribution that would prevail if the unionization rate was increased by one percentage point. The counterfactual distribution is computed by reweighting union and non-union observations to increase the conditional probability of unionization by one percentage point.¹⁷³

Interestingly, both shift-share effects and threat effects (for non-union workers) reported in Figure 4.7 are hump-shaped. The hump-shape feature of the direct (or shift-share) effect is similar to the findings of [Firpo et al. \(2009\)](#) who use RIF-regression to estimate union wage effects (see also Panel A of Appendix Figure C.6). Intuitively, the pattern of union wage effects —positive on average but declining in the upper part of the distribution— is

¹⁷²Appendix C.4 provides comparable results using the RIF-regression methodology of [Firpo et al. \(2009\)](#).

¹⁷³The reweighting factor used in DFL is $\phi(X) = U \frac{Pr^c(U=1|X)}{Pr(U=1|X)} + (1 - U) \frac{Pr^c(U=0|X)}{Pr(U=0|X)}$ where U is a union status dummy and X are covariates. The counterfactual probability of unionization, $Pr^c(U = 1|X)$, used in DFL is based on other years, while we use $Pr^c(U = 1) = Pr(U = 1|X) + 0.01$ (and $Pr^c(U = 0) = Pr(U = 0|X) - 0.01$) in the counterfactual experiment considered here.

consistent with other evidence on the effect of unions on the wage structure. For instance, [Card \(1996\)](#) shows that the union wage premium is positive on average, but declines over the skill distribution.

It is not as intuitive, however, to see why the union effect first grows before reaching a peak around the middle of the distribution. Part of the story is that changes in the rate of unionization have little impact at the bottom of the distribution where wages mostly depend on the minimum wage. Another part of the story is that very few workers are unionized at the bottom of the distribution. The issue is discussed in more detail using an example with uniform distributions presented in Appendix C.5. Note that the hump-shaped pattern of union effects has important implications on how de-unionization affects the shape of the wage distribution. It implies that unionization substantially reduces the 90-50 gap, but also slightly increases the 50-10 gap. Interestingly, DFL reach a similar conclusion using a reweighting approach, as we do with the distribution regression method (see below).

The similarity in the shape of the threat effects and the traditional shift-share effects is remarkable, given that these effects are computed using very different procedures. The results are consistent with non-union employers trying to emulate the union wage structure in response to the threat of unionization. This supports the view that the effects of the unionization rate at the state-industry-year-level capture union threat effects, as opposed to un-modelled state-industry shocks that may affect both wages and unionization. It is also re-assuring to note that, using different approaches and data, [Farber et al. \(2018\)](#) also reach the conclusion that unions have both a direct and indirect (threat) effect on wages. Despite the challenges of finding a credible instrument for the rate of unionization, the evidence presented here and in [Farber et al. \(2018\)](#) strongly suggests that the effect of union on wages goes beyond the direct effect that has been the focus of most of the unions and wage inequality literature.

4.5 Decomposition results

We are now able to estimate how much of the change in the wage distribution over the 1979-2017 period can be accounted for by changes in the minimum wage and the rate of unionization in the presence of spillover effects. In the case of the minimum wage, we first compute counterfactual probabilities by replacing the observed minimum wages in the end period (say 1988) with the minimum wage in the base period (say 1979). For example, for each individual i in year 1988, the predicted cumulative probabilities estimated using the distribution regressions are:

$$\hat{p}_{is88}^k = \Phi \left(Z_{is88} \hat{\beta} + y_k Z_{is88} \hat{\lambda} + \sum_{m=-3}^6 D_{ks88}^m \hat{\phi}_m + \sum_{p=1}^{10} L_{kst}^p \hat{\gamma}_p - \hat{c}_k \right),$$

while the counterfactual cumulative probabilities are:

$$\hat{p}_{is88}^{k,c} = \Phi \left(Z_{is88} \hat{\beta} + y_k Z_{is88} \hat{\lambda} + \sum_{m=-3}^6 D_{ks79}^m \hat{\phi}_m + \sum_{p=1}^{10} L_{kst}^p \hat{\gamma}_p - \hat{c}_k \right).$$

In Section 4.4.2 we introduced \hat{Q}_{ist}^k , the predicted interval probability that individual i is in a given interval $[y_k, y_{k+1}]$, where $\hat{Q}_{ist}^k = \hat{p}_{ist}^k - \hat{p}_{ist}^{k+1}$. Averaging these probabilities over all individuals in 1988 yields the predicted probability \hat{Q}_{88}^k , and its counterfactual counterpart $\hat{Q}_{88}^{k,c}$. We can then compute the various counterfactual statistics of interest in 1988 by reweighting observations with a wage in the interval $[y_k, y_{k+1}]$ using the reweighting factor $\hat{\phi}_{88}^k = \hat{Q}_{88}^{k,c} / \hat{Q}_{88}^k$.¹⁷⁴ We use the same procedure for the periods 1988-2000 and 2000-17.

To isolate the contribution of spillover effects, we use DFL's "tail pasting" procedure where the distribution in the year with a lower minimum wage (say 1988) is replaced by the distribution in the year with a higher minimum wage (say 1979) for wages at or below the higher minimum. The difference between our main estimates (that include spillover effects) and those obtained using this alternative procedure represents the contribution of spillover effects.

In the case of the decline in unionization rate, we also compare the predicted probabilities obtained using observed values of the unionization rates in the end period (say 1988) to the counterfactual probabilities obtained using the base period unionization rate. As such, the procedure is very similar to the one described above for the minimum wage. As in the minimum wage case, and for the sake of comparison with DFL, we first compute the contribution of de-unionization without spillover effects using DFL's reweighting (shift-share) procedure. More specifically, we first reweight data in the end period (say 1988) to have the same distribution of unionization as in the base period conditional on covariates, and then add spillover effects to the reweighted distribution using the procedure we just described.

Figures 5 to 7 report the actual and counterfactual distributions corresponding to the three periods of analysis 1979-1988, 1988-2000, and 2000-2017. In each figure, panel A shows the counterfactual distribution corresponding to a model where the minimum wage is held constant at the base period level, and spillovers are accounted for. Panel B then shows the

¹⁷⁴We use this procedure to preserve the observed wage distribution within each wage bin. With very narrow bins we could simply use mid-points to calculate distributional statistics of interest such as the Gini coefficient, the wage density, etc. We instead use slightly wider wage bins to simplify the estimation and perform a bin by bin reweighting to construct a counterfactual distribution that also includes within-bin wage dispersion.

counterfactual corresponding to the base period's minimum wage and unionization rate, accounting for spillovers in both cases. Thus, a comparison of the two panels highlights the interaction between these two forms of spillovers. The inequality measures corresponding to these distributions can be found in Tables 6a and 6b, respectively, for men and women. The tables report the results of additional models, including counterfactuals without spillover effects. The shaded areas in the figures indicate the range (from the 5th to the 95th percentile) of variation in minimum wages in the base (red area) and end (blue area) years.

As in [Lee \(1999\)](#), spillover effects substantially increase the contribution of the decline in the real minimum wage to increasing inequality over the 1979-1988 period (see Figure 4.8). Comparing our results with spillovers to DFL's "tail pasting" method, we predict a counterfactual with far greater mass above the 1979 minimum wage level and less mass at the minimum wage. This occurs because the model accounts for the fact that with spillover effects, some of the observed 1988 mass below the 1979 minimum wage level results from lower spillover effects and moves above the 1979 minimum wage level in the counterfactual. For women, accounting for these spillovers is particularly important. It doubles the increase in the standard deviation of log wages and the Gini coefficient explained by this institutional factor.

For men, the decline in the unionization rate explains a large share of the declining wage density in the middle of the distribution between 1979 and 1988 (Figure 4.8, panel B). Moreover, because the decline in unionization can explain some of the increasing mass in the lower tail of the 1988 distribution, including unionization (and its spillovers) in the model reduces the share of the mass explained by the minimum wage. The model with only minimum wage spillovers may therefore overfit the 1979 distribution in the counterfactual. For women, the minimum wage effect still dominates. Combined, changes in these two institutional factors account for 101% (74%) of the change in the 50-10 wage gap for men (women) between 1979 and 1988.

Between 1988 and 2000 real minimum wages remain relatively constant (see Figure 4.1). Therefore, the minimum wage cannot explain the decline in inequality at the bottom of the wage distribution (the decline in the 50-10 gap). However, the decline in unionization explains some of the changing mass in the middle of the distribution and accounts for a large share of the increase in the 90-50 gap. Accounting for union spillovers doubles the share of the increase in the 90-50 wage gap explained by unions. This result is consistent with the hump-shaped union threat effects discussed earlier.

Minimum wages rise across a number of states between 2000 and 2017. Figure 4.10 (Panel A) shows that spillover effects can explain some of the wage gains above the 2017 minimum wage values. For men, declining unionization continues to explain a share of the declining mass in the middle of the distribution, and taken together, both institutional

factors explain 99% of the decline in 50-10 wage gap over this period. Women experience very little change in the 50-10 gap over this period.

As in DFL, de-unionization has a modest impact on the female wage distribution, in large part because unionization declines much less for women than men. Table 4.1 shows a relatively modest 6 percentage point decline in unionization rate among women, compared to a 21 percentage points decline for men. Unsurprisingly, we find the largest effects of declining unionization among men, in particular between 1979 and 1988.¹⁷⁵ Moreover, as the unionization rate declines, so does the impact of unions on the wage distribution coming largely from a decline in the unions' threat effect. Over the entire 1979 to 2017 period, declining unionization explains close to 40% of the increase in the 90-50 wage differential for men, with spillover effects accounting for about half of the union effect. Overall, our model explains 53% (28%) of the increase in the standard deviation of log wages for men (women) and 49% (27%) of the increase in the Gini coefficient between 1979 and 2017.

Our results on the contribution of unions to changes in inequality echo those of [Farber et al. \(2018\)](#) who consider a longer time period. Using different data and estimation methods, they find that the direct effect of unionization accounts for 46 percent of the decline in the 90-10 gap between 1936 and 1968, and 16 percent of its increase between 1968 and 2014. The latter figure is similar to our finding that the direct effect de-unionization accounts for 13 percent (0.043 out of a 0.322 change in Panel D of Table 4.7) of the growth in the 90-10 gap for men between 1979 and 2017. Like us, [Farber et al. \(2018\)](#) also find that spillover effects substantially increase the role of unions in the change in inequality. Taken together, these evidences support the conjecture of [Freeman \(1996\)](#) that spillover effects magnify the effect of de-unionization on inequality growth.

4.6 Conclusion

This paper uses an estimation strategy based on distribution regressions to quantify the contribution of union and minimum wage spillover effects to U.S. wage inequality growth

¹⁷⁵In the case of men, the contribution of de-unionization to the growth of inequality is very similar to recent estimates in [Card et al. \(2020\)](#). Table 4.7 shows that de-unionization (without spillover effects) accounts for 0.014 of the 0.118 increase in the standard deviation of log wages between 1979 and 2017. Using a different approach (counterfactual variances in absence of unionization) for a different period (1973 to 2015), [Card et al. \(2020\)](#) find that de-unionization accounts for 0.015 of the 0.121 increase in the standard deviation of log wages (their variance estimates reported in Table 4.1 have been transformed in standard deviations). In the case of women, like [Card et al. \(2020\)](#), we find only small effects of de-unionization on changes in inequality in most periods. One exception is that [Card et al. \(2020\)](#) find a more substantial equalizing effect of unions on female wage inequality in 2015 than in other years. We are unsure of the source of difference between the two studies and suspect it has to do with the control variables used in the estimation (we control for industries and occupation while they do not).

over the 1979-2017 period. The first important finding is that the continuing decline in unionization from 1988 onwards has contributed to continuing wage inequality growth, especially in the upper middle of the distribution. A second important finding is that accounting for spillover effects substantially increases the contribution of both types of institutional changes to wage inequality. These findings confirm and strengthen DFL's conclusion that labour market institutions have played a central role in U.S. wage inequality dynamics since the late 1970s. Our analysis of the impact of minimum wages with spillover effects over a time-period spanning more than 35 years also allows us to understand better why previous findings —[Lee \(1999\)](#) and [Autor et al. \(2016\)](#) — may appear contradictory at first blush. The period from 1979 (or indeed 1973) to 1988 saw a substantial (30 percent) and permanent decline in the value of the federal minimum wage, which was the prevailing one in almost all states at the time. By contrast, after 2005, many states increased their minimum wages above the federal one, resulting in smaller and often transitory changes in the effective minimum wage for a large fraction of the workforce. These important differences in the magnitude and persistence of minimum wage changes over time may help explain why [Lee \(1999\)](#) found larger spillover effects in the pre-1990 period, while [Autor et al. \(2016\)](#) found smaller effects in more recent years. Recent research by [Aaronson et al. \(2018\)](#) suggests that the dynamic employment response to minimum wage changes depends on these changes' magnitude and persistence. Improved understanding of how changes in the wage distribution depend on the dynamics of minimum wage changes should be an important topic of future research.

Likewise, it would be useful to understand better the economic forces behind the spillover effects of unionization estimated in this paper. We interpret these findings as evidence of (declining) union threat effects. An alternative interpretation is that non-union firms that compete with higher-paying union firms need to pay higher wages in imperfectly competitive labour markets than if there were no union employers in their relevant market.¹⁷⁶ In this setting, the rate of unionization positively impacts non-union wages, even if there is no longer a threat of unionization. Consistent with this view, [Benmelech et al. \(2018\)](#) find that firms' market power tends to depress wages, but this connection is substantially weaker when the unionization rate is higher. Future research based on rich employer-employee data could help better understand the connection between the wages paid by union and non-union firms, and shed light on the mechanisms behind the union spillover effects documented in this paper.

¹⁷⁶[Card et al. \(2018\)](#) and [Lamadon et al. \(2020\)](#) present models with imperfect competition where firms pay wages that depend on an index of the wages paid by their competitors. See also [Manning \(2003\)](#).

4.7 Tables and Figures

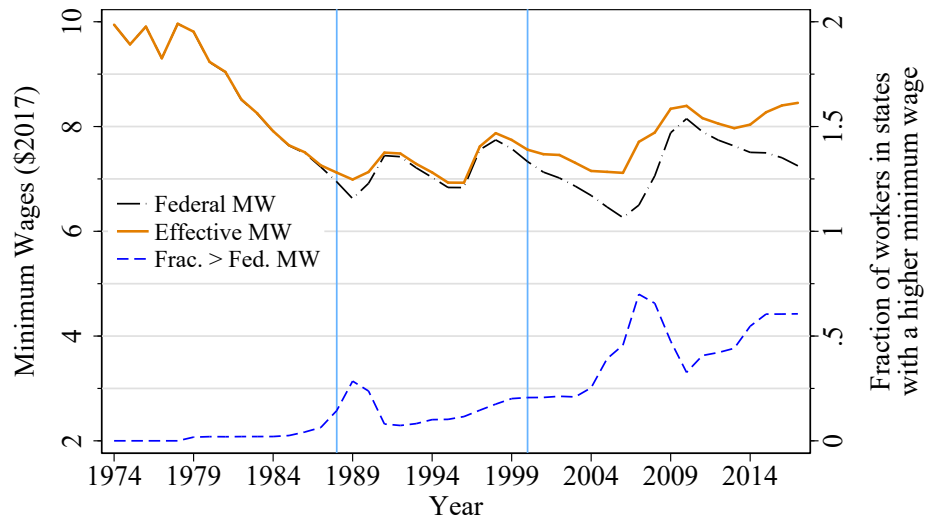


Figure 4.1: Real value (\$2017) of the minimum wage (MW) and fraction of workers in states with a higher minimum

Note: The federal minimum wage is based on official monthly federal minimum wage levels. The effective minimum wage is the maximum of the federal and state minimum wage, where states are weighted according to their relative employed populations, using CPS data. The fraction of workers residing in states with minimum wages above the federal level is calculated using the CPS population weights. The vertical lines demark the periods of analysis: 1979-1988, 1988-2000, and 2000-2017.

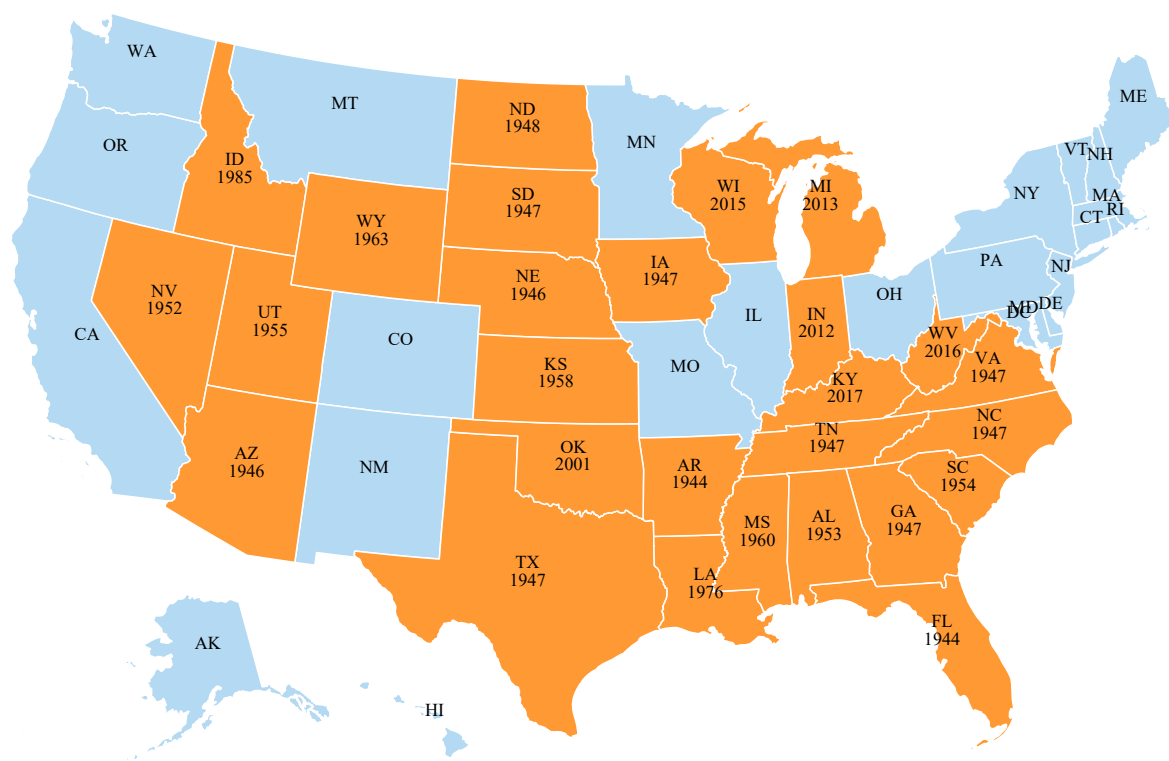


Figure 4.2: Map of Right-to-Work adoption.

Note: Orange states denote those which have adopted RTW laws, while blue states have not. The corresponding year of adoption is shown with the state.

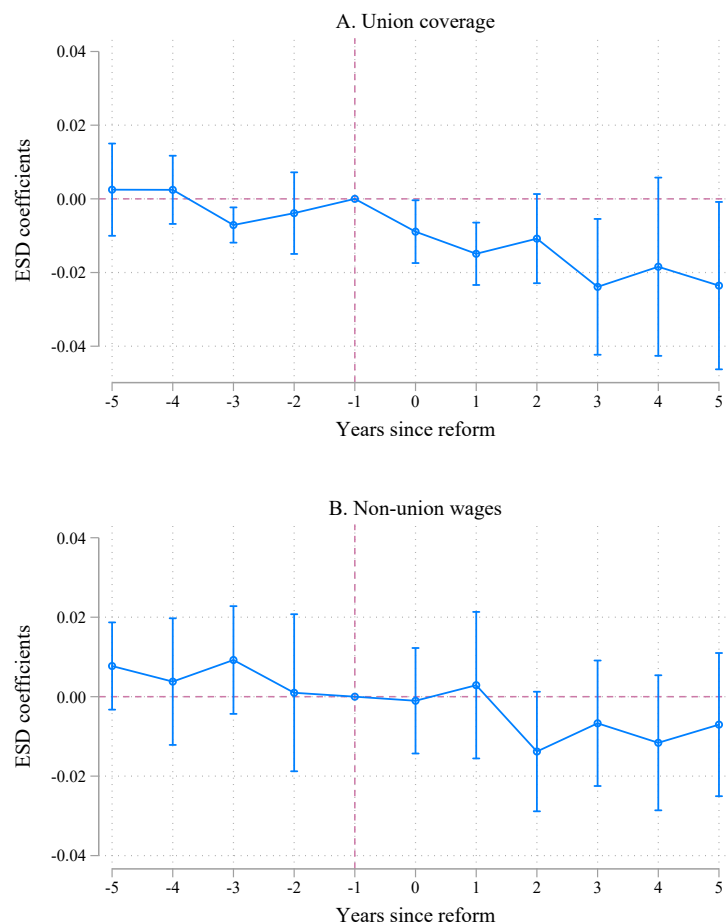


Figure 4.3: Coefficients from event-study of Right-to-Work laws.

Note: The figure depicts coefficients from two event-study designs, estimated using the CPS (1991-2019). The first examines the passing of RTW laws on union coverage, while the second examines on non-union (unallocated) wages. The sample includes all states. The event-time coefficients correspond to passing of laws in Oklahoma, Indiana, Michigan, Wisconsin, Kentucky, and West Virginia. Adjustments are made to account for the passing of Bill 10 in Wisconsin, extending RTW to public sector workers. All models control for the standard set of covariates described in Section 4.2.2 and standard errors are clustered at the state level. Only non-allocated wages are used in Panel B.

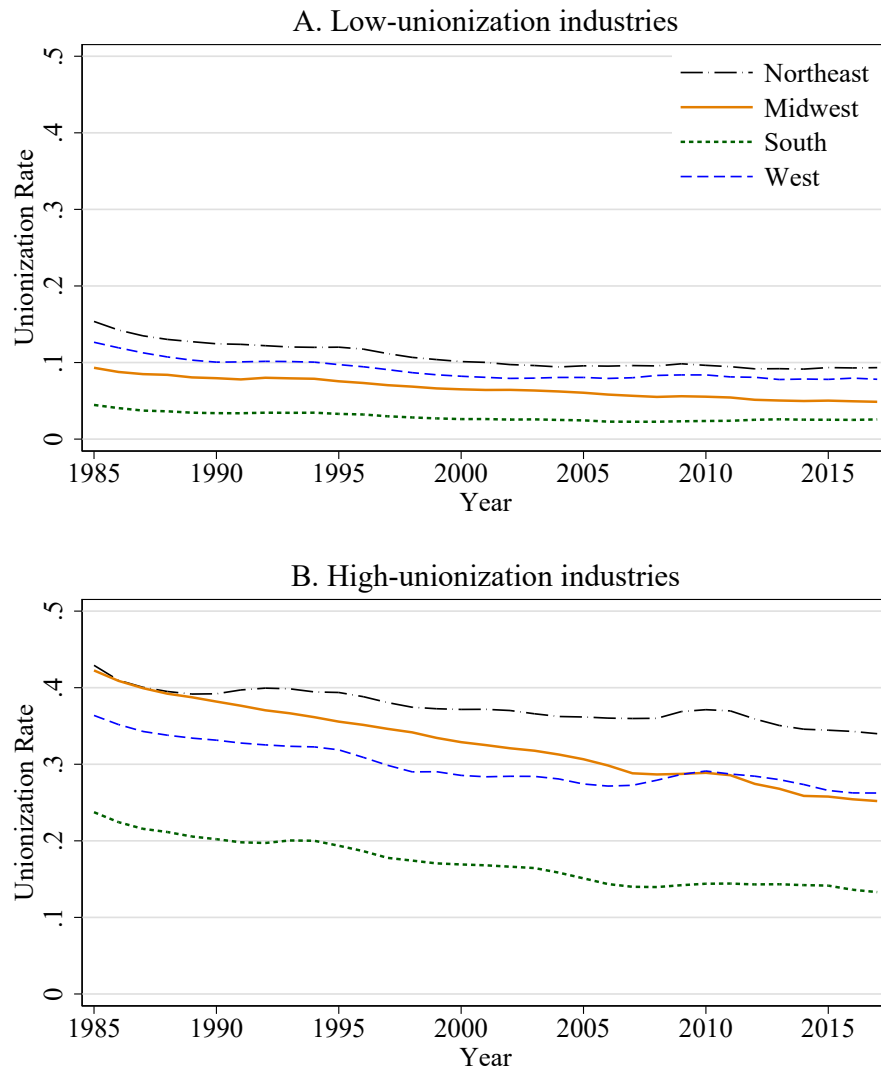


Figure 4.4: Trends in unionization rates across Census Bureau division, and low- and high-unionization industries: Men and women combined

Note: Figure depicts a three-year moving average of the coverage rate of both private and public sector salaried workers by Census Bureau Division, using data from the 1983-2017 CPS. High unionization industries include construction, manufacturing, transportation and utilities, education services, and public administration. Low unionization industries include primary sector, wholesale and retail, financial services, business and professional services, health and welfare services, and other services.

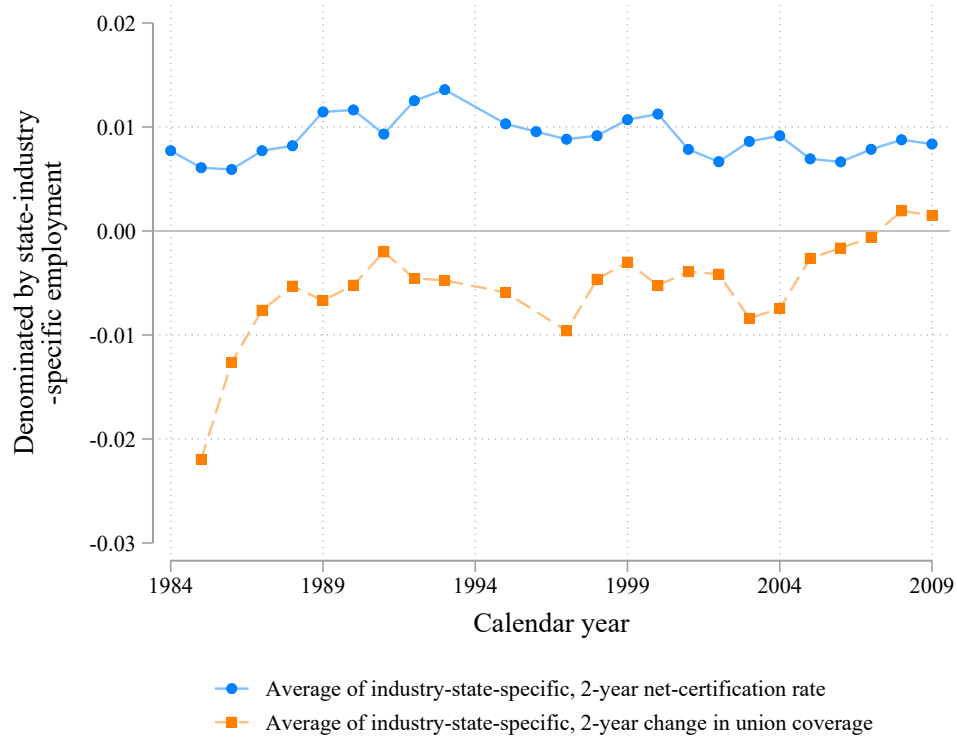


Figure 4.5: Change in union coverage and net certification over time.

Note: Figure depicts two curves. The first is a flow variable, measuring the number of workers unionized through elections in a given state-industry pair (as a share of employment). This is measured over a two year period to smooth out any noise in the data. The second is a change in stock variable, denoting the change in union coverage over a two-year period. The election data comes from NLRB records (1983-2009), while union coverage and employment are measured using the CPS (1983-2009). The measure of union coverage depicted here is the exact measure used in this analysis. Data is missing for 1994 (and 1996) as 1994 does not identify allocated wages and is therefore excluded from the analysis.

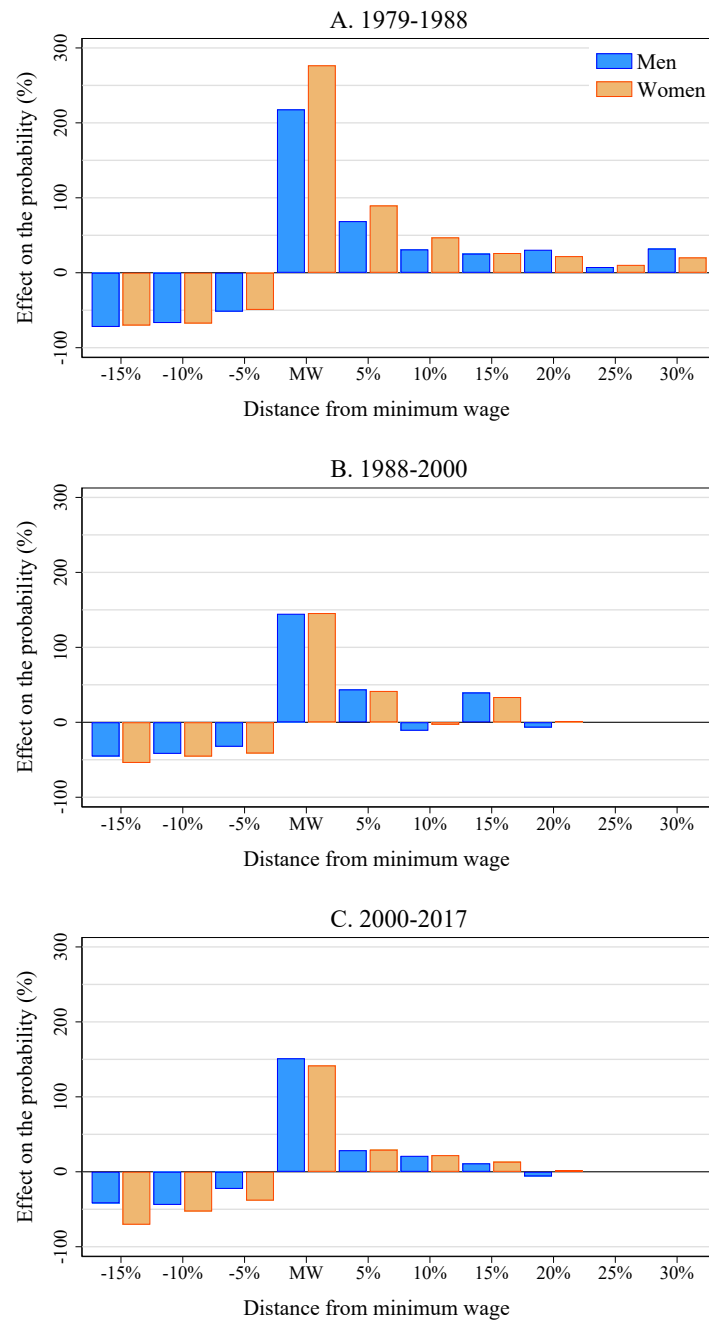


Figure 4.6: Marginal effects of minimum wages

Note: Marginal effects of the minimum wage on the (log) wage distribution are calculated using the difference in the average predicted probability of at each wage-bin under the median minimum wage (during the time period), relative to the counterfactual of no minimum wage.

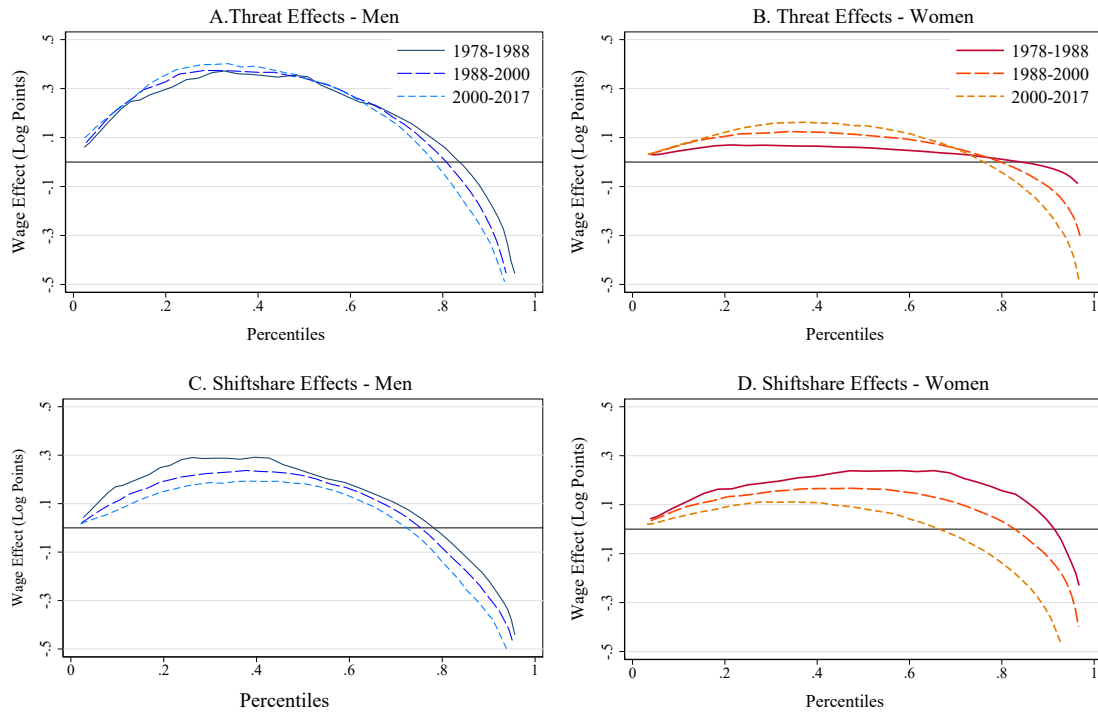


Figure 4.7: Marginal effects of a 1% increase in the unionization rate

Note: The threat effects indicate the wages changes (log points) in the non-union distribution in response to an increase in union coverage of 1%, keeping the union and non-union distributions unchanged. The effects are estimated as the average changes in the predicted probability under the observed coverage rate, and a coverage rate 1%-point higher.

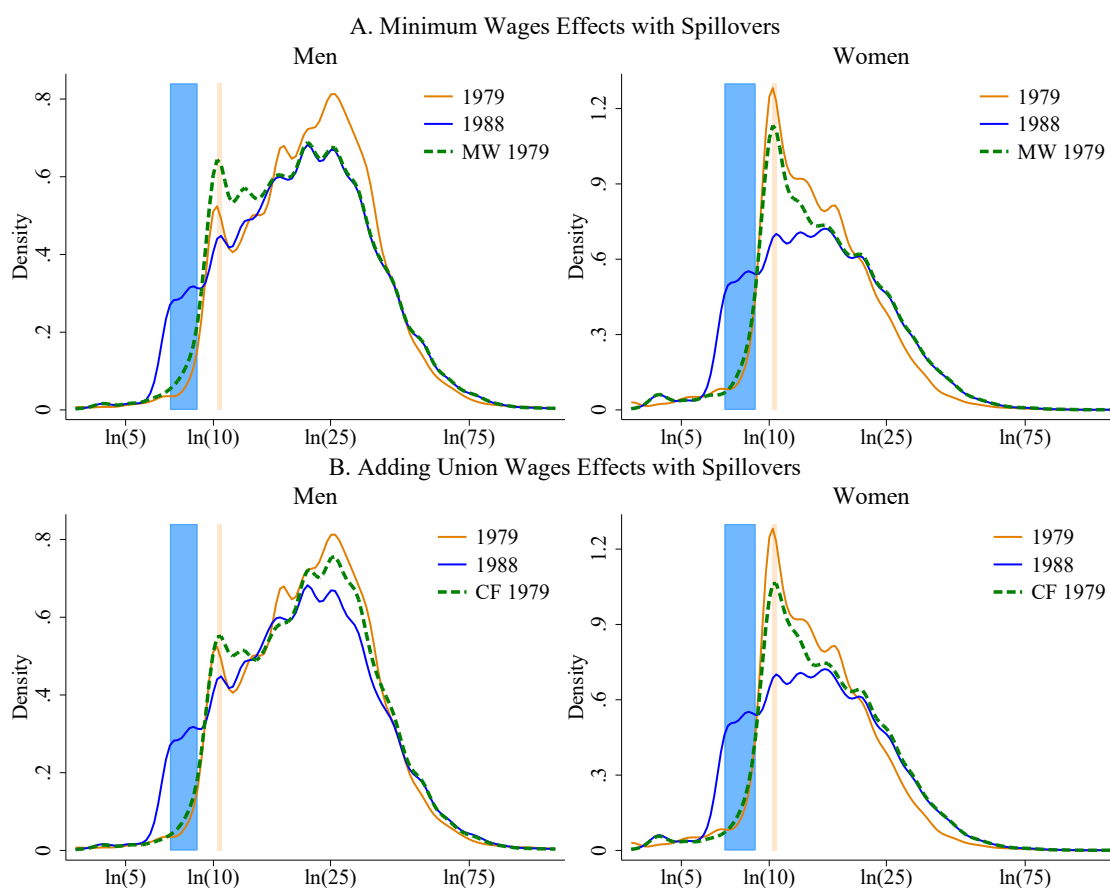


Figure 4.8: Counterfactual densities: 1979-1988

Note: In Panel A, “MW 1979” depicts the 1988 wage density under the counterfactual that the minimum wage level remained at its 1979 level. In Panel B, “CF 1979” depicts the counterfactual density for 1988 if the minimum wage and unionization levels remained at their 1979 levels. The shaded areas of the corresponding color depict the range of state minimum wage levels for that year.

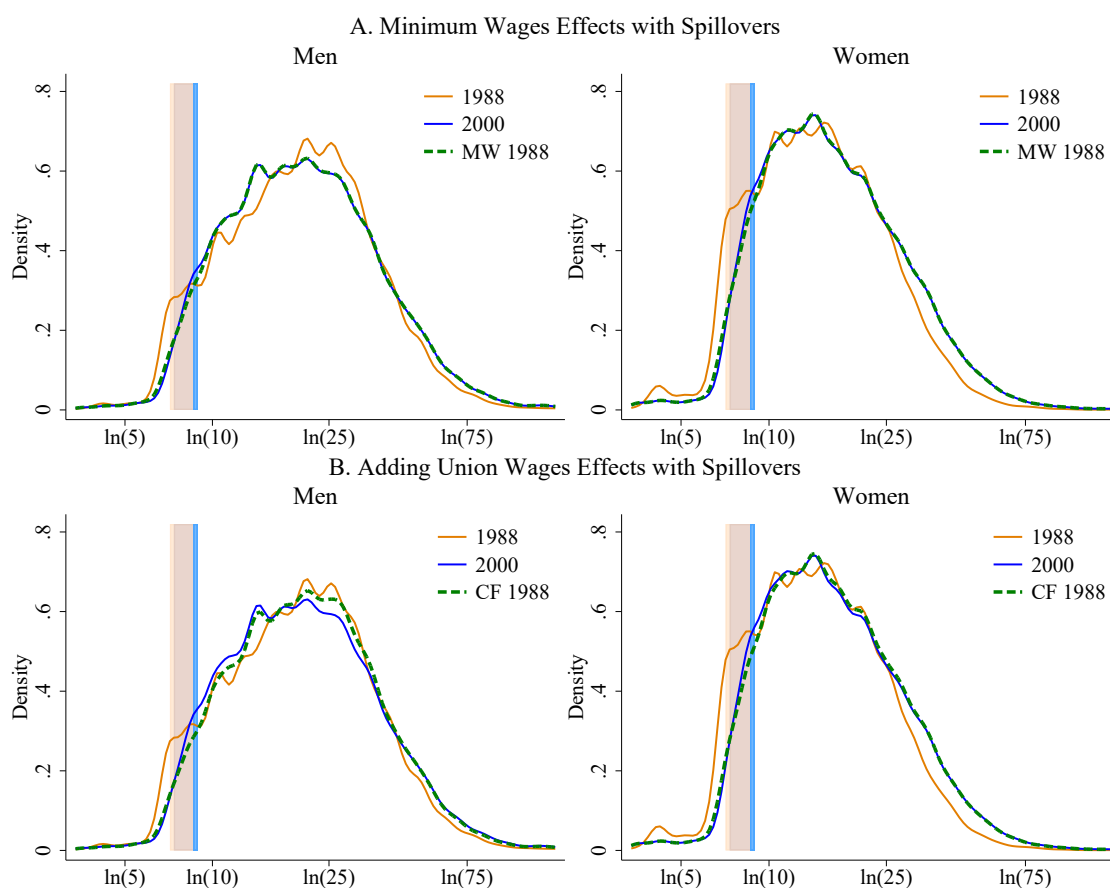


Figure 4.9: Counterfactual densities: 1988-2000

Note: In Panel A, “MW 1988” depicts the 2000 wage density under the counterfactual that the minimum wage level remained at its 1988 level. In Panel B, “CF 1988” depicts the counterfactual density for 2000 assuming the minimum wage and unionization levels remained at their 1988 levels. The shaded areas of the corresponding color depict the range of state minimum wage levels for that year.

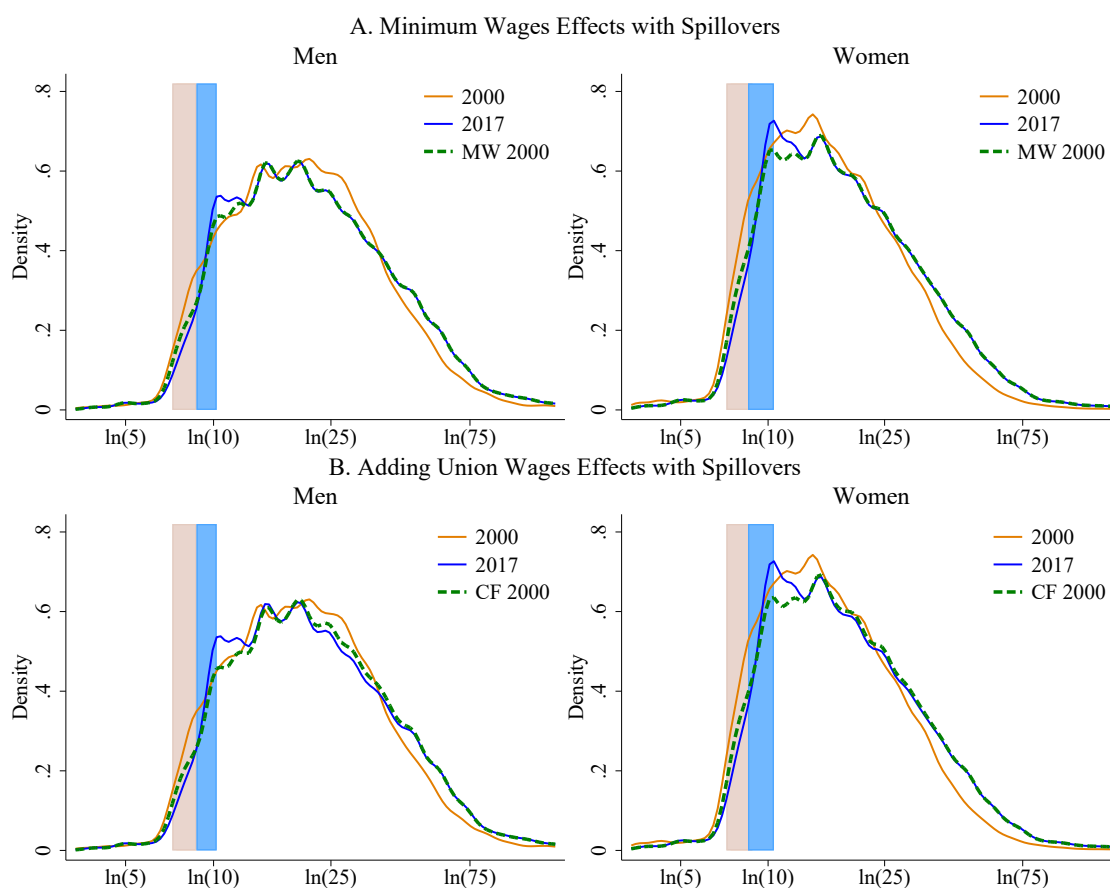


Figure 4.10: Counterfactual densities: 2000-2017

Note: In Panel A, “MW 2000” depicts the 2017 wage density under the counterfactual that the minimum wage level remained at its 2000 level. In Panel B, “CF 2000” depicts the counterfactual density for 2017 if minimum wages and unionization levels remained at their 2000 levels. The shaded areas of the corresponding color depict the range of state minimum wage levels for that year.

Table 4.1: Inequality measures and descriptive statistics

Year	1979	1988	2000	2017
A: Men				
90-10	1.281	1.452	1.521	1.608
90-50	0.588	0.693	0.793	0.901
50-10	0.693	0.759	0.728	0.707
Std(log wages)	0.249	0.326	0.357	0.413
Gini	0.279	0.324	0.355	0.392
Theil	0.142	0.199	0.267	0.326
Unemployment rate	0.051	0.055	0.040	0.045
Unionization rate	0.337	0.229	0.168	0.127
No. of obs.	76213	74020	53037	46342
B: Women				
90-10	0.950	1.286	1.357	1.452
90-50	0.568	0.667	0.746	0.865
50-10	0.382	0.619	0.611	0.588
Std(log wages)	0.172	0.255	0.288	0.357
Gini	0.236	0.287	0.317	0.363
Theil	0.097	0.144	0.204	0.276
Unemployment rate	0.070	0.057	0.042	0.044
Unionization rate	0.176	0.153	0.134	0.115
No. of obs.	62281	69292	52171	45382

Note: 90-10, 90-50, and 50-10 denote corresponding log wage differentials. “No. of obs.” is the number of observations in the unallocated sample used to compute inequality measures. The unemployment and unionization rates are based on the full sample (allocated observations included). For 1979 the unionization rate is derived from the matched May-MORG sample.

Table 4.2: Difference-in-difference estimates of Right-to-Work laws

		By gender		By sector	
	(1) All	(2) Male	(3) Female	(4) Public	(5) Private
A. Union coverage					
Right-to-work	-0.0267*** (0.00865)	-0.0299*** (0.00891)	-0.0232** (0.0101)	-0.0632 (0.0395)	-0.0225*** (0.00609)
Covariates	Yes	Yes	Yes	Yes	Yes
<i>N</i>	4591755	2337504	2254251	784624	3807131
B. Non-union wages					
Right-to-work	-0.0240** (0.00956)	-0.0292** (0.0139)	-0.0191*** (0.00630)	-0.0236** (0.00960)	-0.0244** (0.0105)
Covariates	Yes	Yes	Yes	Yes	Yes
<i>N</i>	2540170	1262956	1277214	301934	2238236

Note: Difference-in-difference estimate of the impact of RTW laws on union coverage and non-union (unallocated) wages. The sample includes all states and is made up of CPS data from 1991-2019. As discussed in Section 4.3.1 Oklahoma, Indiana, Michigan, Wisconsin, West Virginia, and Kentucky adopted RTW laws in 2001, 2011, 2013, 2015 (2010 public sector), 2016, and 2017. Adjustments are made to account for the passing of Bill 10 in Wisconsin, extending RTW to public sector workers. All models control for the standard set of covariates described in Section 4.2.2 and standard errors are clustered at the state level. Only non-allocated wages are used in Panel B.

Table 4.3: Minimum wage effects estimated from distribution regression models

Wage Bins	(1)	(2)	(3)	(4)	(5)	(6)
	1979-1988		1988-2000		2000-2017	
	Women	Men	Women	Men	Women	Men
Minimum wage effects						
More than 10% below	-0.018 (0.004)	-0.025 (0.005)	-0.011 (0.019)	0.011 0.032	-0.030 (0.019)	0.010 (0.020)
5-10% below	-0.016 (0.004)	-0.014 (0.004)	0.000 (0.008)	0.012 (0.013)	-0.006 (0.004)	0.004 (0.005)
0-5% below	0.020 (0.003)	0.022 (0.003)	0.005 (0.006)	0.026 (0.009)	0.014 (0.004)	0.031 (0.004)
At minimum	0.557 (0.017)	0.494 (0.014)	0.341 (0.025)	0.324 (0.029)	0.329 (0.035)	0.293 (0.040)
0-5% above	0.152 (0.003)	0.122 (0.004)	0.095 (0.007)	0.092 (0.005)	0.074 (0.005)	0.059 (0.005)
5-10% above	0.077 (0.002)	0.052 (0.003)	0.003 (0.006)	-0.011 (0.004)	0.053 (0.004)	0.042 (0.003)
10-15% above	0.038 (0.002)	0.033 (0.004)	0.057 (0.007)	0.061 (0.005)	0.024 (0.005)	0.016 (0.004)
15-20% above	0.028 (0.002)	0.031 (0.003)	0.003 (0.010)	-0.011 (0.007)	0.004 (0.013)	-0.010 (0.009)
20-25% above	0.016 (0.002)	0.012 (0.003)				
25-30% above	0.026 (0.002)	0.034 (0.003)				
Spikes at integers						
Dollar values (\$1 to \$10)	0.081 (0.003)	0.083 (0.003)	0.077 (0.004)	0.077 (0.003)	0.063 (0.009)	0.057 (0.008)
Increment at \$5	0.051 (0.004)	0.081 (0.005)				
Increment at \$10			0.004 (0.008)	0.038 (0.005)	0.029 (0.008)	0.049 (0.005)
No. of obs:	698122	787803	715077	741043	946180	955116

Note: The table reports estimates of the models presented in equation (4) in the text. In addition to the variables presented in the table, all models also control for state, year, and quarter effects, state-specific trends, years of education, a quartic in potential experience, experience-education interactions (16 categories plus experience times education), 11 industry categories, 4 occupation categories, and dummy variables for race, marital status, public sector, part-time, and SMSA. Standard errors (clustered at the state level) are in parentheses.

Table 4.4: Estimates of minimum wage impacts along the wage distribution

	Men					Women				
	Fraction affected (FA)	Fraction of affected workers:			Fraction of MW workers moving up (δ_3)	Fraction affected (FA)	Fraction of affected workers:			Fraction of MW workers moving up (δ_3)
		Staying below MW (δ_0)	Moving at MW (δ_1)	Moving above MW (δ_2)			Staying below MW (δ_0)	Moving at MW (δ_1)	Moving above MW (δ_2)	
A. Simple redistribution of the mass of affected workers, no measurement error correction										
1979-1988:	0.071	0.305	0.326	0.369		0.147	0.294	0.337	0.369	
1988-2000:	0.039	0.460	0.316	0.224		0.068	0.456	0.304	0.240	
2000-2017:	0.033	0.479	0.323	0.198		0.051	0.432	0.308	0.259	
B. Redistributing the mass of affected workers and preserving ranks, no measurement error correction										
1979-1988:	0.071	0.305	0.476	0.219	1.000	0.147	0.294	0.458	0.248	1.000
1988-2000:	0.039	0.460	0.534	0.006	1.000	0.068	0.456	0.513	0.031	1.000
2000-2017:	0.033	0.479	0.521	0.000	0.930	0.051	0.432	0.525	0.042	1.000
C. Simple redistribution of the mass of affected workers with partial (50%) measurement error correction										
1979-1988:	0.071	0.207	0.522	0.271		0.147	0.216	0.493	0.291	
1988-2000:	0.039	0.332	0.571	0.097		0.068	0.355	0.506	0.139	
2000-2017:	0.033	0.357	0.565	0.077		0.051	0.340	0.493	0.167	
D. Redistributing the mass of affected workers and preserving ranks with partial (50%) measurement error correction										
1979-1988:	0.071	0.207	0.672	0.121	1.000	0.147	0.216	0.615	0.170	1.000
1988-2000:	0.039	0.332	0.668	0.000	0.443	0.068	0.355	0.645	0.000	0.665
2000-2017:	0.033	0.357	0.643	0.000	0.360	0.051	0.340	0.660	0.000	0.769

Note: The fraction affected, FA, represent the fraction of workers who would have earned less than the minimum wage in absence of a minimum wage. It is computed from the predicted counterfactual distribution that would have prevailed in absence of the minimum wage. The δ parameters indicate the fraction of affected workers staying below the minimum (δ_0), moving to exactly (δ_1) or above (δ_2) the minimum, and workers at the minimum wage who move above the minimum (δ_3). See the text for more details.

Table 4.5: OLS estimates of the union threat effect on log wages with alternative definitions of the unionization rate

Definitions of Unionization Rate:	Non-union Workers				Union Workers			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. 1979-1988								
Industry \times state	0.151 (0.039)	0.152 (0.040)	0.150 (0.040)	0.178 (0.049)	0.337 (0.050)	0.339 (0.051)	0.342 (0.051)	0.326 (0.074)
State		-0.071 (0.085)				-0.149 (0.151)		
Broad industry \times state				-0.049 (0.063)				0.021 (0.102)
B. 1988-2000								
Industry \times state	0.123 (0.035)	0.125 (0.035)	0.125 (0.035)	0.135 (0.044)	0.337 (0.053)	0.341 (0.053)	0.344 (0.054)	0.221 (0.074)
State		-0.152 (0.076)				-0.316 (0.141)		
Broad industry \times state				-0.020 (0.054)				0.203 (0.093)
C. 2000-2017								
Industry \times state	0.110 (0.038)	0.110 (0.038)	0.110 (0.038)	0.069 (0.052)	0.321 (0.053)	0.328 (0.055)	0.332 (0.055)	0.148 (0.093)
State		0.013 (0.066)				-0.380 (0.134)		
Broad industry \times state				0.067 (0.060)				0.257 (0.111)
State \times year dummies	No	No	Yes	No	No	No	Yes	No

Note: The models are estimated for men and women pooled together, and include controls for industry-specific trends and the other variables listed in the note to Table 2. The main industry classification (first row of each panel) consists of 11 industry categories. The broad industry classification (third row of each panel) is based on 3 categories. The number of observations for non-union (union) workers in the three periods are 730778 (197150) in 1979-88, 1201295 (254801) in 1988-2000, and 1636092 (265204) in 2000-2017. Standard errors (clustered at the state-industry level) are in parentheses.

Table 4.6: Unionization rate effects estimated from distribution regression models

Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)
	1979-1988		1988-2000		2000-2017	
	Women	Men	Women	Men	Women	Men
A. Non-Union Workers						
Unionization Rate (UR)	0.060 (0.135)	0.750 (0.118)	0.133 (0.109)	0.724 (0.101)	0.237 (0.084)	0.771 (0.103)
$UR \times y_k$	-0.139 (0.036)	-0.372 (0.038)	-0.293 (0.045)	-0.487 (0.047)	-0.352 (0.057)	-0.498 (0.071)
$UR \times (y_k)^2$	0.002 (0.026)	-0.220 (0.027)	-0.047 (0.027)	-0.190 (0.024)	-0.096 (0.018)	-0.193 (0.016)
$UR \times (y_k)^3$	-0.004 (0.008)	0.023 (0.007)	0.018 (0.009)	0.037 (0.007)	0.044 (0.012)	0.053 (0.009)
$UR \times (y_k)^4$	-0.006 (0.004)	0.008 (0.004)	-0.007 (0.004)	-0.003 (0.003)	-0.006 (0.004)	0.000 (0.003)
No. of obs.:	367769	363009	608892	592403	820847	815245
B. Union Workers						
Unionization Rate (UR)	0.887 (0.217)	1.367 (0.149)	0.870 (0.190)	1.396 (0.128)	0.850 (0.187)	1.399 (0.129)
$UR \times y_k$	-0.374 (0.110)	-0.558 (0.063)	-0.528 (0.098)	-0.487 (0.085)	-0.309 (0.064)	-0.344 (0.057)
$UR \times (y_k)^2$	0.296 (0.053)	-0.072 (0.062)	0.101 (0.039)	-0.043 (0.041)	-0.178 (0.055)	-0.072 (0.051)
$UR \times (y_k)^3$	0.065 (0.022)	0.092 (0.017)	0.098 (0.016)	0.088 (0.019)	0.057 (0.013)	0.063 (0.011)
$UR \times (y_k)^4$	-0.064 (0.007)	-0.029 (0.007)	-0.049 (0.005)	-0.044 (0.005)	-0.005 (0.007)	-0.029 (0.005)
No. of obs.:	73221	123929	106178	148623	125333	139871

Note: The table reports estimates of the models presented in equation (4.5) in the text. The unionization rate is computed at the industry-state-year level. There are 11 industry categories, and industry-specific trends are included in the model in addition to the variables included in Table 2 (minimum wage dummies, integer dummies, and the variables mentioned in the table note). Standard errors (clustered at the state-industry level) are in parentheses.

Table 4.7: Decomposition results: Men

Inequality Measures	(1) Raw Changes	(2) Minimum Wages w/o spill.	(3) w/spill.	(4) Unions w/o spill.	(5) w/spill.	(6) Together w/spill.	(7) Percentage Explained
A: 1979-1988							
90-10	0.213	0.089	0.126	0.027	0.050	0.159	75%
90-50	0.119	0.006	0.004	0.036	0.062	0.065	55%
50-10	0.094	0.083	0.122	-0.009	-0.012	0.094	101%
Std(log wages)	0.073	0.019	0.032	0.009	0.017	0.050	69%
Gini	0.041	0.007	0.011	0.008	0.016	0.026	64%
Theil	0.041	0.007	0.011	0.008	0.017	0.027	66%
B: 1988-2000							
90-10	0.015	0.004	0.004	0.011	0.026	0.031	212%
90-50	0.090	0.001	0.001	0.018	0.037	0.039	43%
50-10	-0.075	0.003	0.002	-0.007	-0.011	-0.008	10%
Std(log wages)	0.013	0.001	0.000	0.004	0.011	0.011	81%
Gini	0.018	0.001	0.000	0.003	0.009	0.009	52%
Theil	0.023	0.001	0.000	0.004	0.011	0.012	49%
C: 2000-2017							
90-10	0.095	-0.007	-0.013	0.004	0.012	-0.001	-1%
90-50	0.121	0.001	0.001	0.011	0.024	0.025	21%
50-10	-0.027	-0.008	-0.014	-0.007	-0.011	-0.026	99%
Std(log wages)	0.032	-0.002	-0.004	0.002	0.005	0.002	5%
Gini	0.020	0.000	-0.001	0.002	0.004	0.004	18%
Theil	0.019	0.000	-0.001	0.002	0.006	0.005	26%
D: 1979-2017							
90-10	0.322	0.086	0.117	0.043	0.088	0.189	59%
90-50	0.330	0.008	0.006	0.065	0.123	0.129	39%
50-10	-0.008	0.078	0.111	-0.023	-0.034	0.060	—
Std(log wages)	0.118	0.018	0.028	0.014	0.034	0.063	53%
Gini	0.079	0.007	0.011	0.013	0.029	0.039	49%
Theil	0.084	0.007	0.010	0.014	0.034	0.044	52%

Note: Column (1) shows the raw changes in inequality measures. Each subsequent column corresponds to a different counterfactual with either minimum wages or unionization turned back to their base period value. Columns (2) and (3) show the contribution of minimum wage changes without (“tail-pasting” only) and with spillover effects. Likewise, columns (4) and (5) show the contribution of changes in unionization without (“shift-share” effect only) and then with spillover effects (threat effects). Column (6) shows the contribution of changes in both the minimum wage and unionization (including spillover effects). Column (7) shows how much of the overall change (column 1) can be explained by institutional change (column 6).

Table 4.8: Decomposition results: Women

Inequality Measures	(1) Raw Changes	(2) Minimum Wages w/o spill.	(3) w/spill.	(4) Unions w/o spill.	(5) w/spill.	(6) Together w/spill.	(7) Percentage Explained
A: 1979-1988							
90-10	0.333	0.141	0.195	0.007	0.007	0.201	60%
90-50	0.087	0.008	0.007	0.011	0.014	0.020	23%
50-10	0.246	0.133	0.188	-0.004	-0.007	0.181	74%
Std(log wages)	0.093	0.017	0.039	0.003	0.004	0.045	48%
Gini	0.050	0.011	0.020	0.002	0.003	0.024	48%
Theil	0.039	0.009	0.016	0.002	0.004	0.020	52%
B: 1988-2000							
90-10	0.045	0.003	-0.003	0.003	0.007	0.004	9%
90-50	0.087	0.002	0.002	0.005	0.010	0.012	14%
50-10	-0.042	0.001	-0.004	-0.002	-0.003	-0.008	20%
Std(log wages)	0.024	0.001	-0.001	0.001	0.003	0.003	11%
Gini	0.021	0.001	0.000	0.001	0.003	0.003	13%
Theil	0.027	0.001	0.000	0.001	0.004	0.004	14%
C: 2000-2017							
90-10	0.110	-0.014	-0.026	0.003	0.007	-0.022	-20%
90-50	0.102	0.001	0.002	0.005	0.011	0.013	13%
50-10	0.008	-0.015	-0.028	-0.002	-0.004	-0.035	—
Std(log wages)	0.047	-0.002	-0.005	0.001	0.003	-0.002	-5%
Gini	0.030	-0.001	-0.002	0.001	0.002	0.001	2%
Theil	0.034	0.000	-0.002	0.001	0.003	0.001	4%
D: 1979-2017							
90-10	0.488	0.130	0.166	0.013	0.021	0.184	38%
90-50	0.276	0.011	0.010	0.021	0.036	0.046	17%
50-10	0.212	0.119	0.156	-0.008	-0.014	0.138	65%
Std(log wages)	0.163	0.017	0.034	0.005	0.011	0.045	28%
Gini	0.102	0.011	0.019	0.004	0.009	0.027	27%
Theil	0.101	0.009	0.015	0.005	0.011	0.026	26%

Note: Column (1) shows the raw changes in inequality measures. Each subsequent column corresponds to a different counterfactual with either minimum wages or unionization turned back to their base period value. Columns (2) and (3) show the contribution of minimum wage changes without (“tail-pasting” only) and with spillover effects. Likewise, columns (4) and (5) show the contribution of changes in unionization without (“shift-share” effect only) and then with spillover effects (threat effects). Column (6) shows the contribution of changes in both the minimum wage and unionization (including spillover effects). Column (7) shows how much of the overall change (column 1) can be explained by institutional change (column 6).

Chapter 5

Conclusion

Both efficiency and equity are central targets in the evaluation and design of economic institutions. For example, designers of present-day tax systems aim to redistribute income, while at the same time minimize distortions to the individual incentive to earn, invest, or consume, as well as mitigate any loss from tax avoidance and evasion. Indeed, the decision to tax households at the individual level - the key institutional detail of Chapter 2 - is motivated by the desire to minimize distortions to secondary earner labour supply, disproportionately represented by female workers. Labour market institutions have a similar underpinning. Both minimum wage and collective bargaining legislation are designed to ensure a more equitable, even just, distribution of economic value for workers, while at the same time considering economic efficiency concerns.

As demonstrated by Chapter 2, these institutions can have unintended consequences. Individual tax structures increase female labour supply but are more burdensome on households with an unequal income distribution. As a result, households have an incentive to find ways to redistribute taxable income. This creates a tax wedge between the self-employed and paid labour markets, where the former benefit from self-reported incomes that can be split with a lower earning spouse. While the literature has identified many such distortions in self-employed households (typically households where the male spouse is self-employed), this is the first paper to show that male spouses become self-employed for this very reason. Moreover, because this distortion depends on the relative earnings of a couple, it is not uniform over the lifecycle. The recent gender inequality literature has emphasized the unequal and long-term burden of parenthood. As a 'shock' to the relative earnings of a couple, the event of childbirth increases this tax wedge between the self-employed and paid labour markets.

By providing causal evidence of the relationship between parental self-employment and the tax code at the point of family formation, the chapter brings together three literatures that have hitherto stood apart: those on tax avoidance in self-employed households, the parent penalty, and gender differences in self-employment. Individual taxation is therefore not neutral; in particular, at the point of family formation it distorts both primary and secondary earner labour supply by incentivizing a joint selection into co-self-employment. The chapter argues that this is not a story of pure tax avoidance. Evidence from survey data

suggests that co-self-employed households are indeed two-worker households in which co-employed mothers benefit from the flexibility of self-employment; e.g., working from home. With childbirth there is a complementarity between the non-pecuniary and tax benefits of self-employment that make selection into co-employment with childbirth even more appealing under an individual tax structure.

The various opportunities for tax avoidance in self-employed households, including income-splitting discussed in Chapter 2, identify an important source of after-tax inequality in the labour market. For self-employed workers at the top of the income distribution - where they are overrepresented - corporate structures can provide an additional source of tax reduction. Limited liability structures exist primarily as a means of reducing entrepreneurial risk; introduced to achieve a more efficient level of risk-taking behaviour. However, their unique tax treatment can also provide a valuable source of tax sheltering for the rich. Chapter 3 is the first paper to provide causal evidence of potential labour-market efficiency gains from incorporation: increased hours worked and increased hiring. A series of legislative reforms in Canada provides a unique opportunity to examine this question in the context of self-employed professionals. In agreement with [Wolfson and Legree \(2015\)](#), the results largely support the claim that incorporation is tax expenditure for the government with little, to no, efficiency gains. Though, for women, incorporation is found to increase the likelihood of hiring staff. A result that is consistent with the after-tax cash flow benefits of incorporation and gender differences in risk-taking behaviour. This chapter once again demonstrates how tax structures can have unintended, but real, labour market consequences.

An ofte- overlooked consequence of minimum wage and collective bargaining institutions is the potential impact they have on non-targetted workers. While unions primarily exist to increase the pecuniary and non-pecuniary benefits of their members, legislation protecting the rights to collective bargaining introduces the potential threat of unionization among uncovered firms. If this leads non-unionized firms to increase benefits, then ignoring this potential threat effect will undercount the impact of collective bargaining on outcomes. [DiNardo et al. \(1996\)](#)'s seminal work on decomposing the impact of labour market institutions on the US wage distribution does not account for potential spillovers from either minimum wages or unions. Chapter 4 extends the analysis of [DiNardo et al. \(1996\)](#) to cover more recent periods and develops a methodology to identify and account for these spillovers. The results demonstrate the overwhelming importance of spillovers from labour market institutions: the minimum wage increases wages above the floor and unionization increases wages in uncovered firms. Chapter 4 finds that accounting for these spillovers significantly increases the contribution of eroding labour market institutions to rising wage inequality in the US.

Bibliography

- Aaronson, D., French, E., Sorkin, I., and To, T. (2018). Industry Dynamics and the Minimum Wage: a Putty-Clay Approach. *International Economic Review*, 59(1):51–84.
- Abraham, K. G. (1988). Flexible Staffing Arrangements and Employers' Short-Term Adjustment Strategies.
- Abraham, K. G., Haltiwanger, J. C., Sandusky, K., and Spletzer, J. R. (2018). Measuring the Gig Economy: Current Knowledge and Open Issues.
- Abraham, S. and Sun, L. (2018). Estimating Dynamic Treatment Effects in Event Studies With Heterogeneous Treatment Effects.
- Acemoglu, D. and Autor, D. (2011). *Skills, tasks and technologies: Implications for employment and earnings*, volume 4. Elsevier B.V.
- Acs, Z. J., Audretsch, D. B., and Evans, D. S. (1994). Why Does the Self-employment Rate Vary Across Countries and Over Time? Technical report, CEPR Discussion Papers.
- Ahn, T. (2010). Attitudes toward risk and self-employment of young workers. *Labour Economics*, 17(2):434–442.
- Allen, W. D. and Curington, W. P. (2014). The self-Employment of men and women: What are their motivations? *Journal of Labor Research*, 35(2):143–161.
- Alstadsæter, B. A., Johannesen, N., and Zucman, G. (2019). Tax evasion and inequality. *American*, 109(6):2073–2103.
- Angelov, N., Johansson, P., and Lindahl, E. (2016). Parenthood and the gender gap in pay. *Journal of Labor Economics*, 34(3):545–579.
- Arnold, T. (1981). Incorporation of professionals in Ohio: Past, present, and future. *Akron Law Review*, 15(2):191–216.
- Aronson, R. L. (1991). *Self-employment: A labor market perspective*. Number 24. ILR Press.
- Astone, N. M., Dariotis, J. K., Sonenstein, F. L., Pleck, J. H., and Hynes, K. (2010). Men's work efforts and the transition to fatherhood. *Journal of Family and Economic Issues*, 31(1):3–13.

- Athey, S. and Imbens, G. (2018). Design-based Analysis in Difference-In-Differences Settings with Staggered Adoption.
- Autor, D. H., Dorn, D., and Hanson, G. H. (2013). The China syndrome: Local labor market effects of import competition in the United States. *American Economic Review*, 103(6):2121–2168.
- Autor, D. H., Manning, A., and Smith, C. L. (2016). The contribution of the minimum wage to US wage inequality over three decades: A reassessment. *American Economic Journal: Applied Economics*, 8(1):58–99.
- Baker, M. and Milligan, K. (2008). *How does job-protected maternity leave affect mothers' employment?*, volume 26.
- Barker, K. and Christensen, K. (1998). *Contingent Work: American Employment Relations in Transition*. Cornell University Press, Ithaca, New York.
- Baron, G. (2013). Selected considerations in the use of professional corporations. *Canadian Tax Journal*, 61(4):1167–1192.
- Bauer, A. M., Macnaughton, A., and Sen, A. (2015). Income splitting and anti-avoidance legislation: evidence from the Canadian "kiddie tax". *International Tax and Public Finance*, 22:909–931.
- Benmelech, E., Bergman, N., and Kim, H. (2018). Strong Employers and Weak Employees: How Does Employer Concentration Affect Wages?
- Bertrand, M. (2011). *New perspectives on gender*, volume 4. Elsevier B.V.
- Bertrand, M., Goldin, C., and Katz, L. F. (2010). Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American Economic Journal: Applied Economics*, 2(3):228–255.
- Bertrand, M., Kamenica, E., and Pan, J. (2015). Gender identity and relative income within households. *The Quarterly Journal of Economics*, 130(2):571–614.
- Biasi, B. and Sarsons, H. (2020). Flexible Wages, Bargaining, and the Gender Gap.
- Binder, A. and Lam, D. (2019). Is There a Male Breadwinner Norm? The Hazards of Inferring Preferences from Marriage Market Outcomes.
- Bjuggren, C. M., Johannson, D., and Stenkula, M. (2010). Using Self-employment as Proxy for Entrepreneurship: Some Empirical Caveats.

- Blanchflower, D. G. (2000). Self-employment in OECD countries. *Labour Economics*, 7(5):471–505.
- Blanchflower, D. G. and Oswald, A. J. (1998). What makes an entrepreneur? *Journal of Labor Economics*, 16(1):26–60.
- Blau, D. M. (1987). A time-series analysis of self-employment in the United States. *Journal of Political Economy*, 95(3):445–467.
- Boden, R. J. (1999). Flexible working hours, family responsibilities, and female self-employment: Gender differences in self-employment selection. *American Journal of Economics and Sociology*, 58(1):71–83.
- Bögenhold, D. and Staber, U. (1991). The decline and rise of self-employment. *Work, Employment and Society*, 5(2):223–239.
- Borusyak, K. and Jaravel, X. (2018). Revisiting Event Study Designs.
- Bound, J. and Krueger, A. B. (1991). The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right? *Journal of Labor Economics*, 9(1):1–24.
- Brochu, P., Green, D. A., Lemieux, T., and Townsend, J. (2018). The Minimum Wage, Turnover, and the Shape of the Wage Distribution. page Working Paper.
- Bruce, D. (1999). Do husbands matter? Married women entering self-employment. *Small Business Economics*, 13(4):317–329.
- Burtch, G., Carnahan, S., and Greenwood, B. N. (2018). Can you gig it? An empirical examination of the gig-economy and entrepreneurship. *Management Science*, pages 1–24.
- Butcher, T., Dickens, R., and Manning, A. (2012). Minimum Wages and Wage Inequality: Some Theory and an Application to the UK.
- Cader, A. and Weinrib, A. (1965). Tax inequities of the professional practitioner in Ontario. *Faculty of Law Review (University of Toronto)*, 23:64–94.
- Cader, H. A. and Leatherman, J. C. (2011). Small business survival and sample selection bias. *Small Business Economics*, 37:155–165.
- Capen, K. (1994). Physicians should weigh advantages, disadvantages of incorporation. *Canadian Medical Association Journal*, 151(10):1482–1483.
- Card, D. (1996). The Effect of Unions on the Structure of Wages: A Longitudinal Analysis. *Econometrica*, 64(4):957.

- Card, D., Cardoso, A. R., Heining, J., and Kline, P. (2018). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics*, 36(S1):S13–S70.
- Card, D. and Krueger, A. B. (1995). *Myth and Measurement: The New Economics of the Minimum Wage*. Princeton University Press.
- Card, D., Lemieux, T., and Craig Riddell, W. (2020). Unions and wage inequality: The roles of gender, skill and public sector employment. *Canadian Journal of Economics*, 53(1):140–173.
- Card, D., Lemieux, T., and Riddell, W. C. (2004). Unions and wage inequality. *Journal of Labor Research*, XXV(4):519–559.
- Carr, D. (1996). Two paths to self-employment? Women’s and men’s self-employment in the United States, 1980. *Work and Occupations*, 23(1):26–53.
- Cengiz, D., Dube, A., Lindner, A., and Zipperer, B. (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics*, 134(3):1405–1454.
- Chaney, T., Sraer, D., and Thesmar, D. (2012). The collateral channel: How real estate shocks affect corporate investment. *American Economic Review*, 102(6):2381–2409.
- Charness, G. and Gneezy, U. (2012). Strong Evidence for Gender Differences in Risk Taking. *Journal of Economic Behavior and Organization*, 83(1):50–58.
- Chernozhukov, V., Fernández-Val, I., and Melly, B. (2013). Inference on counterfactual distributions. *Econometrica*, 81(6):2205–2268.
- Chetverikov, D., Larsen, B., and Palmer, C. (2016). IV quantile regression for group-level treatments, with an application to the distributional effects of trade. *Econometrica*, 84(2):809–833.
- Clain, S. H. (2000). Gender differences in full-time self-employment. *Journal of Economics and Business*, 52:499–513.
- Codazzi, K., Pero, V., and Albuquerque Sant’Anna, A. (2018). Social norms and female labor participation in Brazil. *Review of Development Economics*, 22(4):1513–1535.
- Cooper, G. E. (1978). The Incorporation of Professionals. *Canadian Tax Journal*, 26(1):106–116.
- Cowling, M. and Taylor, M. (2001). Entrepreneurial Women and Men: Two Different Species? *Small Business Economics*, 16(3):167–175.

- Crossley, T., Hurley, J., and Jeon, S.-h. (2009). Physician labour supply in Canada: A cohort analysis. *Health Economics*, 18:437–456.
- Crossley, T. F. and Jeon, S. H. (2007). Joint taxation and the labour supply of married women: Evidence from the Canadian tax reform of 1988. *Fiscal Studies*, 28(3):343–365.
- de Chaisemartin, C. and D’Haultfœuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–2996.
- De Mooij, R. A. and Nicodème, G. (2008). Corporate tax policy and incorporation in the EU. *International Tax and Public Finance*, 15(4):478–498.
- Demirguc-Kunt, A., Love, I., and Maksimovic, V. (2004). Business Environment and the Incorporation Decision.
- Denice, P. and Rosenfeld, J. (2018). Unions and nonunion pay in the United States, 1977-2015. *Sociological Science*, 5:541–561.
- Devereux, M. and Liu, L. (2016). Stimulating Investment through Incorporation.
- DiNardo, J., Fortin, N. M., and Lemieux, T. (1996). Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. *Econometrica*, 64(5):1001.
- Donnelly, M., Magee, J., and Young, A. (2000). Income splitting and the new kiddie tax: Major changes for minor children. *Canadian Tax Journal*, 48(4):979–1018.
- Doumbia, M. and Goussé, M. (2019). Gender Identity and Relative Income within Households: Evidence from Canada.
- Doyle, J. J. J. (2006). Employment Effects of a Minimum Wage: A Density Discontinuity Design Revisited.
- Dube, A., Giuliano, L., and Leonard, J. (2019). Fairness and frictions: The impact of unequal raises on quit behavior. *American Economic Review*, 109(2):620–663.
- Dube, A., Manning, A., and Naidu, S. (2020). Monopsony and Employer Mis-optimization Explain Why Wages Bunch at Round Numbers.
- Duff, D. G. (1999). Neuman and beyond: Income splitting, tax avoidance, and statutory interpretation in the Supreme Court of Canada. *The Canadian Business Law Journal*, 32(3):345–383.
- Dustmann, C., Lindner, A., Schönberg, U., Umkehrer, M., and vom Berge, P. (2020). Reallocation Effects of the Minimum Wage.

- Easterbrook, F. H. and Fischel, D. R. (1985). Limited Liability and the Corporation. *The University of Chicago Law Review*, 52(1):89–117.
- Eckel, C. C. and Grossman, P. J. (2008). Chapter 113 Men, Women and Risk Aversion: Experimental Evidence. *Handbook of Experimental Economics Results*, 1(C):1061–1073.
- Egger, P., Keuschnigg, C., and Winner, H. (2009). Cesifo Working Paper no. 2921.
- Eissa, N. (1995). Taxation and Labor Supply of Married Women: The Tax Reform Act of 1986 as a Natural Experiment.
- Ellwood, D. T. and Fine, G. (1987). The impact of Right-to-Work laws on union organizing. *Journal of Political Economy*, 95(2):250–273.
- Faccio, M. and Xu, J. (2015). Taxes and Capital Structure. *Journal of Financial and Quantitative Analysis*, 50(3):277–300.
- Faggio, G. and Silva, O. (2012). Does Self-Employment Measure Entrepreneurship? Evidence from Great Britain.
- Fama, E. F. and Jensen, M. C. (1983). Separation of ownership and control. *Journal of Law & Economics*, XXVI(June):301–325.
- Farber, H., Herbst, D., Kuziemko, I., and Naidu, S. (2018). Unions and Inequality Over the Twentieth Century: New Evidence from Survey Data.
- Farber, H. S. (2005). Nonunion wage rates and the threat of unionization. *Industrial and Labor Relations Review*, 58(3):335–352.
- Feigenbaum, J., Hertel-Fernandez, A., and Williamson, V. (2018). From the Bargaining Table To the Ballot Box: Political Effects of Right to Work Laws.
- Firpo, S., Fortin, N. M., and Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, 77(3):953–973.
- Firpo, S. P., Fortin, N. M., and Lemieux, T. (2018). Decomposing wage distributions using recentered influence function regressions. *Econometrics*, 6(2):1–40.
- Foresi, S. and Peracchi, F. (1995). The conditional distribution of excess returns: An empirical analysis. *Journal of the American Statistical Association*, 90(430):451–466.
- Fortin, N. M. and Lemieux, T. (1998). Rank regressions, wage distributions, and the gender gap. *Journal of Human Resources*, 33(3):610–643.

- Fortin, N. M., Lemieux, T., and Lloyd, N. (2019). Labor Market Institutions and the Distribution of Wages: The Role of Spillover Effects.
- Fortin, N. M., Lemieux, T., and Lloyd, N. (2021). Labor market institutions and the distribution of wages: The role of spillover effects. *Journal of Labor Economics*, 39(S2):S369–S412.
- Freedman, J. (2000). Limited Liability: Large Company Theory and Small Firms. *Modern Law Review*, 63(3):317–354.
- Freeman, R. B. (1993). How much has de-unionization contributed to the rise in male earnings inequality? In Danziger, S. and Gottschalk, P., editors, *Uneven Tides: Rising Inequality in America*, pages 133–163. Russell Sage Foundation, New York.
- Freeman, R. B. (1996). Labor market institutions and earnings inequality. *New England Economic Review*, (May / June):157–172.
- Freeman, R. B. and Medoff, J. L. (1981). The impact of the percentage organized on union and nonunion wages. *The Review of Economic and Statistics*, 63(4):561–572.
- Fuenmayor, A., Granell, R., and Mediavilla, M. (2018). The effects of separate taxation on labor participation of married couples. An empirical analysis using propensity score. *Review of Economics of the Household*, 16(2):541–561.
- Fuller, S. and Hirsh, C. E. (2019). “Family-friendly” jobs and motherhood pay penalties: The impact of flexible work arrangements across the educational spectrum, volume 46.
- Georgellis, Y. and Wall, H. J. (2005). Gender differences in self-employment. *International Review of Applied Economics*, 19(3):321–342.
- Glauber, R. (2008). Race and gender in families and at work: The fatherhood wage premium. *Gender and Society*, 22(1):8–30.
- Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, 104(4):1091–1119.
- Goldin, C. and Katz, L. F. (2016). A most egalitarian profession: Pharmacy and the evolution of a family-friendly occupation. *Journal of Labor Economics*, 34(3):705–746.
- Goodman-Bacon, A. (2018). Difference-in-differences with Variation in Treatment Timing.
- Goolsbee, A. (1998). Taxes, organizational form, and the deadweight loss of the corporate income tax. *Journal of Public Economics*, 69(1):143–152.

- Gordon, R. H. and MacKie-Mason, J. K. (1994). Tax distortions to the choice of organizational form. *Journal of Public Economics*, 55(2):279–306.
- Gordon, R. H. and Slemrod, J. (1998). Are "Real" Responses to Taxes Simply Income Shifting between Corporate and Personal Tax Bases?
- Graschuk, H. S. (1977). The professional corporation in Alberta. *Canadian Tax Journal*, 25(2):109–118.
- Grossman, J. B. (1983). The impact of the minimum wage on other wages. *The Journal of Human Resources*, 18(3):359–378.
- Gruber, J. and Cullen, J. B. (1996). Spousal Labor Supply as Insurance: Does Unemployment Insurance Crowd Out the Added Worker Effect?
- Gruber, J. and Saez, E. (2002). The elasticity of taxable income: Evidence and implications. *Journal of Public Economics*, 84:1–32.
- Haanwinckel, D. (2020). Supply, Demand, Institutions, and Firms: A Theory of Labor Market Sorting and the Wage Distribution.
- Hakim, C. (1988). Self-employment in Britain: Recent trends and current issues. *Work, Employment and Society*, 2(4):421–450.
- Hamilton, R. W. (1994). Registered Limited Liability Partnerships: Present at the Birth (Nearly). *University of Colorado Law Review*, 66(4):1065–1104.
- Hanglberger, D. and Merz, J. (2015). Does self-employment really raise job satisfaction? Adaptation and anticipation effects on self-employment and general job changes. *Journal for Labour Market Research*, 48(4):287–303.
- Harju, J. and Matikka, T. (2016). Business owners and income-shifting: evidence from Finland. *Small Business Economics*, 46(1):115–136.
- Headd, B. (2003). Redefining business success: Distinguishing between closure and failure. *Small*, 21:51–61.
- Hederos, K. and Stenberg, A. (2019). Gender Identity and Relative Income within Households - Evidence from Sweden.
- Hipple, S. (2004). Self-employment in the United States: An update. *Monthly Labor Review*, 127(7):13–23.

- Hipple, S. (2010). Self-employment in the United States. *Monthly Labor Review*, (September):17–32.
- Hirsch, B. T. and Schumacher, E. J. (2004). Match bias in wage gap estimates due to earnings imputation. *Journal of Labor Economics*, 22(3):689–722.
- Hodges, M. J. and Budig, M. J. (2010). Who gets the daddy bonus?: Organizational hegemonic masculinity and the impact of fatherhood on earnings. *Gender and Society*, 24(6):717–745.
- Hotz, V. J., Johansson, P., and Karimi, A. (2018). Parenthood, family friendly workplaces, and the gender haps in early work careers.
- Humphries, J. E. (2019). The Causes and Consequences of Self-Employment over the Life Cycle.
- Hundley, G. (2000). Male/female earnings differences in self-employment: The effects of marriage, children, and the household division of labor. *Industrial and Labor Relations Review*, 54(1):95–114.
- Hurst, E., Li, G., and Pugsley, B. (2014). Are household surveys like tax forms? Evidence from income underreporting of the self-employed. *Review of Economics and Statistics*, 96(1):19–33.
- Hurst, E. and Pugsley, B. (2010). The Non Pecuniary Benefits of Small Business Ownership.
- Jales, H. (2018). Estimating the effects of the minimum wage in a developing country: A density discontinuity design approach. *Journal of Applied Econometrics*, 33(1):29–51.
- Jeon, S.-H., Liu, H., and Ostrovsky, Y. (2019). Measuring the Gig Economy in Canada Using Administrative Data.
- Jeon, S. H. and Ostrovsky, Y. (2019). Balancing family and work: Transition to self-employment among new mothers. *Oxford Economic Papers*, 71(1):47–72.
- Joona, P. A. (2017). Are mothers of young children more likely to be self-employed? The case of Sweden. *Review of Economics of the Household*, 15(1):307–333.
- Kabátek, J., van Soest, A., and Stancanelli, E. (2014). Income taxation, labour supply and housework: A discrete choice model for French couples. *Labour Economics*, 27:30–43.
- Karoly, L. A. and Zissimopoulos, J. (2004). Self-employment among older U.S. workers. *Monthly Labor Review*, 127(7):24–47.

- Katz, L. F. and Krueger, A. B. (2016). The Rise and Nature of Alternative Work Arrangements in the United States, 1995-2015.
- Katz, L. F. and Krueger, A. B. (2019). The rise and nature of alternative work arrangements in the United States, 1995–2015. *ILR Review*, 72(2):382–416.
- Kleven, H., Landais, C., Posch, J., Steinhauer, A., and Zweimüller, J. (2019a). Child Penalties Across Countries: Evidence and Explanations.
- Kleven, H., Landais, C., and Søgaaard, J. E. (2019b). Children and gender inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, 11(4):181–209.
- Kleven, H. J., Kreiner, C. T., and Saez, E. (2009). The optimal income taxation of couples. *Econometrica*, 77(2):537–560.
- Kleven, H. J. and Waseem, M. (2013). Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan. *The Quarterly Journal of Economics*, 128(December):669–723.
- Kuhn, K. M. (2016). Rise of the "gig-economy" and implications for understanding work and workers. *Industrial and Organizational Psychology*, 30:157–162.
- Kuhn, P. J. and Schuetze, H. J. (2001). Self-employment dynamics and self-employment trends: a study of Canadian men and women, 1982-1998. *The Canadian Journal of Economics*, 34(3):760–784.
- LaLumia, S. (2008). The effects of joint taxation of married couples on labor supply and non-wage income. *Journal of Public Economics*, 92(7):1698–1719.
- Lamadon, T., Mogstad, M., and Setzler, B. (2020). Imperfect Competition, Compensating Differentials and Rent Sharing in the U.S. Labor Market.
- le Maire, D. and Schjerning, B. (2013). Tax bunching, income shifting and self-employment. *Journal of Public Economics*, 107:1–18.
- Lee, D. S. (1999). Wage inequality in the United States during the 1980s: Rising dispersion or falling minimum wage? *The Quarterly Journal of Economics*, 114(3):977–1023.
- Lemieux, T. (2006). Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill? *American Economic Review*, 96(3):461–498.
- Lewis, H. G. (1963). *Unionism and Relative Wages in the United States: An Empirical Inquiry*. University of Chicago Press, Chicago.

- Lin, Z., Picot, G., and Compton, J. (2000). The entry and exit dynamics of self-employment in Canada. *Small Business Economics*, 15(2):105–125.
- Lippmann, Q., Georgieff, A., and Senik, C. (2020). Undoing gender with institutions: Lessons from the German division and reunification. *The Economic Journal*, 130(May):1445–1470.
- Liu, L. (2014). Income taxation and business incorporation: Evidence from the early twentieth century. *National Tax Journal*, 67(2):387–418.
- Lombard, K. V. (2001). Female self-employment and demand for flexible, nonstandard work schedules. *Economic Inquiry*, 39(2):214–237.
- Lundberg, S. and Rose, E. (2000). Parenthood and the earnings of married men and women. *Labour Economics*, 7(6):689–710.
- Lundberg, S. and Rose, E. (2002). The effects of sons and daughters on men’s labor supply and wages. *Review of Economics and Statistics*, 84(2):251–268.
- MacIntyre, J. (1972). The British Columbia Professional Corporations Act. *The University of Toronto Law Journal*, 22(3):207–209.
- MacKie-Mason, J. K. and Gordon, R. H. (1997). How much do taxes discourage incorporation? *Journal of Finance*, 52(2):477–506.
- Macnaughton, A. and Matthews, T. (1999). Is the income-splitting tax needed? Some empirical evidence. *Canadian Tax Journal*, 1164(5).
- Manning, A. (2003). *Monopsony in Motion*. Princeton University Press.
- Manser, M. E. and Picot, G. (1999a). Self-employment in Canada and the United States.
- Manser, M. E. and Picot, G. (1999b). The role of self-employment in U.S. and Canadian job growth. *Monthly Labor Review*, 122(4):10–23.
- Milligan, K. (2016). Canadian Tax and Credit Simulator.
- Moore, W. J., Newman, R. J., and Cunningham, J. (1985). The effect of the extent of unionism on union and nonunion wages. *Journal of Labor Research*, 6(1):21–44.
- Noseleit, F. (2014). Female self-employment and children. *Small Business Economics*, 43(3):549–569.
- Nunn, R., O’Donnell, J., and Shambaugh, J. (2019). The Shift in Private Sector Union Participation: Explanation and Effects. Technical Report August.

- Oxenfeldt, A. R. (1943). New Firms and Free Enterprise: Pre-War and Post-War Aspects. Technical report, American Council on Public Affairs.
- Phillips, J. D. (1962). *The self-employed in the United States*. UNiversity of Illinois Press, Urbana, IL.
- Podgursky, M. (1986). Unions, establishment size, and intra-industry threat effects. *ILR Review*, 39(2):277–284.
- Rauh, J. D. (2006). Investment and Financing Constraints : Evidence. *The Journal of Finance*, LXI(1):33–71.
- Ribstein, L. E. (1995). The emergence of the limited liability company. *Business Lawyer*, 51(1):1–49.
- Rodney, M. (1969). Professional Corporations. *Alberta Law Review*, 7(2):205–217.
- Rosen, S. (1969). Trade union power, threat effects and the extent of organization. *Review of Economic Studies*, 36(2):185–196.
- Rosenfeld, J., Denice, P., and Laird, J. (2016). Union Decline Lowers Wages of Nonunion Workers. Technical report, Economic Policy Institute.
- Saez, E. (2010). Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy*, 2(3):180–212.
- Saridakis, G., Marlow, S., and Storey, D. J. (2014). Do different factors explain male and female self-employment rates? *Journal of Business Venturing*, 29:345–362.
- Scheuer, F. and Slemrod, J. (2020). Taxation and the superrich. *Annual Review of Economics*, 12(1):189–211.
- Schuetze, H. J. (2002). Profiles of tax non-compliance among the self-employed in Canada: 1969 to 1996. *Canadian Public Policy*, 28(2):219–238.
- Schuetze, H. J. (2006). Income splitting among the self-employed. *The Canadian Journal of Economics*, 39(4):1195–1220.
- Schumpeter, J. A. (1921). *The Theory of Economic Development*. Harvard University Press, Cambridge, MA.
- Schumpeter, J. A. (1942). *Capitalism, socialism and democracy*.

- Selin, H. (2013). The rise in female employment and the role of tax incentives. An empirical analysis of the Swedish individual tax reform of 1971. *International Tax and Public Finance*, 21(5):894–922.
- Slemrod, J. (2007). Cheating Ourselves: The Economics of Tax Evasion Tax Evasion. *Journal of Economic Perspectives*, 21(1):25–48.
- Slemrod, J. (2019). Tax compliance and enforcement. *Journal of Economic Literature*, 57(4):904–954.
- Smeltzer, G. (1970). An examination of the Professional Corporations Act in light of the white paper on tax reform. *University of British Columbia Law Review*, 5(2):339.
- Smith, M., Yagan, D., Zidar, O., and Zwick, E. (2019). Capitalists in the twenty-first century. *The Quarterly Journal of Economics*, 134(4):1675–1745.
- Stanford, J. (2017). The resurgence of gig work: Historical and theoretical perspectives. *The Economic and Labour Relations Review*, 28(3):382–401.
- Stephens, M. and Ward-Batts, J. (2004). The impact of separate taxation on the intra-household allocation of assets: Evidence from the UK. *Journal of Public Economics*, 88(9–10):1989–2007.
- Storey, D. J. (1991). The birth of new firms: Does unemployment matter? A review of the evidence the birth of new firms. *Small Business Economics*, 3(3):167–178.
- Taschereau-Dumouchel, M. (2020). The union threat. *Review of Economic Studies*, 87(6):2859–2892.
- Teulings, C. N. (2000). Aggregation bias in elasticities of substitution and the minimum wage paradox. *International Economic Review*, 41(2):359–398.
- van der Zwan, P., Hessels, J., and Rietveld, C. A. (2018). Self-employment and satisfaction with life, work, and leisure. *Journal of Economic Psychology*, 64:73–88.
- van Praag, C. M. (2003). Business survival and success of young small business owners. *Small Business Economics*, 21:1–17.
- Wieber, A. and Holst, E. (2015). Gender Identity and Womens’ Supply of Labor and Non-Market Work: Panel Data Evidence for Germany.
- Wilde, E. T., Batchelder, L., and Ellwood, D. T. (2010). The Mommy Track Divides: The Impact of Childbearing on Wages of Women of Different Skill Levels.

- Wolfson, M. and Legree, S. (2015). Policy forum: Private companies, professionals, and income splitting - Recent Canadian experience. *Canadian Tax Journal*, 63(3):717–737.
- Wolfson, M., Veall, M., Brooks, N., and Murphy, B. (2016). Piercing the veil: Private corporations and the income of the affluent. *Canadian Tax Journal*, 64(1):1–30.
- Zhang, X. (2010). Can motherhood earnings losses be ever regained? Evidence from Canada. *Journal of Family Issues*, 31(12):1671–1688.
- Zinovyeva, N. and Tverdostup, M. (2018). Gender Identity, Co-Working Spouses and Relative Income within Households.
- Zissimopoulos, J. M. and Karoly, L. A. (2007). Transitions to self-employment at older ages: The role of wealth, health, health insurance and other factors. *Labour Economics*, 14(2):269–295.
- Zwick, E. and Mahon, J. (2017). Tax policy and heterogeneous investment behavior. *American Economic Review*, 107(1):217–248.

A Chapter 2 Appendix

A.1 Finite horizon, unitary household model

This section describes a unitary model of family labour supply incorporating the discrete decision of employment sector (paid versus self-employment) along with the nested decision of hours labour supply. In the wage-paying sector contracts are limited to full and part-time, while self-employed workers can make a continuous choice over hours. Earnings in both sectors are pre-determined and known to the family.

The model is gender neutral, providing symmetric labour market options to both parents. Gender asymmetry enters through childbirth with one parent (i.e., the mother) forced to be economically inactive for a single period. The timing of family formation is predetermined and known to the couple, for both the first and second child. After childbirth working parents must solve the problem of childcare, requiring parents to either purchase childcare at a fixed rate, or reduce their labour supply.

Joint self-employment is assumed to be co-employment. As such, the start-up cost to self-employment is shared. These start-up costs are also transferable, allowing spouses to switch ‘management’ roles within the household firm. This sharing of the start-up costs increases co-employment in equilibrium, even without childbirth. Co-employment also provides a positive childcare externality: co-employed parents face a lower cost to childcare arising through flexibility. Note, individually self-employed parents do not benefit from this same lower cost to childcare.¹⁷⁷

Self-employed households also benefit from income-splitting. I assume that income-splitting is legal (without risk of retribution), but only self-employed earnings can be split. For this reason, income-splitting only benefits households where the self-employed worker is also the primary earner. When solving the model, I keep track of both real labour supply and tax reported earnings. In this way I can map the difference between real self-employment and tax-reported self-employment.

Set-up

Households maximize their combined, finite-lifetime utility through a sequence of nested labour supply decisions. In this way households are solving for the optimal path of labour supply, given their known fertility path.

¹⁷⁷While this assumption is inconsistent with the literature that emphasises work-life balance benefits of self-employment, it is consistent with the negative compensating wage-variation such mothers face. I provide evidence in Appendix A.6 that with the exception of co-employed mothers, self-employed mothers face a large, negative wage premium. This likely represents the compensating wage-variation associated with the increase in flexibility. Given that co-employed mothers do not face such a premium, while enjoying similar flexibility benefits, I am justified in limiting the positive externality to co-employed households.

$$V(s_t^1, s_t^2, h_t^1, h_t^2 | s_{t-1}^1, s_{t-1}^2, h_{t-1}^1, h_{t-1}^2) = \underset{s_t^1, s_t^2}{argmax} U(h_t^1, h_t^2 | s_t^1, s_t^2, s_{t-1}^1, s_{t-1}^2, h_{t-1}^1, h_{t-1}^2) - K(s_t^1, s_t^2, s_{t-1}^1, s_{t-1}^2) \\ - \Delta(s_t^1, s_t^2, h_t^1, h_t^2) + \beta V(s_{t+1}^1, s_{t+1}^2, h_{t+1}^1, h_{t+1}^2 | s_t^1, s_t^2, h_t^1, h_t^2)$$

This is a nested value function, whereby internal value of $U(\cdot | \cdot)$ represents the solution to the intensive margin choice of hours of labour supply conditional on the choice of sector and past labour supply decisions. This function represents the value of after-tax earnings less any cost of childcare. To simplify notation, I assume that the fixed start-up costs (K) and income splitting tax savings (Δ) are not included in the nested, intensive margin decision of hours of work. There are therefore no distortions to hours of labour supply resulting from income splitting. Income-splitting generates a pure income-effect that acts as a wedge between the two sectors.

Δ is defined as the difference between the after-tax income of the household under a joint and individual tax structure. In practice, I solve for the minimum transfer required to equate the household members' marginal tax rates, which given a stepped tax structure need not result in perfect income splitting. Only self-employment income can be split, with the result that there are no gains to income splitting in households where the wage-employee earns more than the self-employed parent.

$$U(h_t^1, h_t^2 | s_t^1, s_t^2, s_{t-1}^1, s_{t-1}^2, h_{t-1}^1, h_{t-1}^2) = \underset{h_t^1, h_t^2}{argmax} \left[y^1(h_t^1, s_t^1) - T(y^1(h_t^1, s_t^1)) + y^2(h_t^2, s_t^2, s_{t-1}^2) \right. \\ \left. - T(y^2(h_t^2, s_t^2, s_{t-1}^2)) - N_t(h_t^1, h_t^2, s_t^1, s_t^2) \right] \\ \text{s.t.} \\ y^1(h_t^1, s_t^1) = v_t^1 h_t^1 s_t^1 + w_t^1 h_t^1 (1 - s_t^1) \\ y^2(h_t^2, s_t^2, s_{t-1}^2) = (1 - \iota_t) \cdot (v_t^2 h_t^2 s_t^2 + w_t^2 h_t^2 (1 - s_t^2)) + \iota_t \cdot (\psi \cdot w_{t-1}^2 h_{t-1}^2 (1 - s_{t-1}^2)) \\ \iota_t = \mathbf{1}\{t = \tau_1 \text{ or } t = \tau_2\} \\ h_t^g = \begin{cases} \in (0, 1] & \text{if } s_t^g = 1 \\ \in \{0, 0.5, 1\} & \text{if } s_t^g = 0 \end{cases} \quad g = 1, 2 \\ N_t(h_t^1, h_t^2, s_t^1, s_t^2) = v_t (1 - \rho s_t^1 s_t^2) \cdot \min\{h_t^1, h_t^2\}$$

Conditional on sector of employment, each parent must choose their hours of labour supply. The maximum number of hours are assumed to be equivalent to a full-time contract

and normalized to 1. Self-employed workers can choose any level of hours along the continuum $(0, 1]$, while paid employees must select between a part-time and full-time contract. Non-employed workers are assumed to be in the wage-paying sector, thereby avoiding nominal self-employment. Hours worked in the paid labour market are therefore chosen from the set $\{0, 0.5, 1\}$. Wage rates in each sector are known, denoted here by w_t^g in the wage-paying sector and v_t^g in the self-employed labour market.

Asymmetry enters the model through the event of childbirth (ι_t) which takes place in two predetermined periods. Each couple has two children, and the time between births is known (periods τ_1 and τ_2). The birthing parent (i.e. 'mother': $g = 2$) leaves the labour market for a single period. The mother's earnings function is state dependent. During the period of childbirth, the mother is unable to work, and receives an income supplement (i.e. employment insurance) based on her previous period's earnings. However, only wage earnings are insurable. This assumption is in align with the employment insurance policy that governed Canada for most of the past four decades.¹⁷⁸ The parameter ψ represents the income replacement rate. For example, Canada's federal employment insurance policy has a standard replacement rate of 55% up to a threshold. Finally, the household's budget constraint contains a Leontieff childcare cost function, scaled by the time-varying parameter ν_t . In co-employed households, these costs are reduced by the factor ρ which denotes the childcare externality of co-employment.

If not for the event of childbirth this model would yield an allocation of workers between wage and self-employment based entirely on relative earnings in both sectors (in the absence of income-splitting). In addition, if the average earnings of 'mothers' and 'fathers' are equal in both sectors there would be no 'gender' difference in the rates of self-employment and paid employment. It is the event of childbirth and the subsequent cost of childcare that interacts with other aspects of the model to generate staggered selection into self-employment along 'gendered' lines.

Co-employment after childbirth has obvious childcare benefits. However, selection into self-employment after childbirth is more costly, as there is less time to overcome the fixed start-up costs. For this reason, selection into self-employment will take place early in the life cycle. If mothers enter self-employment before childbirth they lose out on employment insurance benefits when they give birth. Thus, fathers will select into self-employment first, even if the mother has higher self-employment earnings. They can always switch over after the child is born.

The decline in maternal earnings during the childbirth creates an additional incentive for fathers to enter self-employment before childbirth through income splitting. A self-

¹⁷⁸Quebec included the self-employed in their QPIP reform of 2006 and since 2011 federal EI contributions have been optional for the self-employed. These can be claimed for maternity leave; although business revenue must decline by at least 40% to do so.

employed father is able to lower the household's tax burden by shifting self-employment income to the economically inactive mother. This will create a wedge between real and tax-reported self-employment. However, these savings are likely too small to induce a mother to abandon a wage paying position before childbirth, as it is not just one period of earnings she would forego. Thus, future savings from income splitting act as a subsidy to paternal self-employment leading up to childbirth. This, together with the employment insurance structure, means that fathers will lead the household's selection into self-employment. In addition, the childcare savings from co-employment will drive a wedge between the relative earnings in each sector and increase paternal self-employment in all periods and maternal self-employment after childbirth.

Simulation

The optimal labour supply path of each household is solved through backwards induction. This is done for a sample of households with incomes drawn from the distributions and policy parameters described below. Households live for ten periods ($t = -3, -2, \dots, 6$), with the event of initial childbirth occurring in period 0. To simplify matters I denote time in event-time of childbirth.

- Sample of $N=1500$ households. A third of households have their second child in periods 2, 3, and 4 respectively.
- Wage earnings are assuming to constant through time and uncorrelated within the household. They are drawn from the following distribution:

$$\begin{bmatrix} w^1 \\ w^2 \end{bmatrix} \sim N \left(\begin{bmatrix} 10 \\ 8 \end{bmatrix}, \begin{bmatrix} 1.5, 0 \\ 0, 1.5 \end{bmatrix} \right)$$

- Self-employment earnings are dependent on wage earnings in the following way,

$$\begin{bmatrix} v^1 \\ v^2 \end{bmatrix} = \begin{bmatrix} w^1 \\ w^2 \end{bmatrix} + \begin{bmatrix} \eta^1 \\ \eta^2 \end{bmatrix} \quad \text{with} \quad \begin{bmatrix} \eta^1 \\ \eta^2 \end{bmatrix} \sim N \left(\begin{bmatrix} -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0.75, 0 \\ 0, 0.75 \end{bmatrix} \right)$$

- Utility is linear in consumption ($U_t = c_t$) with discounting $\beta = .9$

- Cost functions:

$$N_t(h_t^1, h_t^2, s_t^1, s_t^2) = \nu_t(1 - \rho s_t^1 s_t^2) \cdot \min\{h_t^1, h_t^2\}$$

$$\text{with } \nu_t = \begin{cases} 5 - 1.25(t - \tau_1 - 1) & t \in [\tau_1 + 1, \tau_1 + 4] \\ 5 - 1.25(t - \tau_2 - 1) & t \in [\tau_2 + 1, \tau_2 + 4] \end{cases}$$

and $\rho = 0.3$

$$K(s_t^1, s_t^2, s_{t-1}^1, s_{t-1}^2) = \kappa \cdot (s_t^1 + s_t^2 - s_t^1 s_t^2) \cdot (1 - s_{t-1}^1)(1 - s_{t-1}^2)$$

with $\kappa = 5$

- Tax structure. The marginal tax rate on the first 1.5 units of income is 0, followed by 20% on the next 4.5 units, 30% on the next 4 units, and 40% above 10 units.
- EI replacement rate of $\psi = 0.55$ with no income thresholds.

The EI simulated structure is a simplification of Canada's real EI structure which includes a Working While on Claim (WWC) component. Under WWC the benefits are gradually clawed back if someone has other employment income. In periods 1 and 2 employment is again optional, but the household must pay for childcare if both parents are employed. If the parents are co-employed the cost of childcare is halved. In period 2 the cost of childcare falls to simulate the decline in childcare costs as children reach school age.

Wage earnings are drawn from a normal distribution and held constant over time. The average mother in the sample is assumed to earn 80 cents on the dollar in the wage-paying sector. These earnings are uncorrelated within a couple. For each individual self-employment earnings are equal to wage earnings plus a 'entrepreneurial' shock. This shock is normally distributed with a negative mean to ensure that on average self-employment rate is lower than 50%, as is typically the case in the population. The 'entrepreneurial' shock is also assumed to be uncorrelated within the household. When solving the model, I assume a linear utility function for consumption and no dis-utility of labour.

Solution

Figure A.1 plots the optimal path of *real* self-employment through the event of childbirth. In each panel the right-hand side figure provides a counterfactual in which there is no childbirth. The two curves in each figure demonstrates the increase in *real* self-employment associated with income splitting. Comparing this difference between the figures with and without childbirth, suggests the relative importance of income splitting during the event of

childbirth relative to other periods of the lifecycle. Figure A.2 plots the same curves for *real* co-employment

Panel A of Figure A.1 replicates the discontinuous increase in maternal self-employment after childbirth observed in Figure 2.1. One of the main mechanisms here is the employment insurance structure which incentivizes mothers to remain paid employees leading up to childbirth. For this reason, the level of female self-employment before childbirth is much lower. In contrast, male self-employment is higher before childbirth (Panel B). This is because the start-up costs can be shared, so in households where parents wish to be co-employed after childbirth for childcare reasons, it is best for the father to enter first. Thus, male self-employment increases leading up to childbirth, as in Figure 2.1.

With income-splitting there are now additional benefits to self-employment during the year of childbirth when maternal employment income is lost. This increases paternal self-employment even before childbirth. Moreover, the increase in paternal self-employment is greater with childbirth relative to the counterfactual without childbirth, highlighting the fact that it is the interaction this asymmetric event and the tax structure that leads to such high levels of self-employment and co-employment after childbirth.

In Panel B of Figure A.1 the increase in paternal self-employment under the counterfactual without children arises from the difference between the average income of men and women in the sample. As such, there are income-splitting benefits to self-employment in the average household even before childbirth. For this same reason, there is no increase in real female self-employment under the counterfactual without childbirth.

With childbirth, the increase in real maternal self-employment from income-splitting is dampened by the incentive to return to the paid sector before the arrival of a second child. However, by event-time 6 maternal self-employment (and co-employment) increases by the same margin as paternal self-employment (relative to the counterfactual without children).¹⁷⁹ This highlights a second important point which is that tax-induced paternal self-employment can give rise to a *real* maternal co-employment *within* the context of family formation.

¹⁷⁹The share of women with some reported self-employment income is much higher and increases with childbirth. However, the share of women for whom most of their income is self-employment income matches the real labour supply path.

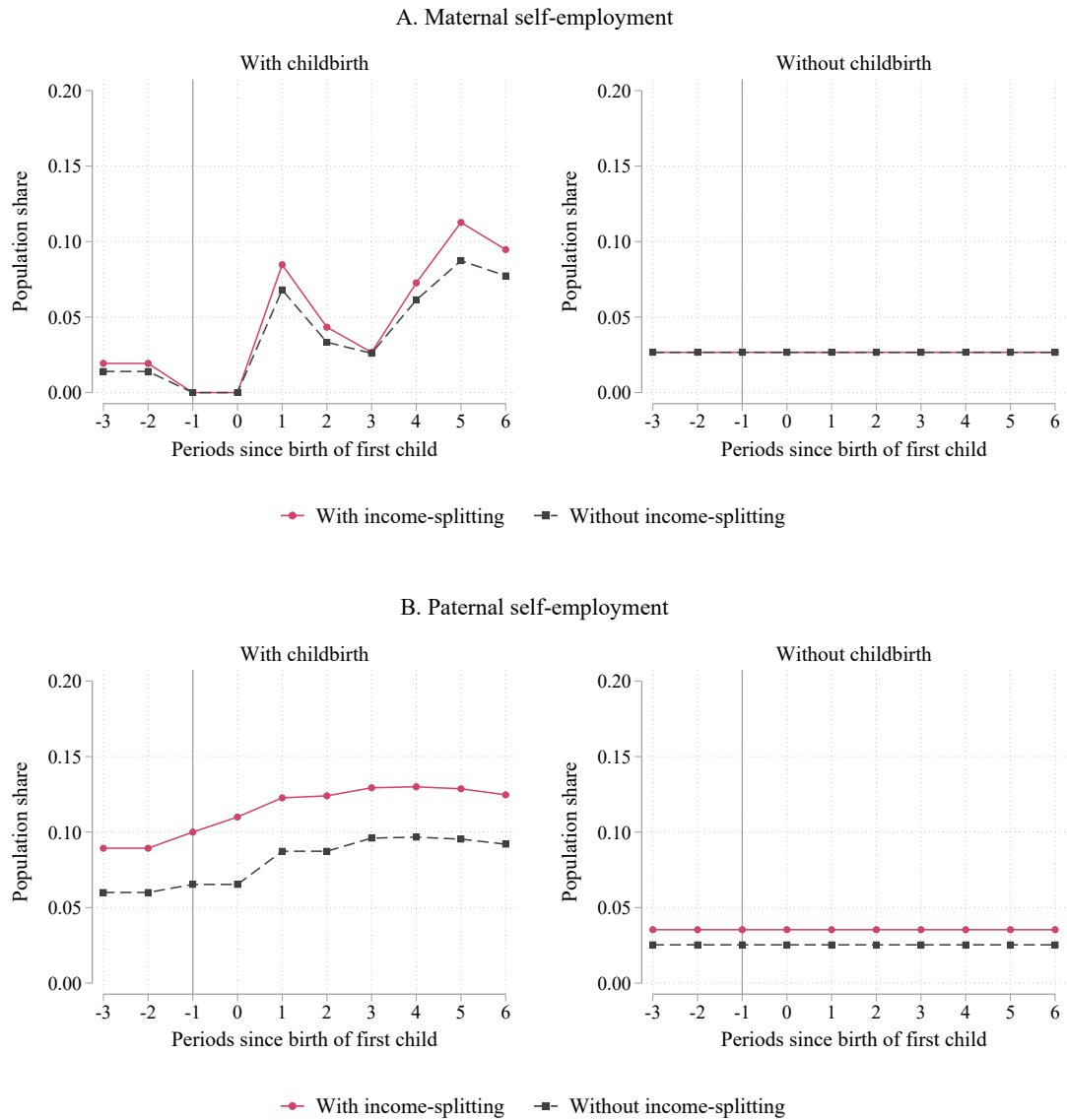


Figure A.1: Simulated *real* self-employment rate by event-time of childbirth

Note: Each marker denotes the self-employment rate (population share) in each event-time period (the model equivalent of years since initial childbirth). The model outcomes are solved under the parameters/distributions described Section A.1. In each panel the right-hand side figures references the counterfactual without childbirth. Each figure contains two curves, referencing the optimal level of self-employment with and without income-splitting.

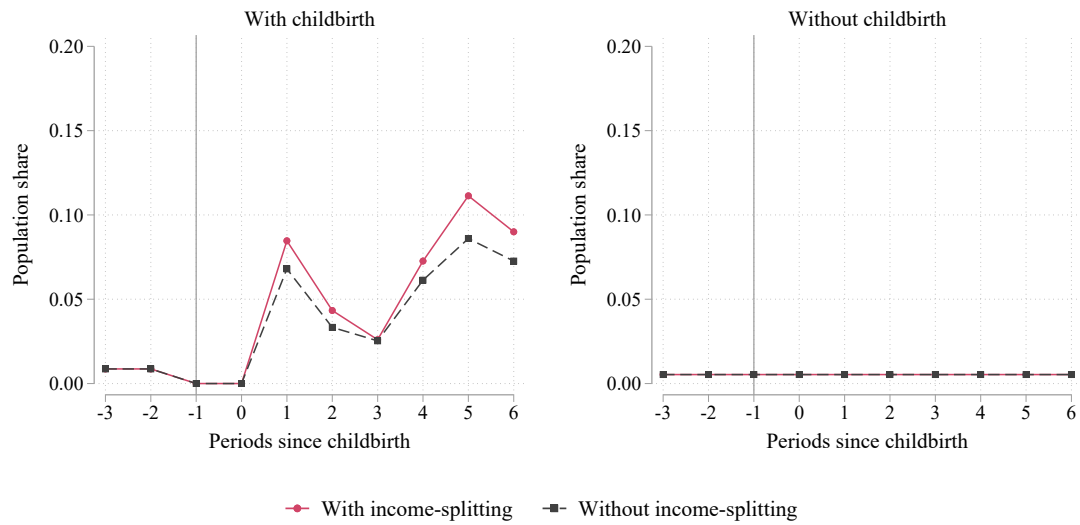


Figure A.2: Simulated *real* co-employment by event-time of childbirth

Note: Each marker denotes the co-employment rate (population share) in each event-time period (the model equivalent of years since initial childbirth). The model outcomes are solved under the parameters/distributions described above. The right-hand side figure references the counterfactual without childbirth. Each figure contains two curves, referencing the optimal level of co-employment with and without income-splitting.

A.2 Identifying childbirth in the LAD (continued)

The working sample is built by identifying all women 20 years of age and older, regardless of marital status. These women may still reside with their parents at the time of sample entry. The year of initial childbirth is then identified, and retroactively assigned to past records. On average I identify about 20,000 *first time* births a year, or 100,000 when weighted. Canada's annual birth rate is around 300,000 for the respective period, but this figure represents all births. Given the selection on age, I am confident that I identify the vast majority of initial births for this demographic. I then select out individuals who are observed at least two years prior to childbirth, to avoid sample selection in the event-study-design analysis. For this reason, the sample may have a slightly higher probability of working prior to initial childbirth.

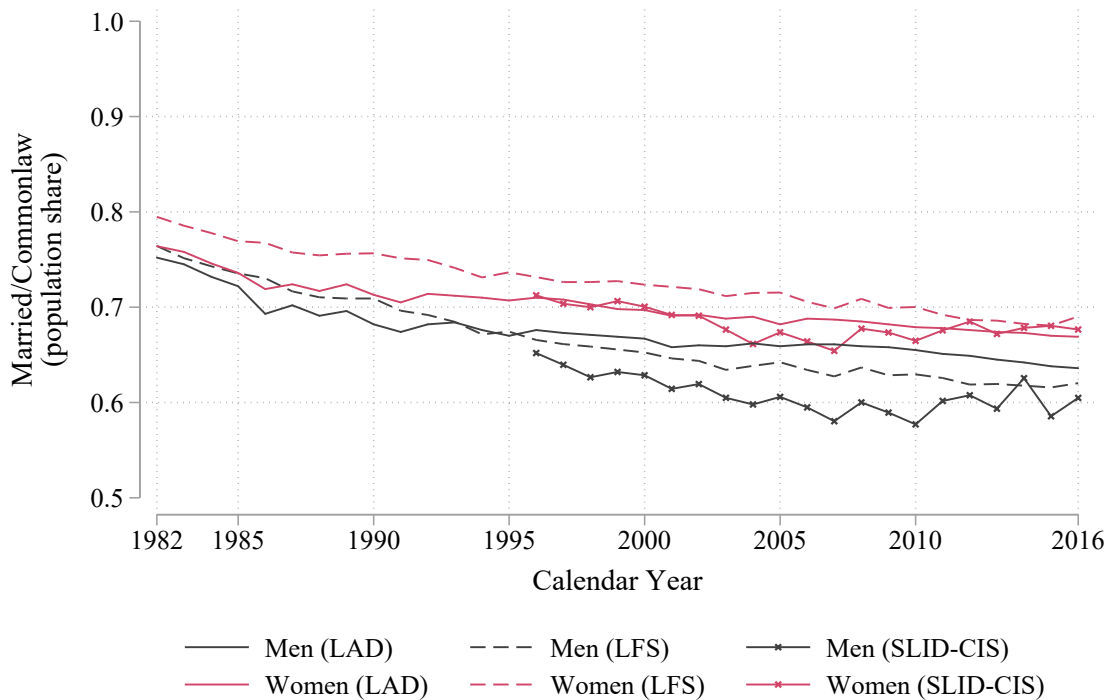


Figure A.3: Comparison of cohabitation rates across the LAD, LFS and SLID/CIS

Note: Estimated using the Longitudinal Administrative Dataset (1982-2016), Labour Force Survey (1982-2016) and SLID/CIS (1996-2016). Sample includes all individuals aged 25-44. Cross sectional weights are applied in the estimation of each series.

Figure A.3 shows the trend in cohabitation (married/common-law couples) in the LAD, alongside comparable estimates using the LFS and SLID/CIS surveys. The LAD series is estimated using the full sample, not the aforementioned estimating sample. There is a

distinct downward trend in cohabitation within this demographic, captured in both the administrative and survey data. While the LAD series slightly underestimates that of the LFS for women, the difference is not stark enough to suggest that the subsample of cohabiting couples is less representative of the overall population.

A.3 Identifying self-employment in the LAD (continued)

As discussed in Section 2.3 I assign individuals to either the wage-paying sector or self-employment based on the primary source of their employment income. While this may appear straight forward, misidentification can easily occur. For example, a worker may be primarily self-employed, but be an employee for part of the year. If their business has either low revenue, or high deductibles, in the short term it is possible that any wage-income they earn is greater than their net self-employment income. Similarly, the need to include dividend income as a form of incorporated self-employment income has the potential of conflating holders of large investment portfolios with self-employed workers.

To deal with these issues of misidentification I use a series of income thresholds. In selecting these thresholds, I exploit the SLID's matched self-reported survey and tax data; where the former includes self-reported labour market status, including main class of employment. In each case I pick the threshold that gives me the best trade off between type one and two errors. In this instance, a type one error means that I identify an individual as self-employed when they are not, while a type two error means I fail to identify an individual as self-employed who is.

Recall that while unincorporated self-employment income is filed separately, incorporated income from a closely held firm is not. For this reason, I use dividend receipts and uninsured wage income above a chosen threshold to identify incorporated income. For uninsured wage receipts this is around \$5,000 (in 2016 CAD), while for dividend income it is closer to \$8,000. In addition, I check the type one error for wage-employment and find that this is minimized by setting the wage threshold for uninsured wages to \$0. That is, according to the matched SLID sample uninsured wages do not represent wage-employment. On this basis, uninsured wage receipts below the threshold of \$5000 are disregarded as employment income. Regarding dividend income, I find that including dividends as a form of employment income vastly improves the type two error on employment status. If you do not include dividends as a form of employment income you underestimate the number of employed individuals in the sample.

Figure A.4 compares my measure of self-employment from the LAD with that of the LFS and SLID/CIS, for the sample of individuals aged 25-44. In the case of the SLID/CIS I use the self-reported presence of any self-employment (incorporated or unincorporated) income in the past year. While the SLID assigns each individual a main job from the past year, the CIS does not. Nevertheless, as is, the SLID/CIS series provides an upper bound on how many individuals had some self-employment work in the past year and may report some self-employment income. As expected, the series exceeds the level of self-employment in the LFS which reflects an individual's self-reported main job. The average difference between the LFS and SLID/CIS of $\sim 3\%$ suggests that there is the potential to overestimate

self-employment using the LAD. This does not appear to be the case. The LAD series is always below the SLID/CIS level, suggesting that by selecting only those individuals with more than 50% of their employment income from self-employment sources does a good job of identifying the main job of an individual.

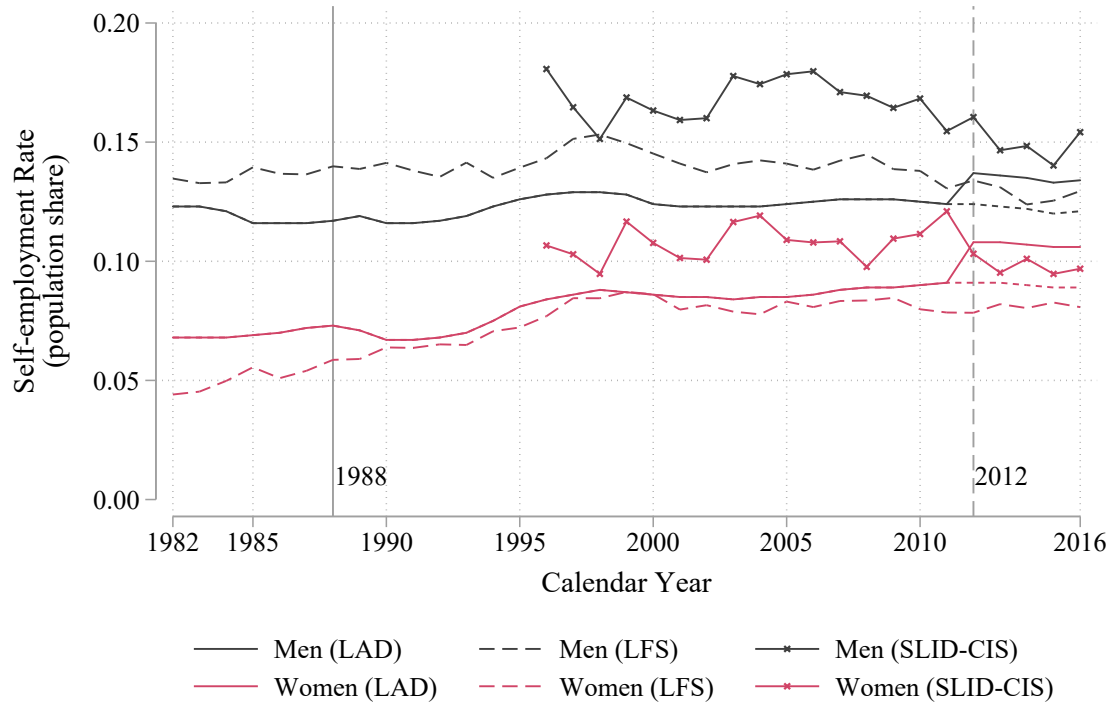


Figure A.4: Comparison of the self-employment rate in the LAD, LFS and SLID/CIS

Note: Estimated using the Longitudinal Administrative Dataset (1982-2016), Labour Force Survey (1982-2016) and SLID/CIS (1996-2016). Sample includes all individuals aged 25-44. Cross sectional weights are applied in the estimation of each series. The LAD series denotes the measured discussed in Section 2.3. The LFS series is based on the individual's self-reported class of employment, associated with their main job (within the last month). The SLID/CIS series indicates the share of individuals who had any self-employment work during the past year.

The LAD series has a discontinuous jump in 2012, that does not appear in either the LFS or SLID/CIS. This jump corresponds to an approximate 2.5 %-point increase in the incidence of uninsured wage income, which I use to identify incorporated self-employment. Using information on the industries of these wage receipts (available from 2000 onwards) I know that the increase is not isolated to industries typically associated with incorporation; for example, a significant share appears in the public sector. Fortunately, I am able to correct for this misallocation in the longitudinal data. In Figure A.4 I plot the continuation of the LAD series without this jump in 2012 using a dashed line.

Figure A.5 depicts the estimated incorporation rate in the LAD and LFS series. A

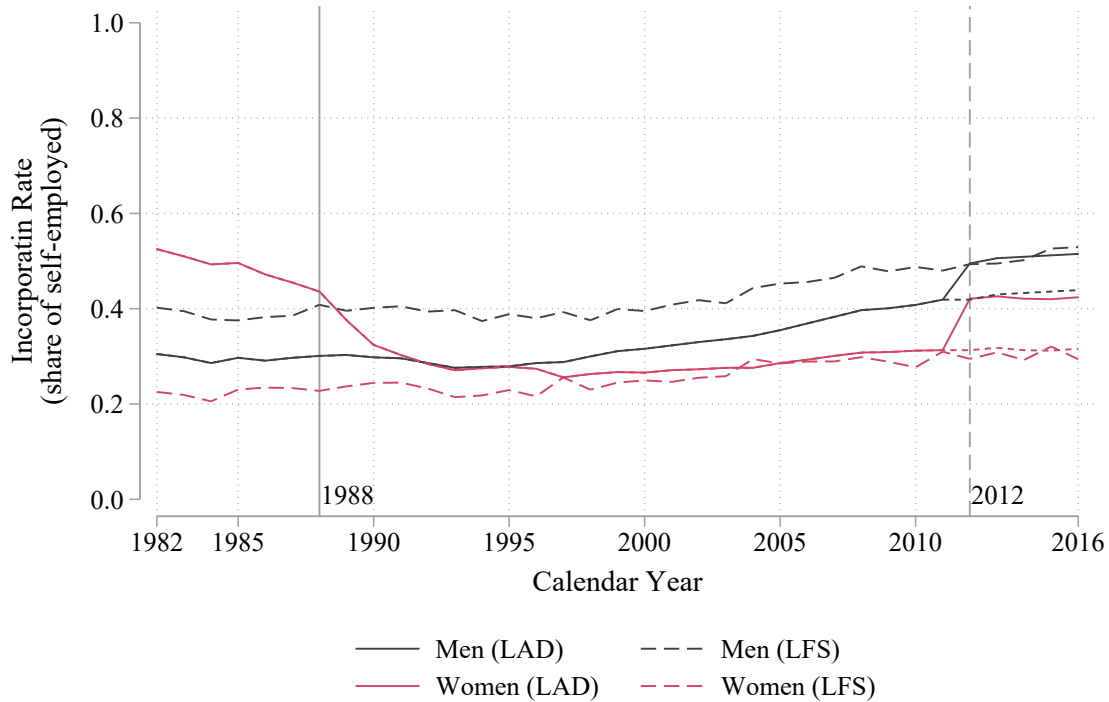


Figure A.5: Incorporation rate of self-employed workers in the LAD and LFS

Note: Estimated using the Longitudinal Administrative Dataset (1982-2016), Labour Force Survey (1982-2016) and SLID/CIS (1996-2016). Sample includes all individuals aged 25-44. Cross sectional weights are applied in the estimation of each series. The LAD series identifies the share of individuals designated self-employed, who derive the majority of their self-employment income from either uninsured wage income or dividends. The LFS series is based on the individual's self-reported class of employment, associated with their main job (within the last month).

self-employed worker is deemed incorporated if more than 50% of their self-employment income is made up of uninsured wage income and dividends (above the chosen threshold). As discussed in Section 2.3 I expect to underestimate incorporation. This is confirmed here by the male series. Note again the correction made after 2012.

For women, I do not underestimate the LFS. In fact, prior to 1988 I vastly overestimate incorporated self-employment. This corresponds to a higher estimated level of female self-employment in the LAD prior to 1988. As with the jump in 2012, Figure A.5 confirms that this is because of a spike in cases of uninsured wage income. I do not believe that this corresponds to a real labour supply adjustment, but rather a change in tax reporting. This is once again a reason why I choose to work with the post 1988 data.

In Section 2.3 I emphasised the fact that the LAD's measure of self-employment is potentially a broader measure of employment within a household firm. Indeed, I find confirmation of this in the fact that, among employed women, type one errors (assigned

self-employed based on tax income when self-reported status is paid employee) are highly correlated with the self-reported self-employment status of their spouse. That is, employed woman married to self-employed men often self identify as a paid employees when their tax records suggest they are self-employed workers. Matching on industry and other employer characteristics suggest that these women are likely employees within the same firm. In this regard, the LAD is a better tool for the analysis of co-employment compared with the LFS.

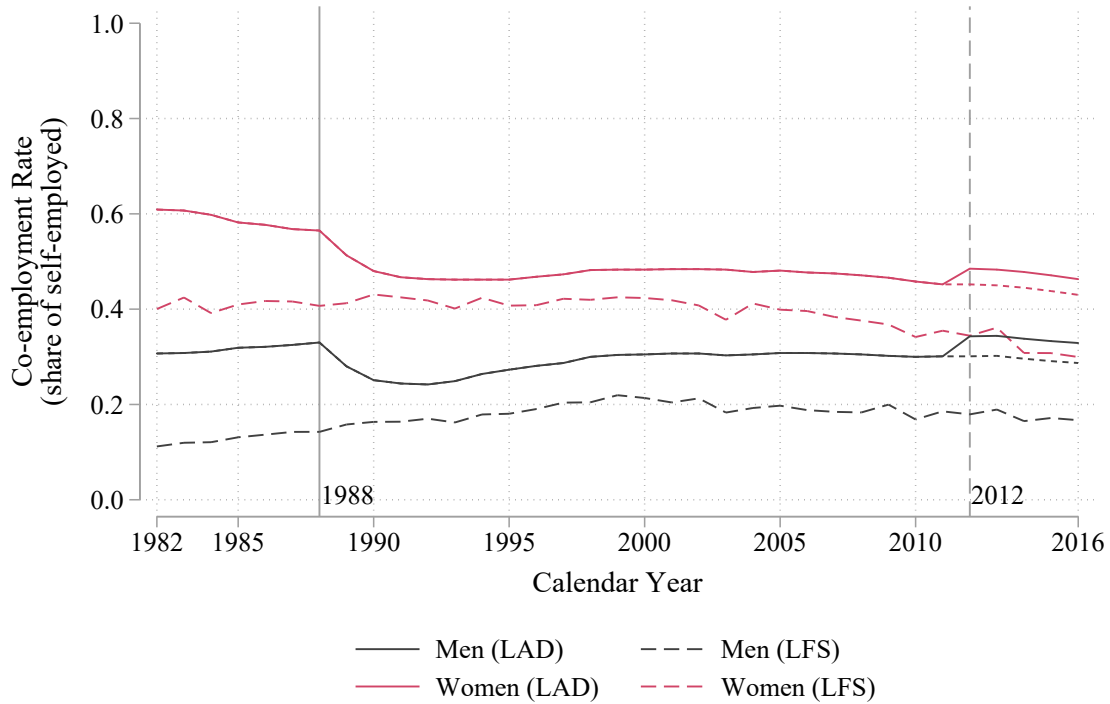


Figure A.6: Comparison of self-employment definitions across LAD and LFS for men and women aged 25-44.

Note: Estimated using the Longitudinal Administrative Dataset (1982-2016), Labour Force Survey (1982-2016) and SLID/CIS (1996-2016). Sample includes all individuals aged 25-44. Cross sectional weights are applied in the estimation of each series. The LAD series denotes the share of married couples where both spouses are designated self-employed. The LFS denotes the share of married couples where both individuals self-report their main job as self-employed (incorporated or unincorporated).

Figure A.6 depicts the trend in joint self-employment (couples where the male and female spouse are both self-employed) across the LAD and LFS. In the LFS co-employment has been falling since the turn of the century, while in the LAD the decline has been less steep. This aligns with the shift towards incorporation among self-employed workers and suggests that co-employed women may report their class of employment differently depending on the incorporation status of their spouse's business.

To summarize, I make use of the SLID's matching of annual tax records with self-reported employment status and class of employment to pick the thresholds used to define uninsured wage income and dividend income as incorporated business income. When combined with unincorporated self-employment income I am then able to assign individuals a self-employed, paid employed, or not employed status. By identifying self-employment in this way, I inadvertently end up with a measure that more broadly encompasses individuals who would typically self-report as self-employed workers as well as co-employed individuals who may self-report as paid employees (or even unpaid family workers). The LAD's measure of self-employment can more broadly be defined as a measure of employment within a closely held firm.

A.4 Event-study-design coefficients

This appendix provides plots of the underlying event-time coefficients corresponding to Figures 2.1-2.6. These plots are shown together with 95% confidence intervals. Note that in many instances the scale of the y-axis has been adjusted to allow for a closer examination of coefficient values.

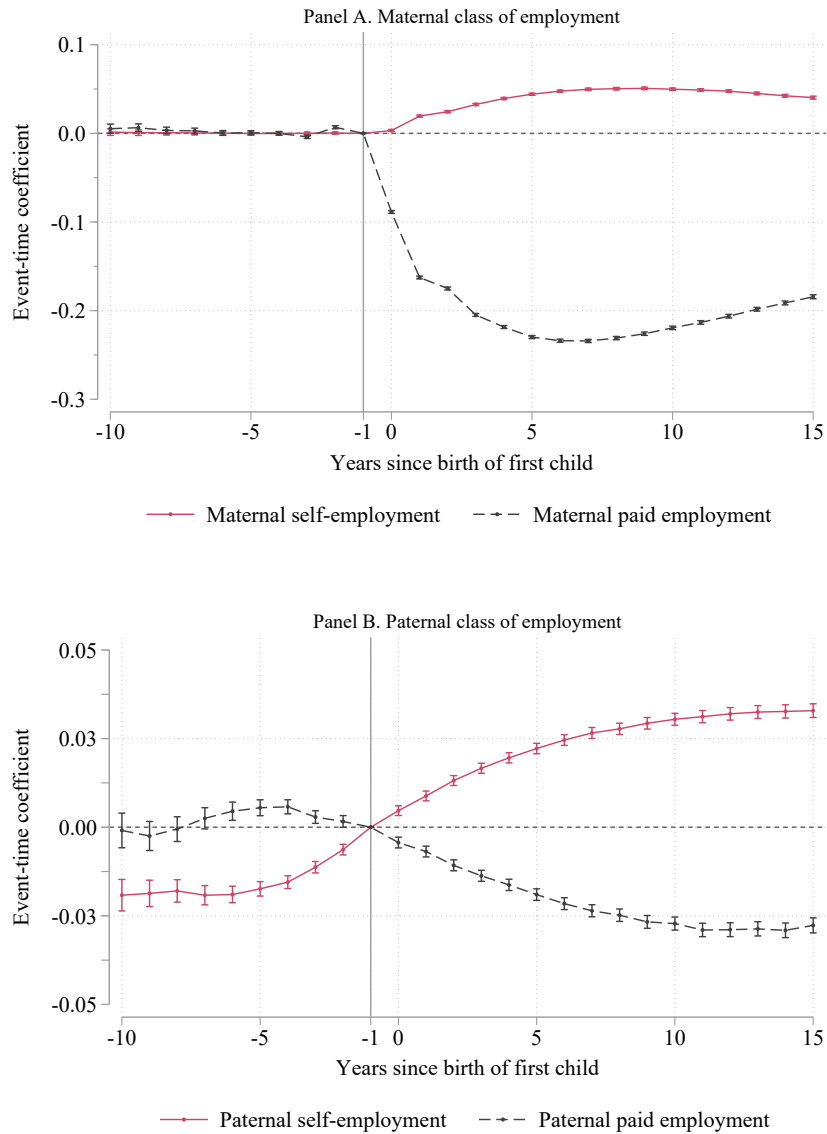


Figure A.7: Event-study-design coefficient: class of employment

Note: Estimated using the Longitudinal Administrative Dataset (1988-2016). Sample includes all married couples where the female spouse is at least 20 years of age. Each data point corresponds to the rescaled event-time coefficient from an event-study-design (ESD) of initial childbirth. In each panel the wage-employment and self-employment curves, correspond to separate ESD models with the discrete outcome of participation in each sector. In both cases the coefficients are normalized by the same value: the predicted level of employment under the counterfactual that the event of childbirth did not occur. See discussion in Section 2.4. Confidence intervals are not shown as they are too small to plot. Panel A corresponds to the maternal response, while Panel B to the paternal response. In both instances the age-dummies included in the underlying models are those of the mother.

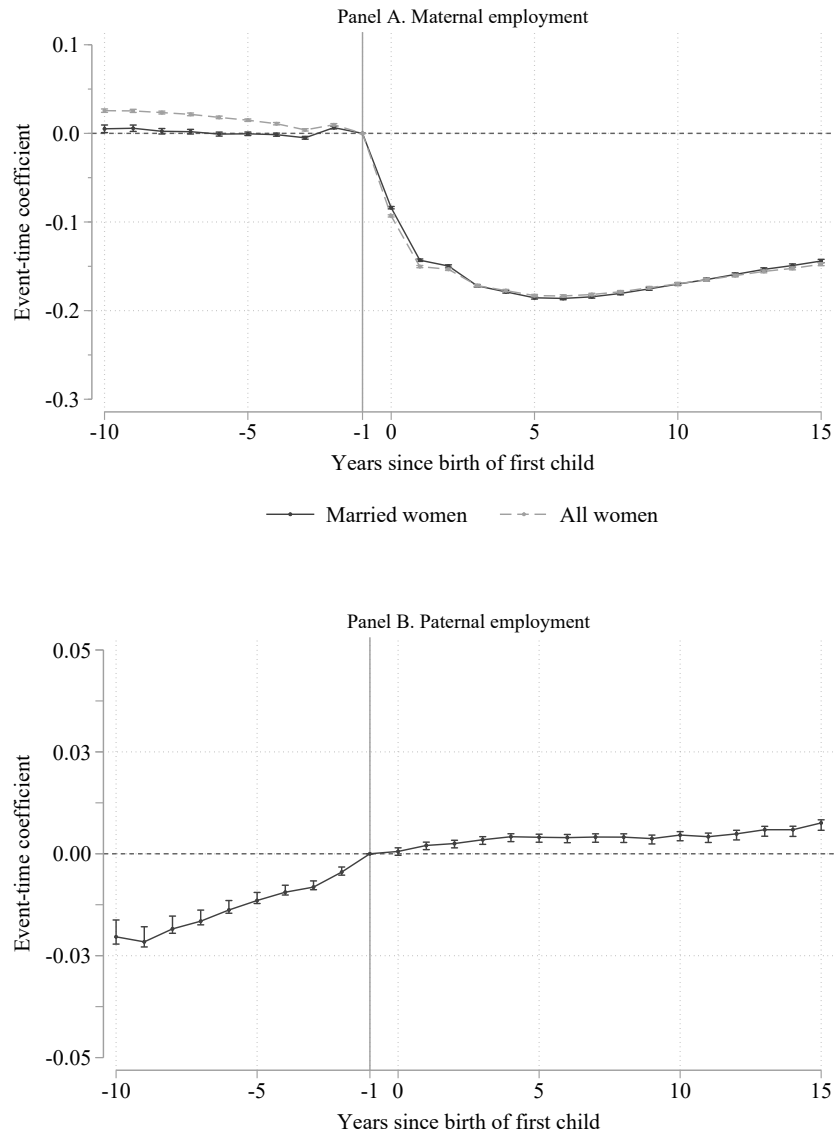


Figure A.8: Event-study-design coefficient: overall employment

Note: Estimated using the Longitudinal Administrative Dataset (1988-2016). Sample includes all married couples where the female spouse is at least 20 years of age. Each data point corresponds to the event-time coefficient from an event-study-design (ESD) of initial childbirth. Corresponding 95% confidence intervals are shown. Both the paternal and maternal curves correspond to a model in which the included age-dummies are those of the mother.

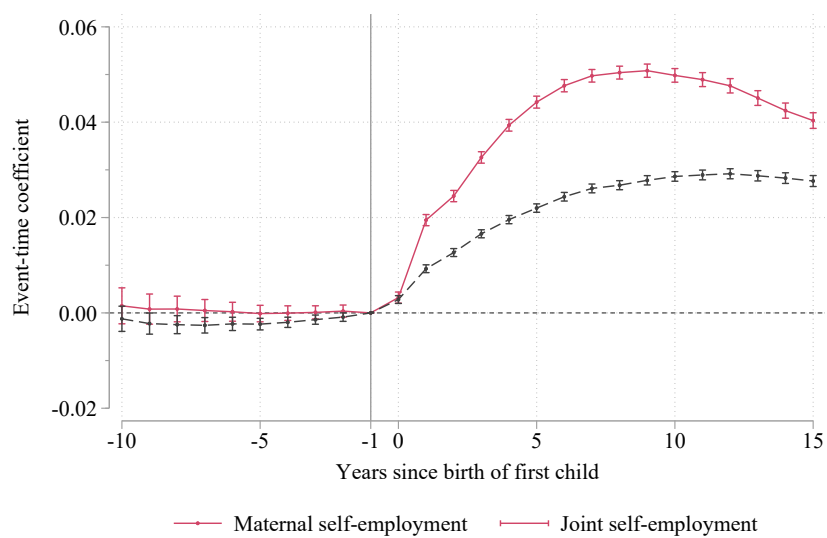


Figure A.9: Event-study-design coefficient: maternal self-employment and joint self-employment

Note: Estimated using the Longitudinal Administrative Dataset (1988-2016). Sample includes all married couples where the female spouse is at least 20 years of age. Each data point corresponds to the event-time coefficient from an event-study-design (ESD) of initial childbirth. Corresponding 95% confidence intervals are shown. Both the paternal and maternal curves correspond to a model in which the included age-dummies are those of the mother.

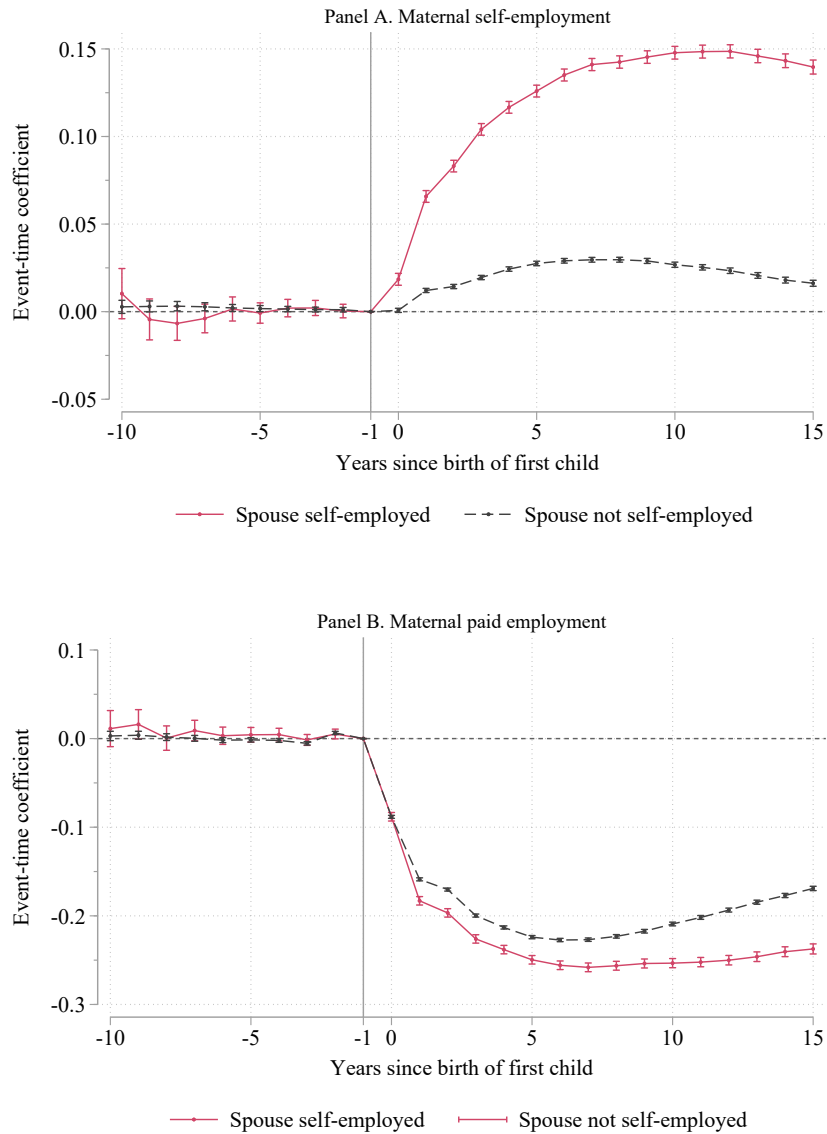


Figure A.10: Event-study-design coefficient: class of employment, by class of spouse

Note: Estimated using the Longitudinal Administrative Dataset (1988-2016). Sample includes all married couples where the female spouse is at least 20 years of age. Each data point corresponds to the event-time coefficient from an event-study-design (ESD) of initial childbirth. Corresponding 95% confidence intervals are shown. Panel A represents the outcome of self-employment, while Panel B the outcome of wage-employment. The coefficients are estimated by interacting the event-time and age dummy variables with the self-employment status of the spouse. The year fixed effects are not interacted.

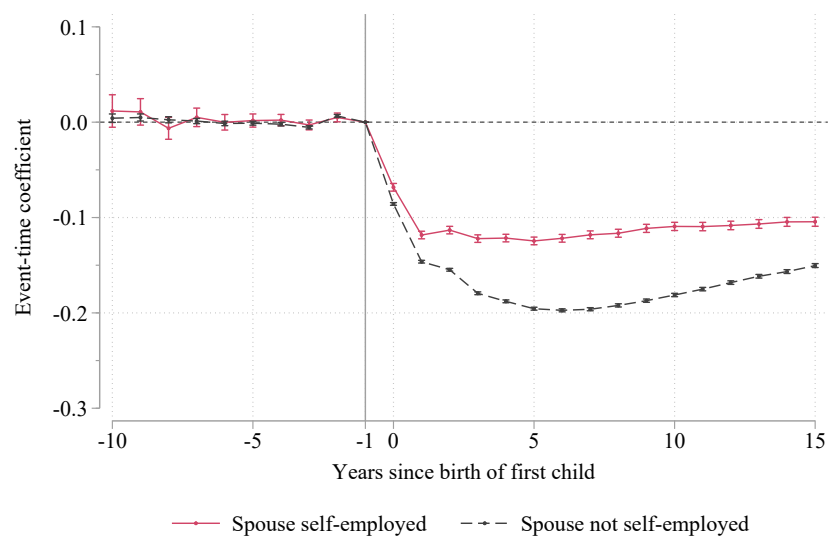


Figure A.11: Event-study-design coefficient: overall employment, by class of spouse

Note: Estimated using the Longitudinal Administrative Dataset (1988-2016). Sample includes all married couples where the female spouse is at least 20 years of age. Each data point corresponds to the event-time coefficient from an event-study-design (ESD) of initial childbirth. Corresponding 95% confidence intervals are shown. Panel A represents the outcome of self-employment, while Panel B the outcome of wage-employment. The coefficients are estimated by interacting the event-time and age dummy variables with the self-employment status of the spouse. The year fixed effects are not interacted.

A.5 Accounting for Regional Differences in Co-employment

In this section, I explore the regional variation in the selection into self-employment and co-employment with childbirth across provincial regions of Canada.¹⁸⁰ A parallel analysis of regional selection into self-employment among men in the LFS corroborates the LAD's event-study results and suggests that accounting for differences in individual and labour market characteristics can explain much of the unconditional variation. Nevertheless, a clear East-West pattern remains. Men in the West of Canada are more likely to be self-employed after childbirth, conditional on observed characteristics. I then provide some preliminary evidence from qualitative surveys that suggests preferences over work, and not gender norms, explain this residual regional variation.

Figures A.12-A.14 plot the LAD-based event-study estimates by region. The results reveal considerable regional variation in the decision to enter self-employment and co-employment with childbirth. In the Prairies and Alberta maternal self-employment increases by over 9% (relative to predicted employment) after childbirth, while in the Atlantic provinces and Quebec this increase is less than 3%. These differences extend to spousal switching between paid and self-employment, with the Prairies and Alberta demonstrating the largest increase in spousal self-employment. Ontario and British Columbia are closer to the federal average. In this section, I demonstrate that a large share of this regional variation can be explained by differences in individual covariates, labour market composition and local labour market conditions. I also provide a discussion of factors that *may* explain the remaining residual regional variation.

As the LAD does not contain certain key individual covariates - namely education, industry¹⁸¹, and occupation - I provide a brief analysis here of male self-employment using the pooled LFS (1988-2016). For reasons discussed in Section 2.4.2 (see also Section A.3), the maternal self-employment response measured using the LAD is only partially captured by self-reported class of employment in survey data. In contrast, the male self-employment response is highly comparable between the LAD and SLID. Hence, I focus here on understanding male selection into self-employment with family formation in the LFS.

Let the family formation gap reflect the difference in the probability of self-employment between individuals with and without children, conditional on age and year. This is the

¹⁸⁰ As is common in Canadian research, I group the four Atlantic provinces of Newfoundland & Labrador, Prince Edward Island, Nova Scotia, and New Brunswick, as well as the Prairie provinces of Manitoba and Saskatchewan.

¹⁸¹ From 2000 the LAD does provide industry codes for the first and second filed T4 slip. However, this excludes individuals with only unincorporated self-employment income, or incorporated workers who issue dividends alone. While unincorporated income is reported within five categories - farming, fishing, professional, business, and commission - these are insufficient to provide a mapping to the industry codes.

closest specification to the event-study-design using cross-sectional data. The relative size of this family formation gap by region should be proportional to the event-study-design estimates in Figures A.13. I estimate a region-specific family formation gap (δ_p^c) using the simple linear probability model described below. The sample includes all employed men, cohabiting with women aged 20-44; the period of the life cycle associated with early family formation. The majority of these men are 30-44 years of age, the period of the life cycle associated with the sharpest increase in self-employment.

$$Q_{it}^2 = \delta_p^c \cdot \mathbf{1}\{child\} + \sum_k \beta_{pk}^2 \cdot \mathbf{1}\{k = age_{it}\} + \delta_p + \delta_t + X'_{it}\beta + \varepsilon_{it}$$

Figure A.15 plots the conditional and unconditional region-specific family formation gap (δ_p^c). The standard (“unconditional”) model includes controls for five-year age-group (interacted with region), province fixed-effects and year fixed-effects. The conditional model includes province-times-year fixed-effects (δ_{pt}), survey month, education, industry, and province-month unemployment rate.¹⁸² The province-times-year fixed-effects will absorb any province specific policy variation; for example, changes to the small business corporate tax rate and province-specific parental leave reforms (Baker and Milligan, 2008). I also include a separate specification with occupation. In this instance the inclusion of occupation is particularly problematic. While occupation status may be an important determinant of an individual’s ability to become self-employed (e.g., professionals), it is also endogenous to an individual’s self-employment state. Self-employed workers are more likely to reporting being in a management occupation.

The aggregate family formation gap in Canada is 2.5% (see Table 2.4), slightly lower than $\sim 3\%$ event-time 5 estimate in Figure 2.1.¹⁸³ By region, the unconditional estimates are indeed proportional to the regional event-study results of Figure A.12. Conditional on age, men with at least one child are 4% more likely to be self-employed in the Prairies and Alberta. The ESD estimates suggest that relative to the year before initial childbirth paternal self-employment increases by 5 %-points by year 5 after childbirth in Alberta and the Prairies. In contrast, event-study estimates for Ontario are closer to 2.5 %-points, similar to those from the LFS. It is evident that the included covariates (excluding occupation)

¹⁸²Survey month is included to account for the seasonality of self-employment, which increases during the winter months in Canada. I include dummies for three education categories: high school or less; school and some tertiary education; Bachelor’s degree or more. The province-month unemployment rate of prime-aged working adults is included to account for local labour market conditions. There is disagreement in literature at to the exact nature of the relationship between self-employment and the business cycle (Zissimopoulos and Karoly, 2007).

¹⁸³We can expect the cross-sectional OLS estimates to underestimate the ESD results for two reasons. First, they include the sample of married households which never have children and may have a lower conditional (on age) probability of selecting into self-employment. Second, this single difference estimate represents an average of the event-time estimates, which increase monotonically with event-time.

explain the large increase in paternal self-employment in Alberta and the Prairies. This no doubt captures, the important primary sector occupations in this region. Nevertheless, the residual regional variation maintains a clear East-West path, with Alberta and British Columbia demonstrating the largest (residual) increase in male self-employment associated with childbirth. Some of this is explained by occupation differences, when comparing Ontario, Alberta, and BC.

What then explains the larger increase in male self-employment with family formation as you move from Eastern to Western Canada? Both the parent penalty and relative income literatures have established a strong correlation between their associated labour market outcomes and cultural gender norms across geographic regions (Bertrand et al., 2015; Kleven et al., 2019a). In Canada, Doumbia and Goussé (2019) establish this same relationship as Bertrand et al. (2015) between the discontinuity in the relative income distributions of Canadian provinces and reported gender norms using data from the European Values Study and the World Value Survey.¹⁸⁴ This paper has argued, in agreement with Zinovyeva and Tverdostup (2018), that self-employed households lie at the root of this relationship. As such, it is possible that gender norms drive selection into self-employment with family formation. Indeed, Binder and Lam (2019) suggest there may even be an ‘equal-earning’ norm that explains bunching in the relative income distribution.

Researchers have identified several questions that illicit such gender norms; for example, how much do you agree with the statement, “being a housewife is fulfilling”? In Table A.1, I provide summary statistics of responses to such questions using data from the Canadian World Value Surveys (WVS) of 1990, 2000, and 2006 (see also Doumbia and Goussé, 2019). I include a limited set of sample characteristics and individuals’ self-reported importance of family, work and leisure. On average, 75% of the sample (of prime-aged working adults) agree with the statement that housewifery is fulfilling, with little evidence of a gender difference. Note, some questions do not appear consistently across each round of WVS. This, in addition to the small sample sizes, emphasise the caution needed in placing too strong an interpretation on these results.

Table A.2 provides these same summary statistics by region of Canada. The question identifying whether individuals agree that both parents should contribute financially to the household has the strongest correlation with the unconditional LAD and LFS results. In the Prairies and Alberta individuals agree least with this statement; the regions where the co-employment response is greatest. However, the other questions are less consist-

¹⁸⁴Using the same data as Bertrand et al. (2015), Binder and Lam (2019) show that equal earning couples drive the estimated discontinuity on either side of the 50 threshold in the relative earning distribution. They suggest that the “the mass of equal-earning couples implies an equal-earning norm, at least for a segment of the population”, but also argue that it may be the result of frictions in the marriage market. For example, if couples are more likely to meet at their place of work.

ent. Moreover, none of the reported gender norm questions suggest a strong East-West correlation consistent with the conditional LFS results.

In this regard, the reported individual preferences over family, work, and leisure demonstrate a far stronger East-West pattern. As you move from East to West coast in Canada individuals place less importance on work in their life. Interestingly, this is not because they appear to place more importance on leisure, or family. As Table A.3 highlights, if you group prime-aged adults by employment status, instead of region, you find that self-employed workers are less likely to say that work is *very* important. While 57% of paid workers agree that work is VERY important, only 48% of self-employed workers think so. Differences in individual characteristics and gender norms are less striking. Unfortunately, I have not been able to establish a more robust, regression-based correlation between self-employment status, gender norms, and work preferences for Canada.

However, regional differences in self-reported reasons for self-employment may help to corroborate the importance of work-based preferences in determining residual selection into self-employment. The 2016 General Social Survey of Canada (Round 30: "Canadians at Work and Home") included such a focused questionnaire. Table A.4 reports these differences by gender. Consistent with the findings of this paper self-employed women are more likely to have a child, and report work-life balance as their reason for self-employment. In contrast, the most common response among men is the freedom to be one's own boss.

When aggregating these responses by region, I combine the responses into three groups: pecuniary ("earning potential"); non-pecuniary ("work-life balance", "freedom to be own boss", and "less stress"); and other reasons. In Alberta and British Columbia non-pecuniary reasons are reported more frequently than in Ontario, the Prairies, and Atlantic provinces. In Quebec self-employed workers are least likely to report pecuniary reasons, and most likely to report non-pecuniary reasons. It is not clear whether this is because of the relatively large sample of female self-employed workers in Quebec.¹⁸⁵ With the exception of Quebec, the emphasis on non-pecuniary factors is consistent with the lower importance of work in these regions. Together, these descriptive results suggest that preferences related to work - in particular, those related to flexibility - may explain the residual regional variation in the selection into self-employment and co-employment with childbirth. While inconclusive, this suggests that preferences related to work, and not gender, may also explain the observed patterns in the relative income distributions of couples and play an important role in the shape and size of the parent penalty. In contrast to the current literature, the presence of more equal earning couples may reflect a preference for work life balance and self-determination.

¹⁸⁵Note, there appears to be under sampling of Quebec among self-employed workers.

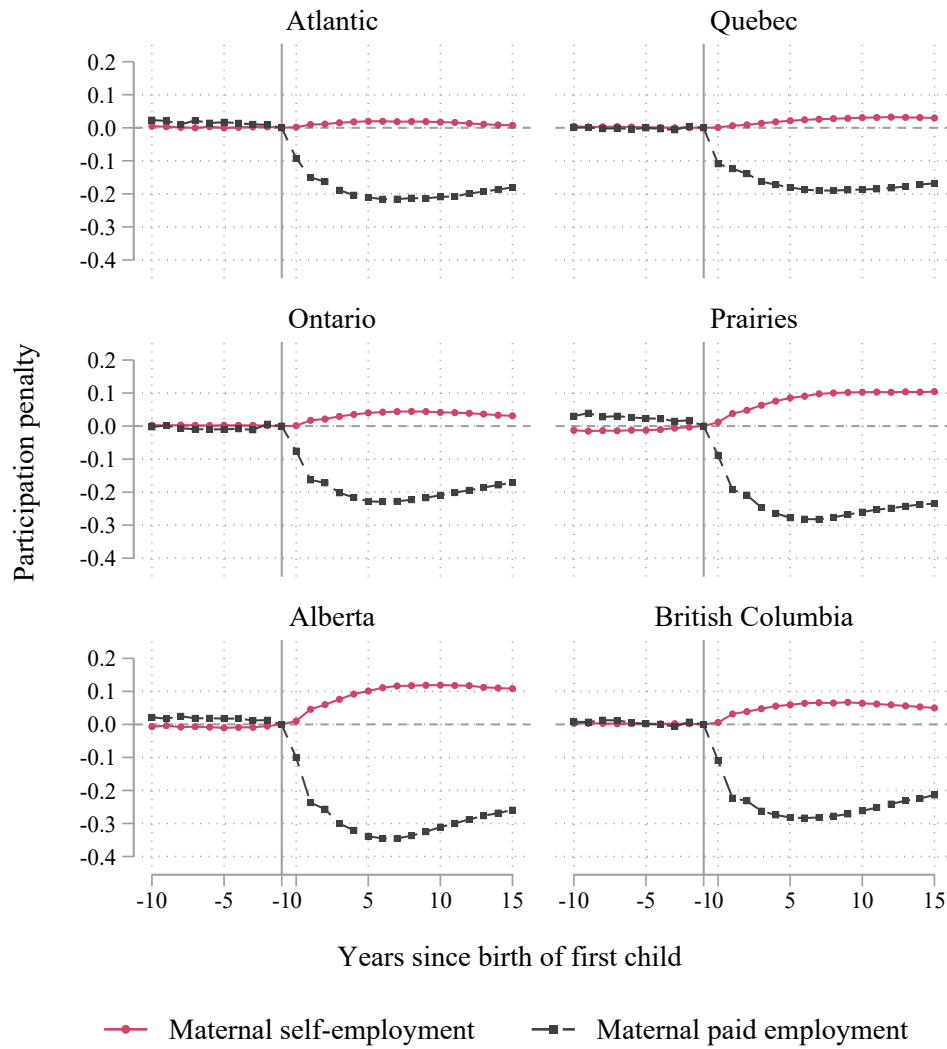


Figure A.12: Participation parent penalty: maternal class of employment by geographical region

Note: Estimated using the Longitudinal Administrative Dataset (1988-2016). The corresponding ESD models are estimated separately for region. The normalization is also region specific: the ESD coefficient is divided by the predicted level of employment under the counterfactual of no childbirth in that specific region. See details in Figure 2.1.

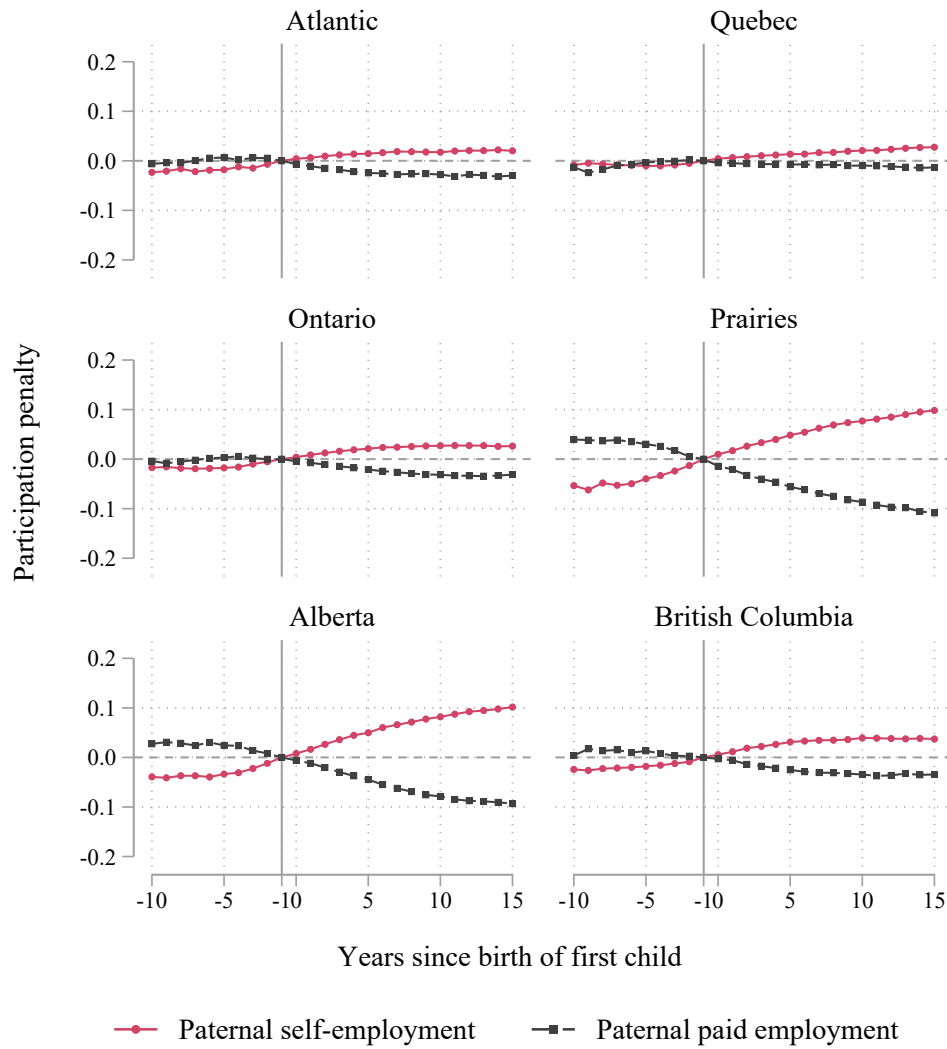


Figure A.13: Participation parent penalty: paternal class of employment by geographical region

Note: Estimated using the Longitudinal Administrative Dataset (1988-2016). The corresponding ESD models are estimated separately for region. The normalization is also region specific: the ESD coefficient is divided by the predicted level of employment under the counterfactual of no childbirth in that specific region. See details in Figure 2.1.

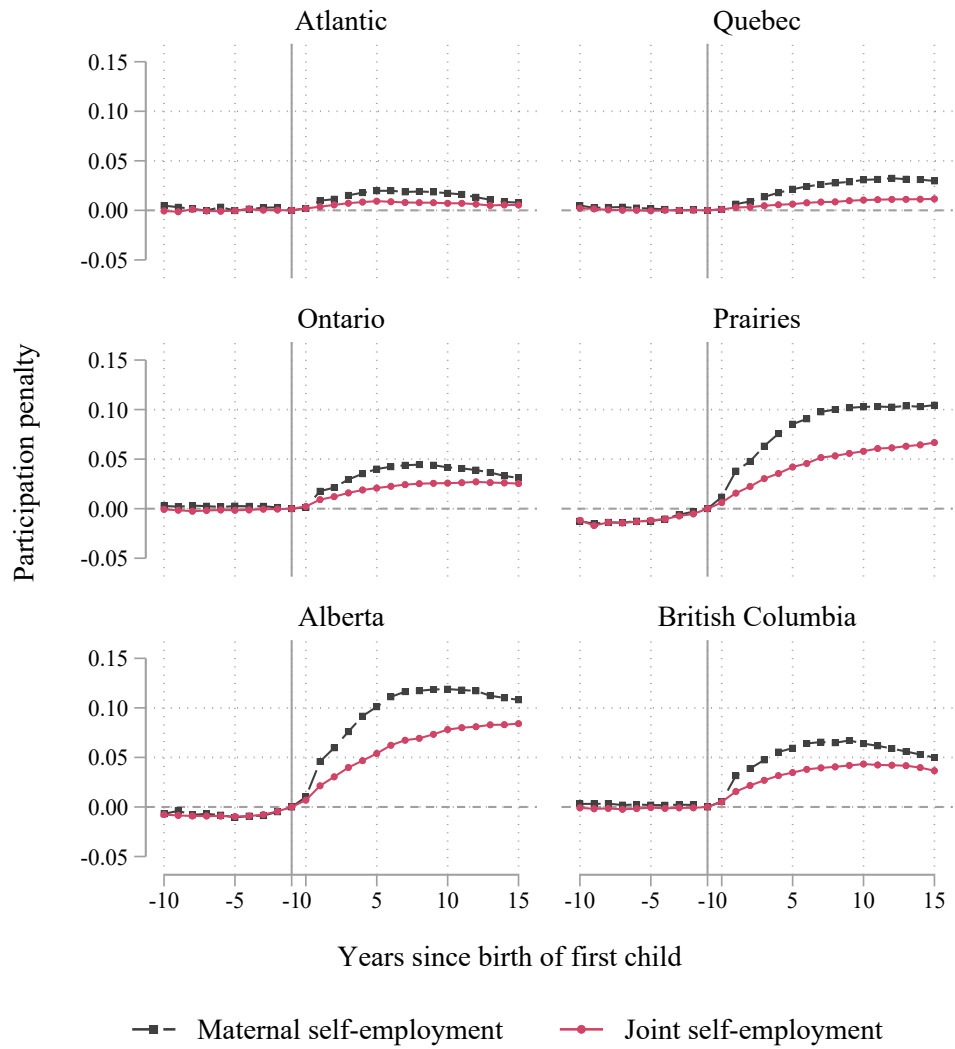


Figure A.14: Participation parent penalty: maternal self-employment and joint self-employment by geographical region

Note: Estimated using the Longitudinal Administrative Dataset (1988-2016). The corresponding ESD models are estimated separately for region. The normalization is also region specific: the ESD coefficient is divided by the predicted level of employment under the counterfactual of no childbirth in that specific region. See details in Figure 2.1.

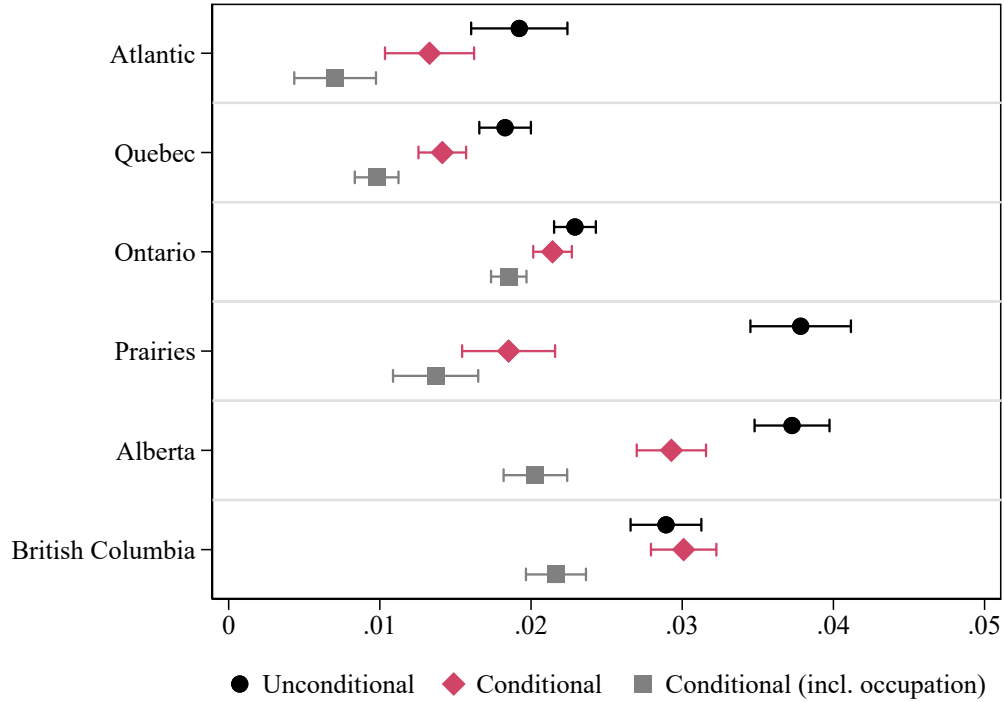


Figure A.15: Region-specific child fixed effect.

Note: Estimated using the public release version of the Labour Force Survey (1988-2015). Sample includes all couples, where the female spouse is aged 20-44, and male spouse is aged 20-54 and is employed. Each marker represents the coefficient on the indicator for the presence of children younger than 18 years of age in the household. Estimates are weighted using cross-sectional weights, and 95% confidence intervals are shown.

Table A.1: Summary statistics and gender norms, by gender: WVS 1990, 2000, and 2006

	<i>All</i>	<i>Men</i>	<i>Women</i>
<i>Characteristics</i>			
At least 1 child	0.75	0.69	0.82
Married	0.73	0.72	0.73
Employed	0.77	0.86	0.68
Self-employed	0.07	0.08	0.07
<i>Areas that are VERY important</i>			
Family	0.93	0.89	0.97
Work	0.55	0.61	0.49
Leisure	0.41	0.40	0.42
<i>Agree or strongly agree with the statement</i>			
Housewifery is fulfilling	0.75	0.74	0.76
Men have right to scarce work	0.12	0.12	0.13
Both parents should contribute [†]	0.70	0.71	0.69
Men make better leaders [‡]	0.21	0.25	0.16
<i>N</i>	3286	1386	1898

Note: Own calculations using the European Value Survey (1990) and World Value Survey (2000, 2006). The sample includes adults aged 25-54. Survey weights applied.

[†] Only asked in 1990 and 2000

[‡] Only asked in 2000 and 2006

Table A.2: Summary statistics and gender norms, by region: WVS 1990, 2000, and 2006

	<i>Atlantic</i>	<i>Quebec</i>	<i>Ontario</i>	<i>Prairies</i>	<i>Alberta</i>	<i>British Columbia</i>
<i>Characteristics</i>						
Male	0.49	0.50	0.49	0.49	0.52	0.50
At least 1 child	0.81	0.76	0.75	0.76	0.74	0.74
Married	0.72	0.72	0.74	0.73	0.72	0.72
Employed	0.68	0.76	0.78	0.84	0.83	0.74
Self-employed	0.05	0.06	0.07	0.10	0.09	0.12
<i>Areas that are VERY important</i>						
Family	0.96	0.89	0.95	0.96	0.93	0.94
Work	0.63	0.58	0.57	0.51	0.48	0.49
Leisure	0.50	0.31	0.45	0.39	0.47	0.39
<i>Agree or strongly agree with the statement</i>						
Housewifery is fulfilling	0.77	0.81	0.70	0.82	0.76	0.72
Men have right to scarce work	0.10	0.18	0.11	0.12	0.10	0.09
Both parents should contribute [†]	0.70	0.78	0.70	0.63	0.59	0.66
Men make better leaders [‡]	0.13	0.17	0.23	0.21	0.19	0.25
<i>N</i>	616	822	980	222	300	346

Note: Own calculations using the European Value Survey (1990) and World Value Survey (2000, 2006). The sample includes adults aged 25-54. Survey weights applied.

[†] Only asked in 1990 and 2000

[‡] Only asked in 2000 and 2006

Table A.3: Summary statistics and gender norms, by employment: WVS 1990, 2000, and 2006

	<i>Not employed</i>	<i>Self-employed</i>	<i>Wage-employed</i>
<i>Characteristics</i>			
Male	0.30	0.55	0.55
At least 1 child	0.81	0.72	0.74
Married	0.66	0.78	0.74
<i>Areas that are VERY important</i>			
Family	0.94	0.89	0.93
Work	0.52	0.48	0.57
Leisure	0.35	0.46	0.42
<i>Agree or strongly agree with the statement</i>			
Housewifery is fulfilling	0.78	0.79	0.74
Men have right to scarce work	0.19	0.07	0.11
Both parents should contribute [†]	0.67	0.68	0.71
Men make better leaders [‡]	0.21	0.20	0.21
<i>N</i>	823	224	2232

Note: Own calculations using the European Value Survey (1990) and World Value Survey (2000, 2006). The sample includes adults aged 25-54. Survey weights applied.

[†] Only asked in 1990 and 2000

[‡] Only asked in 2000 and 2006

Table A.4: Reason for self-employment among prime-aged adults, by gender: GSS 2016

	<i>All</i>	<i>Men</i>	<i>Women</i>
<i>Characteristics</i>			
At least 1 child	0.56	0.53	0.61
Married	0.72	0.72	0.72
Business owner	0.88	0.88	0.86
<i>Reasons for self-employment</i>			
No alternative	0.22	0.24	0.19
Work-life balance	0.15	0.10	0.23
Freedom to be own boss	0.24	0.26	0.20
Earning potential	0.12	0.14	0.09
Less stress	0.01	0.01	0.02
Nature of job	0.12	0.10	0.15
Family business	0.07	0.09	0.04
Other	0.07	0.06	0.09
<i>N</i>	804	476	328

Note: Own calculations using the General Social Survey (Round 30, 2016). The sample includes adults aged 25-54 who are currently self-employed. Survey weights applied.

Table A.5: Reason for self-employment among prime-aged adults, by region: GSS 2016

	<i>Atlantic</i>	<i>Quebec</i>	<i>Ontario</i>	<i>Prairies</i>	<i>Alberta</i>	<i>British Columbia</i>
<i>Characteristics</i>						
Male	0.62	0.55	0.57	0.64	0.57	0.63
At least 1 child	0.61	0.54	0.55	0.55	0.66	0.47
Married	0.75	0.65	0.71	0.77	0.77	0.70
Business owner	0.85	0.84	0.87	0.88	0.94	0.91
<i>Reasons for self-employment</i>						
Pecuniary	0.12	0.07	0.12	0.16	0.14	0.13
Non-pecuniary	0.34	0.52	0.32	0.36	0.47	0.45
Other	0.54	0.40	0.55	0.48	0.39	0.42
<i>N</i>	135	130	220	103	98	118

Note: Own calculations using the General Social Survey (Round 30, 2016). The sample includes adults aged 25-54 who are currently self-employed. Survey weights applied.

Pecuniary reasons include "Earning potential". Non-pecuniary reasons include "Work-life balance", "Freedom to be own boss", and "Less stress". Other reasons include "No alternative", "Nature of job", "Family business", and "Other".

A.6 Characteristics of co-employment mothers

This section examines some of the characteristics of co-employed mothers. As discussed in Section 2.3 the identification of self-employed individuals in the LAD may extend to individuals who are paid employees within a closely held firm (including, a self-employed spouse's firm) as such employees need not pay employment insurance premiums. This potential outcome is validated by the higher incorporation rate of self-employed women in the LAD sample (see Appendix A.3), as joint self-employment is more likely with unincorporated businesses.

For this reason, when working with the SLID survey I defined co-employed couples as those who are self-employed in their main jobs, and those who are paid employees but match on employer characteristics with their paid spouse.¹⁸⁶ This is not a perfect identifier, but as demonstrated in Section 2.4.2 including such an adjustment improves the fit of the event-study-design using the SLID with that of LAD (see Figure 2.10). It is also consistent with the fact that matching on industry increases over the lifecycle in households where the male spouse is self-employed in proportion with the increase in joint self-employment (see Figures 2.11 & 2.4). Here I show how the characteristics of employed mothers, aged 20-44, differ dependent on their class of employment and whether they are co-employed.

Table A.6 provides summary statistics of these employed mothers, separated by whether they are paid employees or self-employed. The former excludes co-employed paid employees. The table also separates out self-employed mothers into those who are *individually* self-employed and those who are co-employed. The figures related to labour supply do not suggest a large difference across the group of employed mothers. If anything, self-employed women work more weeks and hours of the year.

Compensation (or income) is rather different. While individually self-employed women report lower income and than paid employees, co-employed mothers report similar income statistics to paid employees. The exception being the effective wage of co-employed mothers which is far higher than any other group; however, there are large outliers for this variable.

Regarding flexibility, both self-employed mothers and co-employed mothers are more likely to work part-time (36%) and work from home (64%). Only 9% of paid employees work from home, while 25% are part-time employees.

Table A.7 tests some of these differences within a regression framework while controlling for differences across in age, education, and industry. Conditional on these covariates self-employed mothers face a steep negative wage premium of 77%. In contrast, there is no evidence of a positive or negative wage premium among co-employed mothers. Thus, co-employed mothers benefit from the flexibility of working from home - a characteristic of

¹⁸⁶These include industry, sector, number of employees, and number of firm locations.

other self-employed mothers - while earning the wages of paid employees. This suggests that co-employment is path to avoiding any compensating wage variation associated with flexibility.

Table A.6: Labour supply characteristics of employed, married mothers in two earner households

	<i>All workers</i>	<i>Paid employees</i>	<i>Self-employed workers</i>		
			<i>All</i>	<i>Self-employed</i>	<i>Co-employed</i>
<i>Labour supply</i>					
Annual weeks	48.41	48.22	49.55	48.66	50.44
Annual hours (main job)	1449.42	1443.14	1491.71	1478.70	1504.44
Annual hours (all jobs)	1525.91	1516.68	1588.19	1569.13	1606.87
Weekly hours	31.24	31.17	31.72	31.72	31.72
<i>Compensation</i>					
Total income	37379.48	38519.18	30451.01	23230.91	37786.72
Income share	0.39	0.40	0.33	0.24	0.43
Market income	33376.40	34484.01	26642.99	20171.88	33217.72
Effective wage	31.98	30.83	39.73	22.16	56.92
<i>Flexibility</i>					
Part-time	0.26	0.25	0.36	0.35	0.37
Work from home	0.17	0.09	0.64	0.68	0.59
<i>N</i>	66849	57406	9443	4759	4684

Note: Estimated using the public use version of the Survey of Labour and Income Dynamics (1996-2010). Sample includes married mothers aged 20-44 in two-earner households, whose husbands are within the aged 20-54. Both spouses were employed at some stage in the past year. As in Table A.7 'self-employed' mothers are those who report that their main job was self-employed while their spouse was a paid employee. Similarly, 'co-employed' mothers refer to women who identify their main job as self-employed and whose spouse is self-employed, as well as those who are paid employees match on job characteristics of their self-employed spouse (see discussion in Appendix A.6). Effective wage is the ratio of total market income and annuals hours worked (across all jobs). With the exception of relative income, the income variables do include negative values. Estimates are unweighted.

Table A.7: Labour market characteristics of self-employed and co-employed mothers, relative to paid employees in two earner households

	<i>ln(Annual hours)</i>	<i>ln(Market income)</i>	<i>ln(Hourly wage)</i>	<i>Income share</i>
	(1)	(2)	(3)	(4)
<i>Base category: paid employees</i>				
Self-employed	-0.134*** (0.0132)	-0.941*** (0.0161)	-0.851*** (0.0146)	-0.160*** (0.00284)
Co-employed	-0.0968*** (0.0131)	-0.238*** (0.0158)	-0.148*** (0.0140)	0.0357*** (0.00286)
Age-group	Yes	Yes	Yes	Yes
<i>N</i>	65021	65146	63743	66051
<i>Base category: paid employees</i>				
Self-employed	-0.0946*** (0.0136)	-0.856*** (0.0158)	-0.797*** (0.0142)	-0.149*** (0.00290)
Co-employed	-0.0461*** (0.0142)	-0.0719*** (0.0163)	-0.0206 (0.0143)	0.0459*** (0.00307)
Age-group	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes
<i>N</i>	63647	63739	62386	64634

Note: Estimated using the public use version of the Survey of Labour and Income Dynamics (1996-2010). Sample includes married mothers aged 20-44 in two-earner households, whose husbands are within the aged 20-54. Both spouses were employed at some stage in the past year. As in Table A.7 'self-employed' mothers are those who report that their main job was self-employed while their spouse was a paid employee. Similarly, 'co-employed' mothers refer to women who identify their main job as self-employed and whose spouse is self-employed, as well as those who are paid employees match on job characteristics of their self-employed spouse (see discussion in Appendix A.6). Age-group refers to 5-year age-group. All models include state and year fixed effects. Additional covariates include education group (at most high school, some tertiary, Bachelor's or more) and industry code (16-group). Effective wage is the ratio of total market income and annuals hours worked (across all jobs). With the exception of relative income, the income variables do include negative values. Estimates are unweighted.

A.7 Relative earnings

This paper provides evidence of extensive bunching at 0.5 in the relative employment earnings of self-employed households (see Figure 2.13). This is consistent with [Zinovyeva and Tverdostup \(2018\)](#)'s view that co-employed households explain the presence of a 'breadwinner' norm ([Bertrand et al., 2015](#)). [Zinovyeva and Tverdostup \(2018\)](#)'s claim is more relevant within the context of an individual tax structure, where self-reported earnings can be manipulated to lower the household's tax liability.¹⁸⁷ As wage and salary employment earnings are typically reported by a third party, bunching in the relative income distribution cannot be the result of such manipulation. However, non-labour market income - including, investment and rental income - can be adjusted through asset allocation.

Figure A.16 depicts the distribution of total taxable income and total employment income for married, non-self-employed couples with and without children. As households with children typically have a more unequal employment income distribution, there is more room for the shifting of non-labour market income. Indeed, Figure A.16 demonstrates such an outcome. Households with children demonstrate a bigger difference between distribution of relative employment and total taxable income, suggesting that such reallocation of non-labour market income takes place. This, once again, highlights the value of income-splitting for households after childbirth, even non-self-employed households.

¹⁸⁷It should be noted that [Bertrand et al. \(2015\)](#)'s analysis concerns US couples who do not face this same incentive. However, there are non-tax reasons to expect co-employed households to bunch at 0.5.

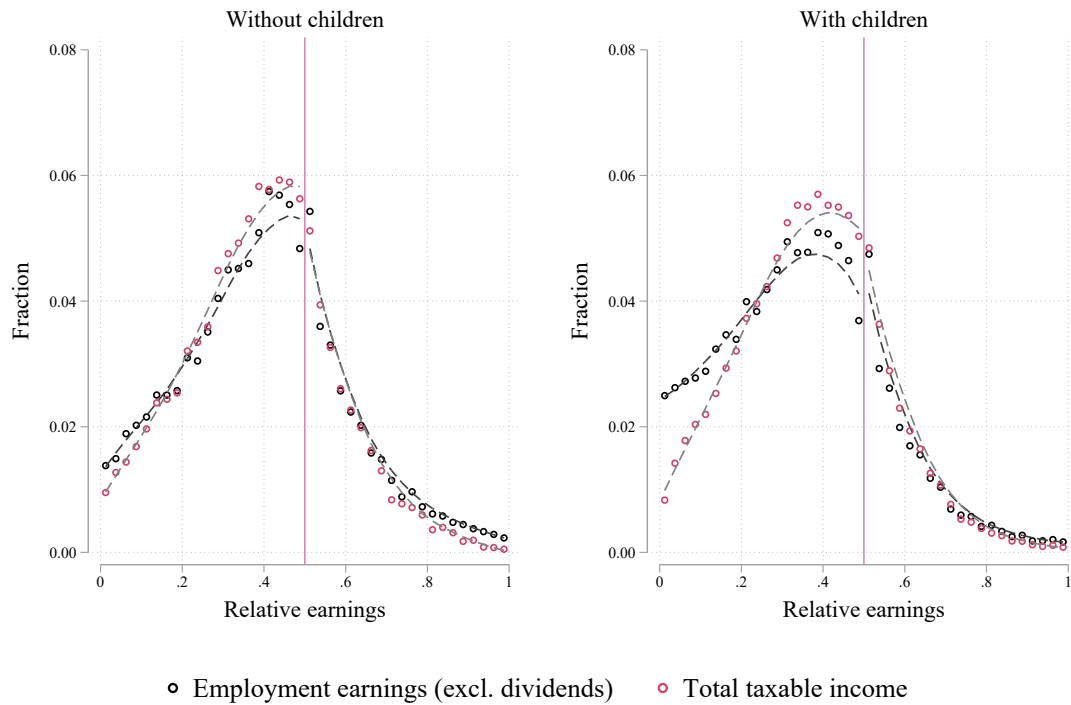


Figure A.16: Relative employment earnings (excl. dividends) of married women, ages 25-54, with and without children; SLID/CIS 1996-2016. Markers represent fraction of couples in binds of 2.5%.

Note: Estimate using public release version of the Survey of Labour and Income Dynamics (1996-2011), and Canadian Income Survey (2012-2016). Sample includes married, two-earner households where the female spouse is working aged (25-54). Employment earnings include dividend income. The male spouse's self-employment status is based on a self-reported indicator for any self-employed work in the past year. This work need not reflect the individual's main job. Each marker depicts fraction of individuals/couples within a 0.025 bin. The dashed curve depicts the LOWESS smoother; estimated separately for households on either side of the 0.5 cutoff. Cross-sectional weights have been applied.

A.8 Income-splitting: Regulatory setting and tax reforms

In Canada, shifting income to a family member for the sole purpose of lowering one's tax liability is considered tax avoidance.¹⁸⁸ Moreover, the third party reporting of wage income means that income splitting is largely confined to the domain of self-employed income (Duff, 1999).¹⁸⁹ Acting to dissuade income-splitting through child dependents, the Canadian government adjusted the tax treatment of dividends paid to children of incorporated business owners in 2000 (Bauer et al., 2015; Donnelly et al., 2000; Macnaughton and Matthews, 1999). However, limiting income splitting between spouses in a married household is more difficult, as joint ownership of a business or employment within a spouse's firm is certainly legal.¹⁹⁰ Legally, payments to a spouse for services rendered need not be at the market rate but must be 'reasonable'; leaving much room for tax optimization.

The question remains, how much can a couple save through income splitting? The answer depends on a number of factors, including the households relative and absolute incomes, as well as the progressivity of the tax structure. It also fundamentally depends on the incorporation status of the business, as a legal corporation introduces a third entity with which income can be split. For my sample, I estimate that the average household tax rate of self-employed households falls by an additional 3.5% after childbirth relative to non-self-employed households, consistent with many of the simulations provided here (see Figure 2.14). Only in cases where one spouse has no income, can the savings be as high as 10%.¹⁹¹

Consider the example of an unincorporated business. In 1990 an Ontario household with a combined income of \$61,060 (\$100,000 in 2016 CAD's) and a 30:70 income split could lower their income tax burden by \$1,200 by splitting their income 50:50; a 2.7% increase in after tax income. In 2000 the same household with a combined income of \$74,300 (\$100,000 in

¹⁸⁸Income splitting is governed by the following act:

"A person who, after the first day of August, 1917, has reduced his income by the transfer or assignment of any real or personal, movable or immovable property, to such person's wife or husband, as the case may be, or to any member of the family of such person, shall, nevertheless, be liable to be taxed as if such transfer or assignment had not been made, unless the Minister is satisfied that such transfer or assignment was not made for the purpose of evading the taxes imposed under this Act or any part thereof." 4(4) of the Income War Tax Act, S.C. 1917, c. 28

¹⁸⁹There are many ways in which households may act to lower their average tax burden by shifting non-employment income - such as investment or rental income - even in wage-employed households. I provide some evidence of such behaviour in Appendix A.7. However, this paper concerns the shifting of employment income.

¹⁹⁰In contexts where both spouses participate in a household business it may not be obvious how to allocate the taxable profits in accordance with individual effort. The simplest solution may be the only practical solution: to split the business income. This just happens to be the most tax-efficient solution too. It is for this reason, that we should not think that income-splitting among co-employed couples is purely a characteristic of individual tax jurisdictions.

¹⁹¹Such savings were only possible prior to the 2000 federal income tax reform.

2016 CAD's) would save only \$152; an increase of only 0.3% in disposable income. Similarly, in 2010 the potential savings are only 0.4% of disposable income. In contrast, a wealthier Ontario household with a combined income of \$150,000 (in 2016 CAD's) and the same 30:70 split could increase their disposable income 0.8% in 1990, 2.3% in 2000, and 1.7% in 2010. These, and all subsequent, simulated estimates are made using Milligan's (2016) CTaCS calculator.

The value of income splitting has changed dramatically during the past four decades and that these changes are highly dependent on the households absolute and relative income. This variation also arises from a number of key policy reforms.¹⁹² During the 1990's both the federal and provincial government's adjusted their income tax structures; first increasing the progressivity of the tax structure, and later reducing it.¹⁹³ Beginning in 1996 the government embarked on an extensive reform of the Canadian Pension Plan. The reform introduced a staggered increase to the contribution rate from 2.8% in 1996 to 4.95% in 2001.¹⁹⁴ CPP contributions are paid on all income above a basic exempt amount and below a maximum pensionable threshold (adjusted annually): \$3,500 and \$54,900 in 2016, respectively. Because the self-employed pay both the employer and employee component of the CPP the reform increased their contribution rate from 5.6% to 9.9%; a 4.3% increase on the marginal tax rate of all pensionable earnings.¹⁹⁵ As demonstrated by Figure A.17, this dramatically changed the progressivity of tax structure for the self-employed workers.¹⁹⁶

A third important source of variation is the federal income tax reform of 2001, which coincided with the introduction of an independent provincial tax structure.¹⁹⁷ In 2001 the federal government shifted its highest income tax rate of 29% from income above \$60,000 to income above \$100,000, lowering the marginal tax rate between \$60,000 and \$100,000 to 26%. The combined effect of this reform and the CPP reform is demonstrated by Figure A.17. Between 2000 and 2001 Canada's provinces adopted reforms to switch from a tax-on-tax structure to a fully independent income tax structure. These reforms

¹⁹²As previously discussed this paper does not extend to the pre-1988 period. As such, it does not incorporate the important 1988 federal tax reform which removed a family tax benefit that generated income splitting tax savings for married households (Crossley and Jeon, 2007).

¹⁹³Many of these reforms came in the form of surtaxes, and changes to the provincial tax-on-tax rate. For example, Ontario's tax-on-tax rate increased from 0.53 in 1990 to 0.58 in 1995, and then down to 0.395 in 1999. The Canadian provinces were unable to set their own income tax brackets until reforms in 2000.

¹⁹⁴In 1980 the contribution rate was 1.8% and increased by 0.1% on an annual basis between 1986 and 1996. Until the recent reform in 2019, the contribution rate has remained constant at 4.95%.

¹⁹⁵The 1996 reform of the CPP was required to ensure the sustainability of the fund but did not increase the pension benefits of participants. Hence, there was no corresponding increase to the lifetime savings of contributing workers, only an increase in contributions, amounting to an effective tax increase.

¹⁹⁶In certain provinces the regressivity that the CPP introduces can exceed the difference in the marginal tax rate brackets, resulting in a locally regressive tax structure and higher taxes under income splitting. However, these losses are generally very small, and a 50:50 income split generally yields the lowest tax burden.

¹⁹⁷This excludes Quebec which had an independent income tax structure prior to these reforms.

allowed provinces to set tax brackets and rates independent of the federal structure. For example, Alberta adopted a flat 10% provincial tax rate. For self-employed Albertans, this flat rate exacerbated the lack of progressivity in the underlying federal (and CPP) structure, resulting in a very flat income tax rate with regressive sections. The result was a very large drop in the potential savings from income splitting.

To demonstrate the combined effect of these various reforms I plot in Figure 2.15 the difference in after tax income of two households, one with a relative income of 50:50 and the other with a relative income of 35-65, both with a total income of \$80,000 in (2016 CAD). The difference reflects the benefit in real dollars (2016 CAD's) to income-splitting. The value increased through the 1990s but beginning in 1998 begins to fall as a result of the aforementioned reforms. The level also differs across provincial jurisdiction. The effect of Saskatchewan and Alberta's flatter tax reforms in 2001 is evident in the steep decline in 2001. As a first test, the dynamics of these reforms closely tracks the rise and fall in unincorporated, own-account self-employed workers in Canada (see Figure 2.2).

Estimating the potential savings from income splitting for an incorporated business owner is more complex. The channel through which income is parsed from the business to the household and the extent to which income is saved within the firm will matter. Parsing income through wages will trigger CPP payments on both the firm and employee side, while dividends are not considered employment income and are therefore exempt from CPP contributions.¹⁹⁸ The income tax implications of wages versus dividends are relatively minor. Wages are deducted from corporate profits and taxed at individual income tax rate, while dividends are first adjusted to account the corporate taxes that have been paid before being taxed at the individual level. The net result should be that regardless of the channel the income is taxed at the same rate.¹⁹⁹ For small business owners, the primary tax benefit to incorporation lies in long term savings. If a household can afford to leave profits within the incorporated firm it can accumulate interest at the lower (corporate) tax rate, which in the long run generates significantly higher returns. For this reason, incorporation is popular among high income self-employed professionals such as doctors and lawyers. However, I am unable to identify these savings in my data and focus instead on the savings from income splitting within an unincorporated firm.²⁰⁰ In general, this paper remains agnostic about the incorporation status of self-employment given the endogeneity of incorporation.

¹⁹⁸In Section 2.3 I make the case that excluding dividend income as a signal of taxpayer's employment status results in the underestimation of total employment. Avoiding CPP payments may be a primary reason why incorporated business owners choose to allocate profits through dividends.

¹⁹⁹Slippage can occur when at the provincial level the dividend adjustment rates are not corrected for changes in the provincial or federal corporate tax rates. There therefore may be periods where arbitrage is possible; however, these savings are typically not large.

²⁰⁰One can think of savings from income-splitting for an unincorporated business as a lower bound for incorporated businesses.

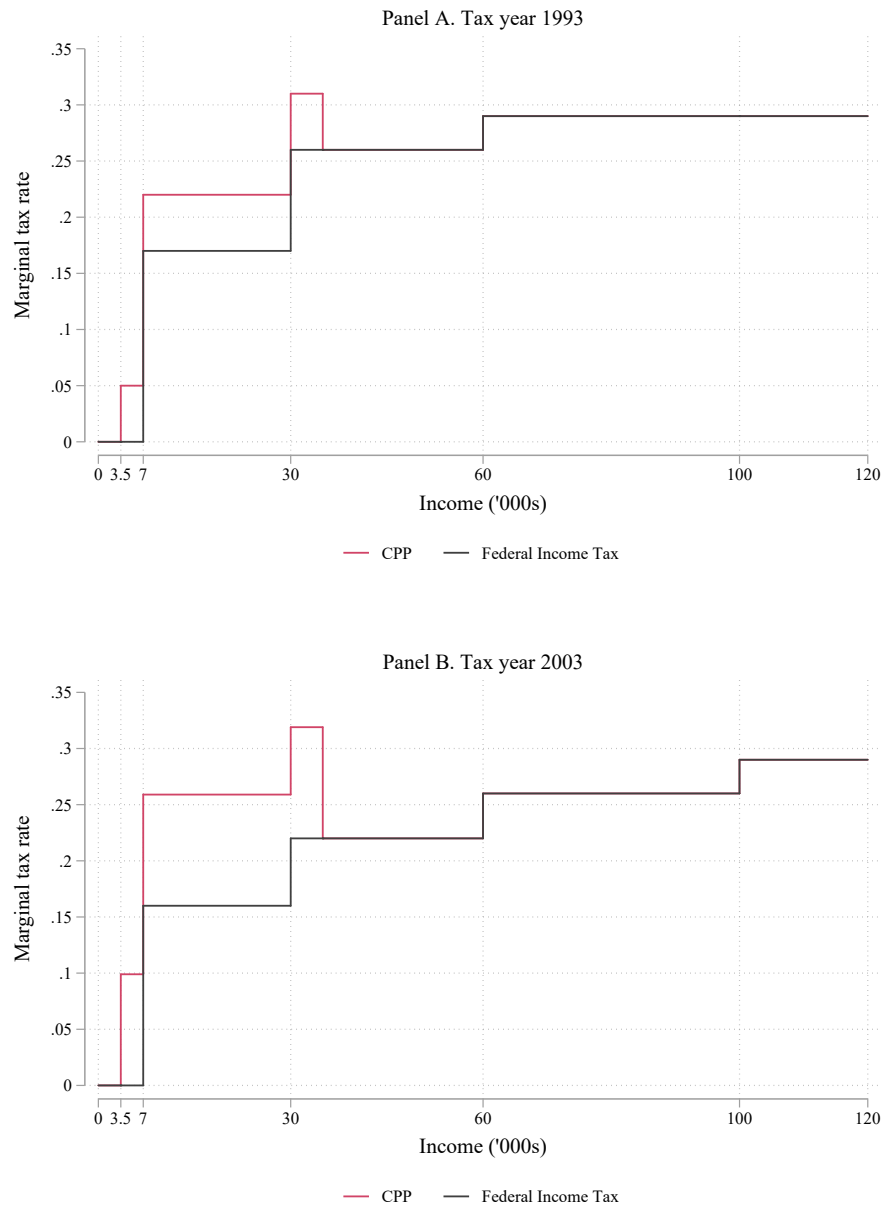


Figure A.17: Federal income tax brackets and Canadian Pension Plan contributions for 1993 and 2003; assuming self-employed filer.

B Chapter 3 Appendix

B.1 Proof of IV decomposition

Consider a case with a single endogenous variable X_1 .

$$Y = \beta_0 + \beta_1 X_1 + U$$

Suppose there exists a set of event-time instruments W that predict X_1 and collectively form the set of $2 + d$ instruments $Z = [\iota, W]$; with $d \geq 0$. The IV estimator of β_1 is given by,

$$\hat{\beta}_1 = \frac{X_1^T P_Z M_\iota P_Z Y}{X_1^T P_Z M_\iota P_Z X_1} = \frac{X_1^T (P_Z - P_\iota) Y}{X_1^T (P_Z - P_\iota) X_1} \quad (1)$$

To demonstrate the decomposition in the text, it suffices to show that the projection matrix $P_Z - P_\iota$ can be decomposed into a weighted average of other projection matrices. This proof is dependent on the fact that the excluded instrument, W , is a set of J orthogonal event-time coefficients:

$$W = [D_1, \dots, D_J]$$

First, note that $P_Z = P_{\tilde{Z}}$, where $\tilde{Z} = [D_0, D_1, \dots, D_J]$, and

$$D_0 = \iota - \sum_{j \in J} D_j$$

As $D_0 \perp D_1 \perp \dots \perp D_J$, it follows that

$$P_Z = P_{\tilde{Z}} = P_0 + P_1 + \dots + P_J = P_0 + \sum_{j \in J} P_j$$

Each of the above projection matrices reflects a block diagonal matrix with values $1/N_j$ within the block, 0 outside. Similarly, P_ι is matrix of cells $1/N$. Suppose $k = J + 1$ (the number of excluded indicators plus the constant) then there are $k^2 - k$ off-diagonal, symmetric blocks in P_ι . Given this characteristic, we can decompose P_ι as follows,

$$\begin{aligned} P_\iota &= \frac{N_0}{N} P_0 + \frac{\sum N_j}{N} P_0 + \sum_{j \in J} \frac{N_0 + N_j}{N} \left[P_{0j} - \frac{N_0}{N_0 + N_j} P_0 - \frac{N_j}{N_0 + N_j} P_j \right] \\ &= \frac{N_0}{N} P_0 + \frac{\sum N_j}{N} P_0 - \sum_{j \in J} \left[\frac{N_0}{N} P_0 + \frac{N_j}{N} P_j \right] + \sum_{j \in J} \frac{N_0 + N_j}{N} P_{0j} \end{aligned}$$

where P_0 is the projection of the pooled excluded variables: $\sum_{j \in J} D_j$. And P_{0j} is the

projection of vectors $[D_0 + D_j]$. To simplify notation, let $\sum N_j = N_J$; $N_J + N_0 = N$. It follows that

$$\begin{aligned} P_Z - P_t &= P_0 + \sum_{j \in J} P_j - \frac{N_0}{N} P_0 - \frac{N_J}{N} P_0 + \sum_{j \in J} \left[\frac{N_0}{N} P_0 + \frac{N_j}{N} P_j \right] - \sum_{j \in J} \frac{N_0 + N_j}{N} P_{0j} \\ &= \frac{N_0 + N_J}{N} \left[P_0 + \sum_{j \in J} P_j \right] - \frac{N_0}{N} P_0 - \frac{N_J}{N} P_0 + \sum_{j \in J} \left[\frac{N_0}{N} P_0 + \frac{N_j}{N} P_j \right] - \sum_{j \in J} \frac{N_0 + N_j}{N} P_{0j} \\ &= \frac{N_J}{N} \left[\sum_{j \in J} P_j - P_0 \right] + \sum_{j \in J} \frac{N_0 + N_j}{N} [P_0 + P_j - P_{0j}] \end{aligned}$$

The second line shows that you can split $P_0 + \sum_{j \in J} P_j$ into two separate terms using the weights $\frac{N_0}{N}$ and $\frac{N_J}{N}$. The third line re-organizes the terms. To the left-hand side parenthesis we can add and subtract P_0 as well as P_t .

$$\begin{aligned} P_Z - P_t &= \frac{N_J}{N} \left[(P_0 + \sum_{j \in J} P_j - P_t) - (P_0 + P_0 - P_t) \right] + \sum_{j \in J} \frac{N_0 + N_j}{N} [P_0 + P_j - P_{0j}] \\ &= \frac{N_J}{N} [(P_Z - P_t) - (P_J - P_t)] + \sum_{j \in J} \frac{N_0 + N_j}{N} [P_0 + P_j - P_{0j}] \\ \Rightarrow P_Z - P_t &= \sum_{j \in J} \frac{N_0 + N_j}{N_0} [P_0 + P_j - P_{0j}] - \frac{N_J}{N_0} (P_J - P_t) \end{aligned}$$

where P_J is the projection matrix of $[t, \sum_{j \in J} D_j]$, the IV estimator that pools all excluded event-times into a single excluded instrument. The first term contains, $P_0 + P_j - P_{0j}$ which is shown below to be the projection from the regression that excludes on event-time j . Consider projection P_{ij^c} the projection matrix of $[t, D_1, \dots, D_{j-1}, D_{j+1}, \dots, D_J]$. This projection matrix is the same as that Z , with the exception that it excludes D_j . Then,

$$\begin{aligned} P_Z - P_{ij^c} &= P_0 + \sum_{k \in J} P_k - (P_{0j} + \sum_{k \neq j} P_k) \\ &= P_0 + P_j - P_{0j} \end{aligned}$$

Thus, the projection matrix that defines the IV estimator with an event-time first stage is,

$$P_Z - P_t = \sum_{j \in J} \frac{N_0 + N_j}{N_0} [P_Z - P_{ij^c}] - \frac{N_J}{N_0} (P_J - P_t)$$

yielding the weighted estimator described in the text. Note, in the text I substitute N_0 with N_{jc} . Here, the base period (D_0) is the only non-excluded event-time dummy. However, this decomposition can be extended to consider the case where there are other orthogonal event-time dummy variables that are not included in the excluded set of instruments (e.g., pre-event event-time dummies). In this instance, N_0 is replaced with N_{jc} : the number of observations that are not included in the excluded set of event-time coefficients.

B.2 IV decomposition with exogenous covariates

Consider a case with a single endogenous variable X_1 and $k - 2$ additional covariates X_2 .

$$Y = \beta_0 + \beta_1 X_1 + X_2 \beta_2 + U$$

Suppose there exists a set of event-time instruments W that explain X_1 and collectively form the set of $k + d$ instruments $Z = [\iota, W, X_2]$; with $d \geq 0$. The IV estimator of β_1 is given by,

$$\hat{\beta}_1 = \frac{\tilde{X}_1^T P_{\tilde{Z}} Y}{\tilde{X}_1^T P_{\tilde{Z}} \tilde{X}_1} \quad (2)$$

where

$$\begin{aligned} \tilde{X}_1 &= M_{\iota 2} X_1 \\ P_{\tilde{Z}} &= M_{\iota 2} Z (Z^T M_{\iota 2} Z)^{-1} Z^T M_{\iota 2} \end{aligned}$$

$M_{\iota 2} = I - P_{\iota 2}$ is the orthogonal projection matrix of $[\iota, X_2]$. You can show that the projection $P_{\tilde{Z}}$ is equivalent to the projection $P_{\tilde{W}}$ where,

$$\tilde{W} = [\iota, W, M_{\iota w} X_2]$$

and $M_{\iota w}$ is the orthogonal projection matrix of $[\iota, W]$. As $[\iota, W] \perp M_{\iota w} X_2 = \tilde{X}_2$, it follows that

$$P_{\tilde{W}} = P_{\iota W} + P_{\tilde{X}_2}$$

Now that $P_{\iota W}$ is isolated, the decomposition in Appendix B.1 follows. As above, this allows us to decompose $P_{\iota W}$ into its orthogonal event-time components, resulting in

$$P_{\tilde{W}} = \frac{N_{J^c}}{N} P_{\tilde{X}_2} + \sum_{j \in J} \frac{N_{J^c} + N_j}{N} (P_{\iota W} - P_{\iota j}) + \frac{N_j}{N} (P_{\iota j} - P_{\iota})$$

Incorporating this decomposition in the denominator and numerator of equation 2 yields the weighted average estimator

$$\hat{\beta}^{IV} = \rho \hat{\beta}_{\tilde{X}_2}^{IV} + \sum_{j \in J} \psi_j \hat{\beta}_j^{IV} - \phi \hat{\beta}_I^{IV}$$

with weights,

$$\begin{aligned}\rho &= \frac{N_{J^c} \tilde{\Omega}_{\tilde{X}_2}}{N_{J^c} \tilde{\Omega}_{\tilde{X}_2} + \sum_{j \in J} (N_{J^c} + N_j) \tilde{\Omega}_j - N_J \tilde{\Omega}_J} \\ \psi_j &= \frac{(N_{J^c} + N_j) \tilde{\Omega}_j}{N_{J^c} \tilde{\Omega}_{\tilde{X}_2} + \sum_{j \in J} (N_{J^c} + N_j) \tilde{\Omega}_j - N_J \tilde{\Omega}_J} \\ \phi &= \frac{N_J \tilde{\Omega}_J}{N_{J^c} \tilde{\Omega}_{\tilde{X}_2} + \sum_{j \in J} (N_{J^c} + N_j) \tilde{\Omega}_j - N_J \tilde{\Omega}_J}\end{aligned}$$

As in Appendix B.1, $\tilde{\Omega}$ references the difference in the explained sum of squares from implied first-stage. However, in this instance the outcome variable in the first-stage is always $\tilde{X}_1 = M_{I2}X_1$, and not X_1 . For example, $\tilde{\Omega}_{\tilde{X}_2}$ references the increase in the explained sum of squares in a regression of \tilde{X}_1 on \tilde{X}_2 . The interpretation of $\tilde{\Omega}_j$ is the same as in B.1.

B.3 Alternative specifications

In this section I provide an additional set of IV estimates based on alternative first-stage specifications. As discussed in Section 3.4, the event-study-design (ESD) is a more flexible difference-in-difference (DiD) estimator. Naturally, this leaves the DiD as one alternative. One advantage of this more ‘constrained’ estimator is that one can, and should, control for fixed effects at the treatment unit level. Here I consider the first-stage model,

$$Z_{it} = \beta D_{it} + \delta_{pn} + \delta_t + X'_{it}\gamma + \varepsilon_{it}$$

where $D_{it} = \mathbf{1}\{s \geq 0\}$ and s is event-time (as specified in Section 3.4). Applying a DiD specification requires a control group: either ‘always treated’ or ‘never treated’. As in Table 3.7, I make use business and natural science professionals as an ‘always treated’ control. This is not ideal, given the nature of the treatment. Untreated professionals cannot incorporate, while an ‘always treated’ control can respond to changing incentives; for example, changes in the corporate-individual income tax wedge. For this reason, the parallel trends assumption need not hold. The ESD presented in models (1)-(3) of Table 3.7 includes the same control group within an ESD framework. It therefore gives evidence of pre-trends. Comparing the pre-trends from Tables 3.2 and 3.7 we see that the inclusion of a control group increases the pre-trend for medical professionals, while for legal professionals the pre-trend is ‘flatter’. This may be related to similarities between legal and accounting professionals (in the control group); in particular, their adoption of alternative limited liability partnership structures (as discussed in Section 3.3).

These first-stage results can be found in Table B.1. The pooled DiD estimate suggests a 12.2% increase in incorporation, which we know from the event-study specification is similar the increase by event-time 6 or 7 (see Table 3.2). For medical professionals, the DiD estimate of 18.9% is closer to the event-time 9 estimate under the main specification, but also event-time 6 or 7 when the control group is included (model (2) of Table 3.7). The legal profession estimate of 4.3% is also closer to the event-time 6 or 7 from Table 3.7 (see model (3)), which incorporates the control group.

With a DiD first-stage, the IV estimates are very different to those reported in Table 3.3. They point to an increase in labour supply among medical professionals and a decrease in hiring. For legal professions, the results suggest a sharp decline in hiring. These results are evidently different to those in the main specification and more work is needed to clarify these differences. The primary concern with this specification is the nature of the control group. More weight would be put on these estimates if the control group is never treated.

As a second alternative, I consider a parametric version of the ESD and include only the treated sample of medical and legal professionals (referred to here as the break-in-

trend specification). Event-time (s) is included in the model as a linear term as well as an interaction with a the above post-treatment dummy (D_{it}). Given the approximately linear adoption of incorporation after the reform (see Figure 3.7), this model should capture the impact of the reform well. Moreover, it has the advantage of reducing noise in the event-time estimates that arises from small sample sizes at the province-profession level.

$$Z_{it} = \beta_0 s + \beta_1 D_{it} + \beta_2 D_{it} \cdot s + \delta_t + \delta_n + X'_{it} \gamma + \varepsilon_{it},$$

In this specification, treatment-unit FEs (δ_{pn}) are perfectly colinear.²⁰¹ Assuming there to be no pre-trend, we can drop $\beta_0 s$ from the specification and include unit FEs (δ_{pn} or δ_p). In practice, I include only province FEs in both the DiD and restricted break-in-trend specification.

$$Z_{it} = \beta_1^n D_{it} + \beta_2^n D_{it} \cdot s + \delta_t + \delta_{pn} + X'_{it} \gamma + \varepsilon_{it},$$

These first-stage estimates from these two ‘break-in-trend’ specifications are presented in Table B.3. Given the early treatment of some provinces (e.g., Alberta), a simple linear trend in event-time risks underestimating the initial rate of adoption if incorporation reaches a plateau. Evidence of such a plateau is seen in Alberta for which we only observe $s > 15$ in this sample. Incorporation in Alberta is higher than any other province and does not increase as dramatically over the sample period. I therefore include an additional break-in-trend at event-time 10 to ensure that the break at 0 identifies the initial take up and the results are more comparable with the ESD first-stage that excludes event-times 2 through 10.²⁰² The pooled estimate for β_0 is -0.0035 (model (1), Table B.3), a remarkably flat slope in keeping with the flat pre-trend (see model (4) in Table 3.2 for a comparable ESD specification). The model suggests a post reform event-time slope of 0.021, suggesting a 21% increase in incorporation over ten years. This is just larger than the event-study estimate of 20.1% for event-time 10. When you exclude the pre-trend and include province FEs, the estimates remain fairly similar. As with the ESD specification, when you assume a zero pre-trend and include province FEs the model estimates a marginally faster take-up by medical professionals and marginally slower take up by legal professionals.

The IV estimates for hours of work suggest a decline in medical hours, when province FEs are excluded. Including province FEs finds no significant relationship. There is a negative relationship with part-time status, with and without FEs for medical professionals. This opposes the hours result (which is not robust to the inclusion of FEs, but supports the

²⁰¹This helps to demonstrate the issue of underidentification in the more flexible ESD model (see discussion in [Borusyak and Jaravel, 2018](#)).

²⁰²Only the initial break-in-trend ($D_{it} \cdot s$) and post treatment dummy (D_{it}) are excluded from the second stage estimation. The LATE should therefore correspond to take-up during the first 10 years after treatment.

DiD results. Finally, the hiring result is negative without FEs, but positive with for medical professionals with FEs. The positive result in in model (5) for health professionals is larger and highly significant for the female subsample ($\sim 20\%$).²⁰³ None of the IV coefficients are significant for legal professionals under this parametric specification.

Taken together, these results do not lend strong support to those in the main specification. More work is needed to find a specification that is well suited to this reform setting. In future work I intend to explore specifications that take into account the staggered adoption of Limited Liability Partnerships as well as the potential interaction of treatment and the corporate tax rate. The former may help account for the weaker first-stage among lawyers (and even accountants), while the latter may assist in controlling for pre-trends in specifications with an ‘always treated’ control group.

Table B.1: First stage: generalized difference-in-difference, by occupation

	(1) Pooled	(2) Health	(3) Legal
Post treatment	0.122*** (0.00291)	0.189*** (0.00359)	0.0434*** (0.00426)
Corporate rate	-0.678*** (0.0476)	-0.799*** (0.0532)	-0.354*** (0.0593)
Province FE	Yes	Yes	Yes
N	365078	301653	255581

Note: Estimated using Canadian Labour Force Survey (1989-2019). Sample includes non-education self-employed professionals, aged 25-59. The table reports the post treatment coefficient from the specified difference-in-difference model, as well as the coefficient on the corporate tax rate. Models are estimated without sampling weights and include covariates: sex, age, education, marital status, age of youngest child, industry, and corporate tax rate. As standard, all models include year FEs. Estimates that pool professionals also include occupation FEs.

²⁰³Results not shown here.

Table B.2: Labour supply and hiring: IV results based on DiD first stage

	(1) Pooled	(2) Health	(3) Legal
Panel A. Log of weekly hours			
Incorporated	0.0625*** (0.0242)	0.0553*** (0.0189)	0.140 (0.103)
Province FE	Yes	Yes	Yes
<i>N</i>	365078	301653	255581
C-D Wald F-stat	1763.7	2776.2	103.5
Panel B. Part-time status			
Incorporated	-0.0536*** (0.0171)	-0.0636*** (0.0135)	-0.0407 (0.0718)
Province FE	Yes	Yes	Yes
<i>N</i>	365078	301653	255581
C-D Wald F-stat	1763.7	2776.2	103.5
Panel C. Hiring status			
Incorporated	-0.114*** (0.0235)	-0.0394** (0.0182)	-0.925*** (0.147)
Province FE	Yes	Yes	Yes
<i>N</i>	365078	301653	255581
C-D Wald F-stat	1763.7	2776.2	103.5

Note: Estimated using Canadian Labour Force Survey (1989-2019). Sample includes health and legal professionals, aged 25-59. Models are estimated without sampling weights and include covariates listed in Table B.1. See Table B.1 for corresponding first-stage estimates.

Table B.3: First stage: break-in-trend, by occupation

	With pre-trend			Without pre-trend		
	(1) Pooled	(2) Health	(3) Legal	(4) Pooled	(5) Health	(6) Legal
Event-time	-0.00350*** (0.000509)	-0.00867*** (0.000721)	0.00647*** (0.000763)			
Post treatment	-0.00156 (0.00563)	-0.0126 (0.00774)	-0.00578 (0.00907)	-0.0144*** (0.00517)	-0.00519 (0.00787)	0.00466 (0.00908)
Event-time x Post	0.0248*** (0.000897)	0.0379*** (0.00123)	0.00563*** (0.00139)	0.0220*** (0.000771)	0.0405*** (0.00129)	0.00471*** (0.00140)
>10 yrs from treatment	0.213*** (0.00646)	0.329*** (0.00809)	0.0528*** (0.0114)	0.216*** (0.00647)	0.344*** (0.00813)	0.0405*** (0.0117)
Event-time x >10 yrs	-0.0166*** (0.000719)	-0.0262*** (0.000926)	-0.00563*** (0.00123)	-0.0167*** (0.000719)	-0.0274*** (0.000930)	-0.00405*** (0.00124)
Corporate rate	-0.0965 (0.0630)	-0.360*** (0.0817)	0.476*** (0.104)	-0.156** (0.0651)	0.0135 (0.0921)	0.284*** (0.109)
Province FE	No	No	No	Yes	Yes	Yes
N	172922	109497	63425	172922	109497	63425

Note: Estimated using Canadian Labour Force Survey (1989-2019). Sample includes health and legal professionals, aged 25-59. The table reports event-time trend and break-in-trend coefficients from the specified break-in-trend model, as well as the coefficient on the corporate tax rate. Models are estimated without sampling weights and include covariates listed in Table B.1.

Table B.4: Labour supply and hiring: IV results based on break-in-trend first stage

	With pre-trend			Without pre-trend		
	(1) Pooled	(2) Health	(3) Legal	(4) Pooled	(5) Health	(6) Legal
Panel A. Log of weekly hours						
Incorporated	-0.147*** (0.0371)	-0.0975*** (0.0308)	-0.403 (0.303)	0.0517 (0.0333)	-0.00453 (0.0297)	-0.288 (0.331)
Province FE	No	No	No	Yes	Yes	Yes
<i>N</i>	172922	109497	63425	172922	109497	63425
C-D Wald F-stat	386.9	477.1	8.231	450.9	495.4	6.419
Sargan P-value	0.312	0.020	0.649	0.085	0.904	0.709
Panel B. Part-time status						
Incorporated	-0.0440* (0.0258)	-0.0534** (0.0225)	-0.102 (0.186)	-0.0772*** (0.0239)	-0.0487** (0.0221)	-0.347 (0.223)
Province FE	No	No	No	Yes	Yes	Yes
<i>N</i>	172922	109497	63425	172922	109497	63425
C-D Wald F-stat	386.9	477.1	8.231	450.9	495.4	6.419
Sargan P-value	0.272	0.219	0.391	0.126	0.190	0.603
Panel C. Hiring status						
Incorporated	-0.0744** (0.0358)	-0.0817** (0.0327)	-0.256 (0.257)	0.331*** (0.0317)	0.0527* (0.0310)	-0.133 (0.277)
Province FE	No	No	No	Yes	Yes	Yes
<i>N</i>	172922	109497	63425	172922	109497	63425
C-D Wald F-stat	386.9	477.1	8.231	450.9	495.4	6.419
Sargan P-value	0.000	0.000	0.087	0.000	0.000	0.153

Note: Estimated using Canadian Labour Force Survey (1989-2019). Sample includes health and legal professionals, aged 25-59. Models are estimated without sampling weights and include covariates listed in Table B.1. See Table B.3 for corresponding first-stage estimates.

C Chapter 4 Appendix

C.1 Right-to-Work results using NLRB data

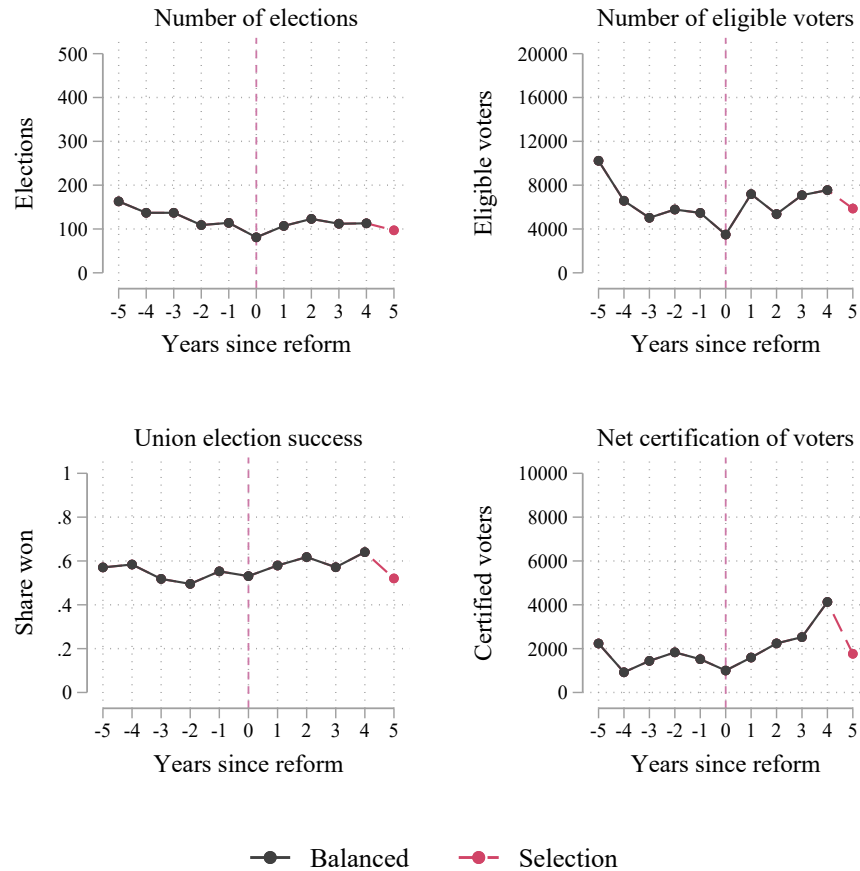


Figure C.1: Average of NLRB election data by event-time: Indiana, Michigan, and Wisconsin

Note: The figure plots aggregate election data by event-time of RTW law passing in the three states of Indiana, Michigan, and Wisconsin. As public and private sector unions are not identified in the data, we set the year of treatment to 2015 for Wisconsin. The final period is denoted by a separate “selection” label, because Wisconsin is not counted in the statistics.

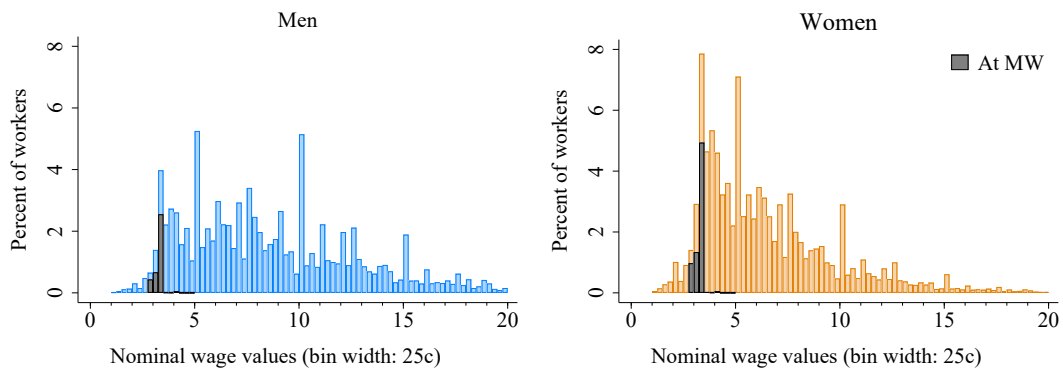
C.2 Issues in the estimation of the direct and spillover effects of the minimum wage for distribution regressions

Figure C.2 illustrates the potential importance of heaping issues by presenting histograms of the lower part of the distribution of wages. The figures differentiate the mass (in gray) at actual minimum wages from mass that arises due to heaping issues. As can be gleaned from the figure, mass at minimum wages is substantially more important in the period 1979-88 than in the other periods.

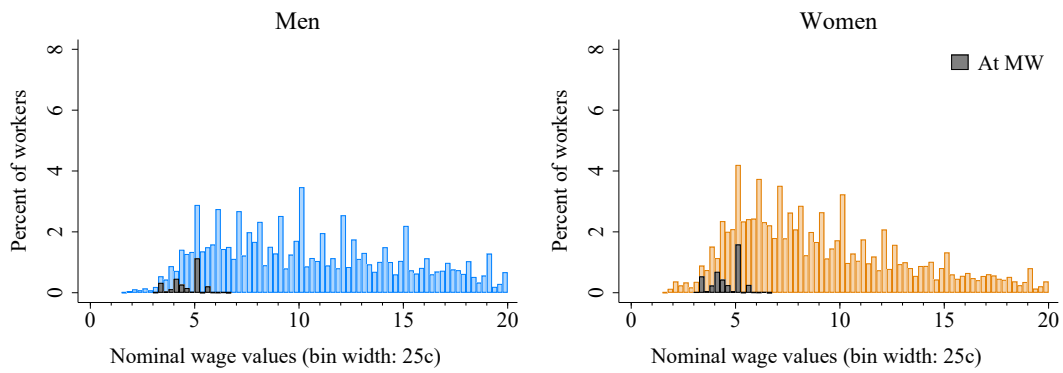
Figure C.3 illustrates the estimated wage distributions for 1979-88 with and without the minimum wage. The two distributions are reported in Figure C.3, Panel A, correspond to the values of the unconditional probability \hat{Q}_t^k , where $\hat{Q}_t^k = \frac{1}{N_t} \sum_i \hat{Q}_{ist}^k$, and the predicted probabilities $\hat{Q}_t^{k,c}$ recentered around the median value of the minimum wage. As explained in the text, $\hat{Q}_{ist}^{k,c} = \hat{p}_{ist}^{k,c} - \hat{p}_{ist}^{k+1,c}$ and $\hat{Q}_t^{k,c}$ is defined as $\hat{Q}_t^{k,c} = \frac{1}{N_t} \sum_i \hat{Q}_{ist}^{k,c}$. Panels B and C of Figure A2 illustrates the impact of the measurement error correction on the wage distribution for the 1979-88 period.

Figure C.4 compare our estimated spillover effects where we do control for rounding off with the set of heaping dummies L_{kst}^p to what we would find without controlling for these dummies. The results of this comparison for 1979-88 are reported in Figure C.4.

A. 1979-1988



B. 1988-2000



C. 2000-2017

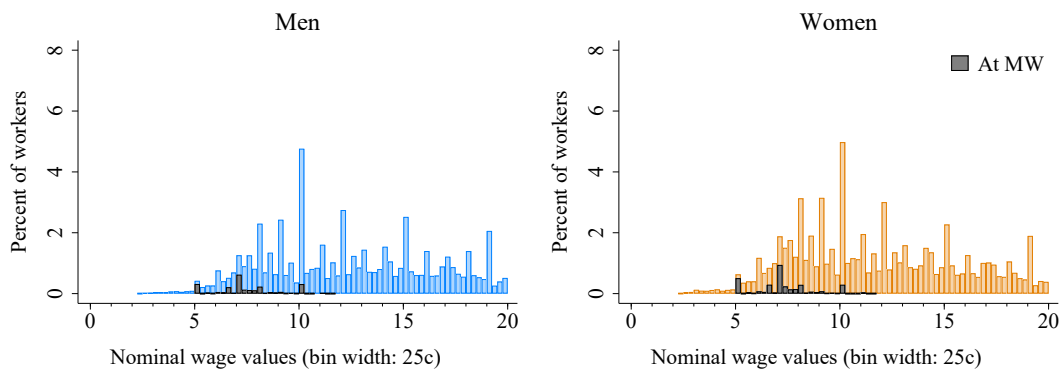


Figure C.2: Histogram of nominal wage bins

Note: The histogram uses 25 cents wage bins (starting from 0) and displays the distribution only up to \$20/ hrs but remains proportional to total distribution. The darker wage bins show the share of workers earning their state's minimum wage. This fraction is calculated by identifying workers who report an hourly wage within 10 cents of their state's minimum wage.

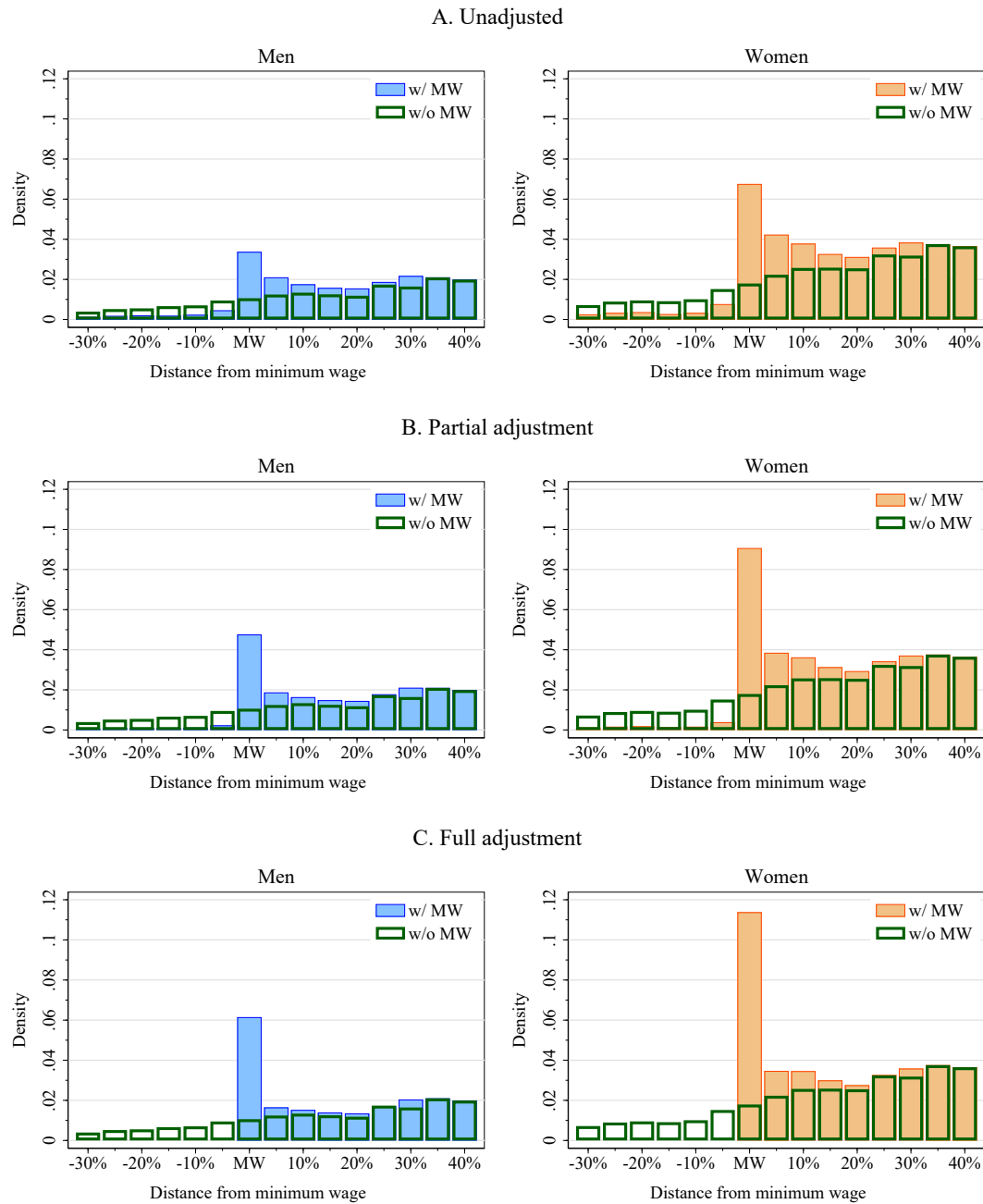


Figure C.3: Estimates of minimum wages impact: 1979-88

Note: The bars depict the predicted density of each wage bin around the minimum wage, given the estimates of the model. In each panel, the counterfactual probabilities without a minimum wage (“w/o MW”) are obtained by setting the minimum wage coefficients to zero (see discussion in Section 4.4.2). Panels B and C reflect the partial and full adjustments made to account for measurement error discussed in the text (Section 4.4.3).

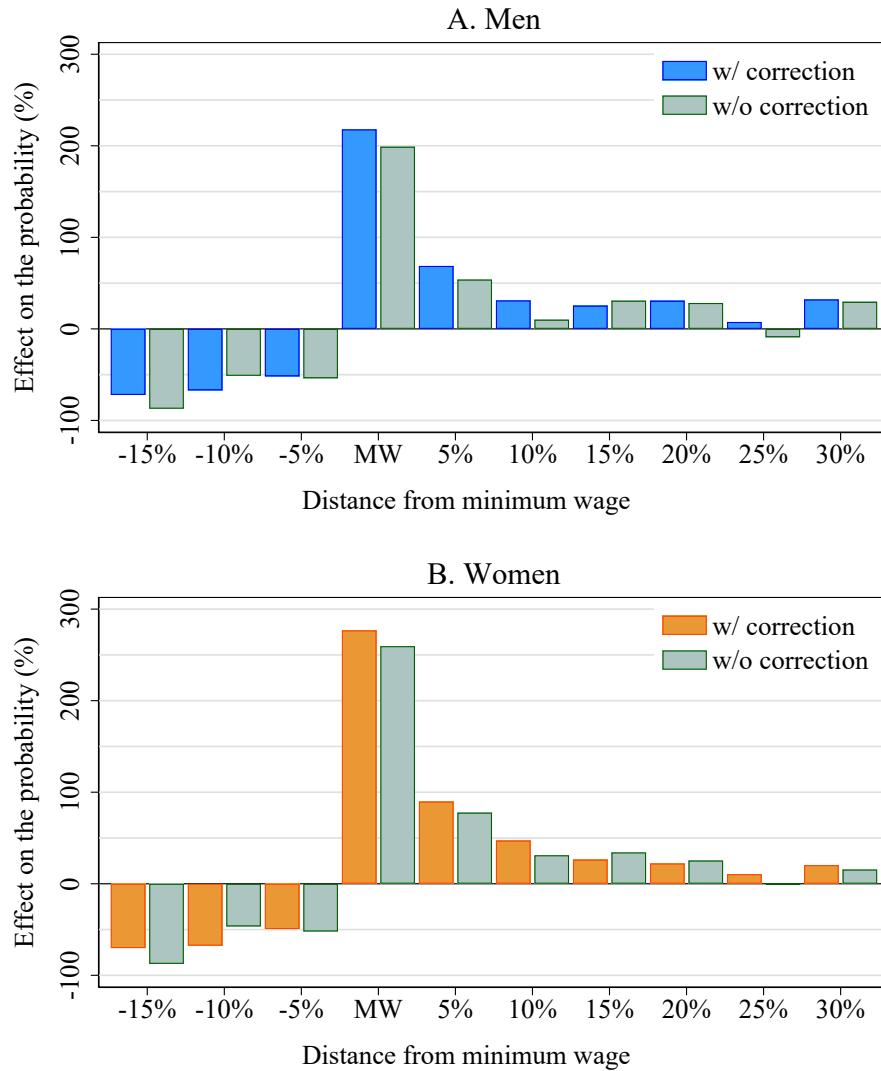


Figure C.4: Estimates of minimum wages impact: 1979-88

Note: As in Figure 4.6, each bar represents the marginal effect of the minimum wage, shown here for the 1979-1988 period. Each panel includes two sets of estimates: with and without the correction for heaping. The model “with correction” includes the controls for bunching at integer, nominal wage-values (as in Figure 4.6). In contrast, the set of marginal effects “without correction” is constructed from a separate model estimated without these additional controls. See discussion in 4.4.3.

C.3 An illustrative example of the identification of minimum wage effects

This appendix is offered as a complement to section 4.2.3. Here, we illustrate how a new minimum wage affects wages below (loss of mass), at (spike) and above (spillover) the new minimum. Importantly, we show how these changes to the wage distribution map into the parameters φ_m of the model. Consider a latent normal wage distribution in Figure C.5a (blue line). We now add a minimum wage (red line) that creates a large spike at the minimum, adds some mass slightly above the minimum wage (spillover effects), and dramatically reduces the probability of being at values below the minimum wage.

Figure C.5a shows that the probability of being in the “spillover zone” just above the minimum wage increases from A to A+C, while the probability of being at the spike increases from B to B+D. In this simple example, the parameters φ_1 and φ_0 are the horizontal values (illustrated by arrows in Figure C.5b) by which the cutoff points have to be moved to increase the two probabilities by an amount of C and D, respectively.

Next, Figure C.5c and B1d illustrate a case with two states that differ in terms of mean wages. If we use the dummy variable X in equation (4.2) to indicate if an observation comes from the high-wage state, the parameter β will capture the mean wage differences between the two states. The three key parameters to be estimated in this example are β (the difference in means) and the minimum wage parameters φ_1 and φ_0 . As discussed at the beginning of this section, these parameters are jointly estimated in our estimation approach, while corresponding parameters are estimated in two separate steps in [Lee \(1999\)](#) and [Autor et al. \(2016\)](#).²⁰⁴

To better understand how φ_1 and φ_0 are estimated in the two states example, Appendix Figure C.5d shows the recentered densities obtained using the parameter – or adjustment factor – β . The recentering clearly shows how the same federal minimum wage bites at different points of the distribution in the two states. A precisely similar graph would be obtained if the two states had the same latent wage distribution but different state wage minimum wages. Thus, from an identification perspective, it does not matter whether the variation in the relative minimum wage is driven by differences in mean wages across states (as in [Lee, 1999](#)), or difference in state minimum wages (as in [Autor et al., 2016](#)).²⁰⁵ The parameters φ_1 and φ_0 correspond again to horizontal moves in cutoff values (arrows in Appendix Figure C.5d) required to fit the change in probabilities induced by the minimum

²⁰⁴The model parameters are quite different in the two approaches since we are modeling the probability distribution, while [Lee \(1999\)](#) and [Autor et al. \(2016\)](#) are modeling quantiles of the wage distribution. The β parameters in equation (4.2) are, nonetheless, closely connected to the “first-step” median used in these two papers to compute the relative value of the minimum wage. The measurement error linked to plugging in estimates of the medians does not apply given that we are jointly estimating similar centrality parameters and minimum wage effects.

²⁰⁵As long as the state-fixed effect is constant, or changes smoothly with the wage-distribution.

wage. Interestingly, the same horizontal shift has a larger impact on probabilities when the minimum wage is relatively higher up in the distribution (low-wage state case in Figure C.5d). This convenient property is linked to the well-known fact that marginal effects in a probit model are directly proportional to the density at the point where the marginal effects are computed. As in Lee (1999), the relative bite of the minimum wage—its distance relative to the median—also plays a central role in the estimation in the distribution regression model.

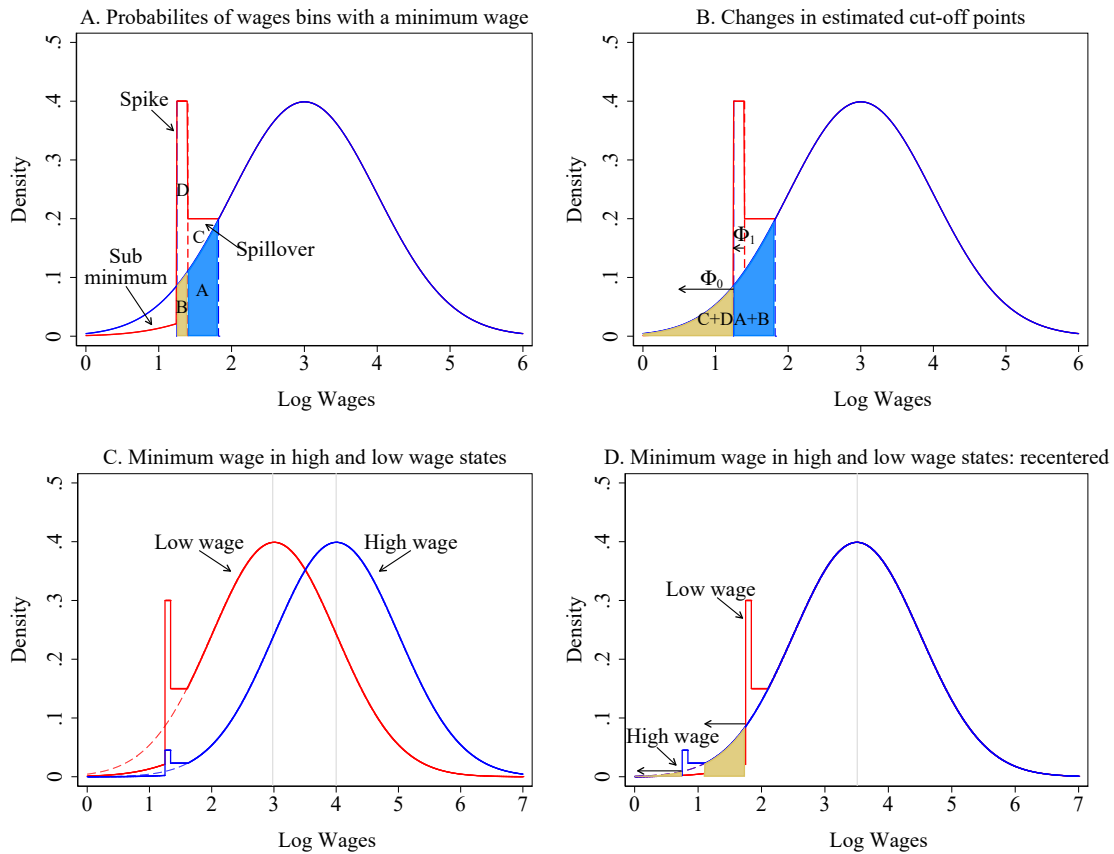


Figure C.5: Illustrative example of the identification of minimum wages effects

C.4 RIF-regression estimates of the impact of unionization

As discussed in Section 4.3, the estimates of union threat effects used in the decompositions are based on estimates of the distribution regressions where the unionization rate at the state-industry-year level is included as an additional regressor. Here we present more straightforward estimates based on OLS and RIF-regression models where it is easier to estimate the effects of the unionization rate at different points of the distribution. The results from these simple regressions are reported in Figure C.6.

To compare our results with earlier studies, we report in the first panel estimates of the effect of the union status on wages. The OLS estimates yield the typical union wage premium, while the RIF-regression coefficients indicate how the union effect varies at different point of the distribution.²⁰⁶ We use the same set of covariates but estimate the model over the entire 1979-2017 period. These results are comparable to those of the Figure 4.7 which provides marginal effects of a 1 %-point increase in union coverage (“shift share”) and a 1 %-point increase the industry-state unionization rate (“union threat”).

Consistent with the existing literature, Panel A of Figure C.6 shows that the union wage premium (horizontal red line) is about 20% for men, and a bit smaller for women.²⁰⁷ As in [Firpo et al. \(2009\)](#), the union effect estimates obtained using RIF-regressions are hump-shaped. For both men and women, they peak around the middle of the distribution, and steadily decline in the upper part of the distribution.

Intuitively, the pattern of union wage effects – positive on average but declining in the upper part of the distribution – is consistent with other evidence on the effect of unions on the wage structure. For instance, Card (1996) shows that the union wage premium is positive on average but declines over the skill distribution.

It is not as intuitive, however, to see why the RIF-regression estimates first grow before reaching a peak around the middle of the distribution. Part of the story is that changes in the rate of unionization have little impact at the bottom of the distribution where wages mostly depend on the minimum wage. Another part of the story is that very few workers are unionized at the bottom of the distribution. The issue is discussed in more detail using an example with uniform distributions in Appendix C.5. Note that the hump-shaped pattern of RIF-regression coefficients has important implications on how de-unionization affects the shape of the wage distribution. Panel A of Figure C.6 indeed indicates that unionization

²⁰⁶As discussed in [Firpo et al. \(2009\)](#), RIF-regression estimates can be interpreted as the impact of a small change in the probability of unionization on the unconditional quantiles of the wage distribution. As such, RIF-regressions are one among several possible ways of computing the counterfactual distribution obtained by changing the probability of unionization. The alternative approach used in Section 4.4.4 consists of reweighting the data to slightly increase the fraction of union workers (as in DFL), and see how it affects the various wage quantiles.

²⁰⁷For instance, [Card et al. \(2020\)](#) find a union wage premium of 0.16 for men and 0.09 for women in 2015.

substantially reduces the 90-50 gap, but also increases the 50-10 gap. Interestingly, DFL reach a similar conclusion using a reweighting approach, as we do with the distribution regression method (see Section 4.4.4).

Panel B shows corresponding estimates of the effect of the state-industry-year unionization rate on the wages of non-union workers. Interestingly, in the case of men the shape and magnitude of the estimated effects are qualitatively similar to those for the union status reported in Panel A. In the case of women, the OLS estimate is substantially smaller, and the RIF-regression estimates are a bit unstable across the various percentiles of the distribution. This is consistent with the main model, where we find that the threat effect is much smaller. Moreover, here we pool the three periods. Given the rapidly changing composition of female labour supply over this period, it may be preferable to estimate separate models.

Taken together, the results reported in Figure C.6 support the view that the threat of unionization has a positive effect on the wages of non-union workers. Although the shape of the RIF-regression coefficients varies across the specifications reported in Panels B, the estimates tend to be small and often negative at the top of the distribution. As discussed in Section 4.3, this supports the view that declining unionization rates (or success rates of union elections) capture declining threat effects instead of spurious state-industry shocks that both reduce wages and unionization rates.

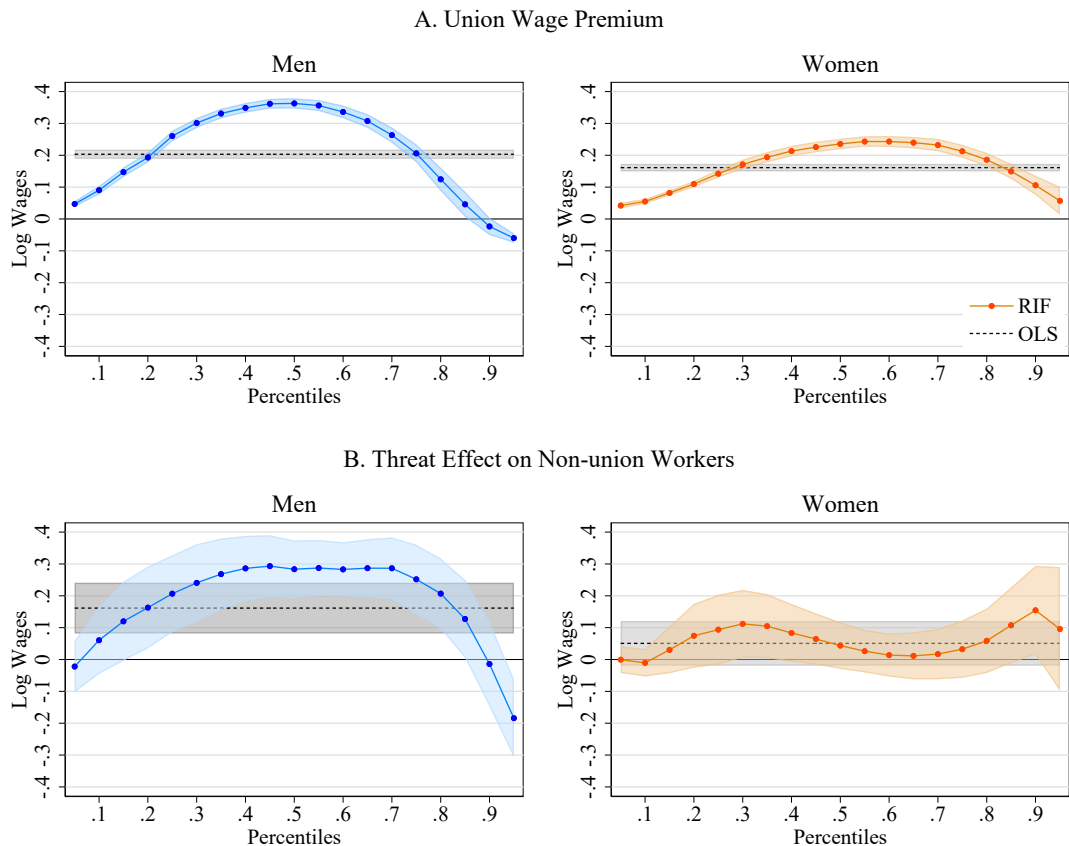


Figure C.6: OLS and RIF regression model: Spillover effects based on unionization rates

Note: Each figure plots the coefficient and 95% confidence interval for the OLS and RIF regression model. Panel A corresponds to the direct union-wage gap using the sample of non-allocated wages of covered and uncovered salary workers. Panel B corresponds to the coefficient on the unionization rate at the industry-state level (3 year moving average) using the sample of uncovered salary workers. All models pool data from 1979-2017, for which union coverage and non-allocated wages are observed; this excludes data from 1980-1982, 1994, and part of 1995. The density and RIF are estimated separately by year, while the RIF regression is estimated once, pooling all available years of data. Each model includes state and industry FE, as well as state- and industry-specific trends, as in the larger distribution model. Additional covariates include our standard set of controls: year FE, marital status, part-time status, experience (quartic), education, education-experience bins, public sector, part-time, occupation, and CMSA location.

C.5 Understanding the hump shaped effect of unionization on the wage distribution

In this appendix, we discuss a simple example to illustrate why a change in the rate of unionization is likely to have a “hump shape” or “inverse U-shape” impact on wage quantiles. The hump shape effect has been documented empirically using RIF-regressions (e.g. Appendix Figure C1 or Firpo, Fortin, and Lemieux, 2009) and distribution regressions (Figure 4.7).

For the sake of simplicity, consider a case where non-union wages follow a uniform distribution between zero and one ($Y \sim U(0, 1)$). Union wages follow a $U(0.6, 0.8)$ distribution, which has a higher mean but lower variance than the non-union distribution. The two distributions are illustrated in Figure C.7a.

Now consider a counterfactual experiment where the unionization rate increases from 0.2 to 0.3. Figure C.7b shows the wage densities for all workers combined, while C.7c shows the corresponding cumulative distribution functions (CDF). Raising the rate of unionization increases the mass in the upper middle of the distribution and reduces the mass in the two tails of the distribution. While this reduces overall wage dispersion (the variance goes from 0.074 to 0.068), the impact is uneven at different points of the wage distribution.²⁰⁸ To see this, recall that the effect of increasing the unionization rate on wage quantiles is the horizontal distance between the two CDFs plotted in Figure C.7c. The effect on wage quantiles is zero at the very bottom of the distribution, but grows linearly until the 40th percentile. The effect of changing the unionization rate on wage quantiles then starts declining before turning negative around the 80th percentile. This non-monotonic effect of the unionization rate on wage quantiles is illustrated in Figure C.7d that plots the (smoothed) change in wage quantiles over the whole distribution, and exhibits the hump-shaped feature discussed above.

The intuition for why unionization increases wage quantiles at the bottom of the distribution, but reduces wage quantiles at the top is straightforward. Increasing the rate of unionization shrinks the wage distribution towards the upper middle (0.6-0.8 range in Figure C.7b), which pulls up wage quantiles at the bottom and pulls down wage quantiles at the top. What is not as intuitive is why the effect first grows at the bottom of the distribution before declining later on. In the case of the uniform distribution, the lowest quantiles cannot move much in response to a change in the rate of unionization as they are “pinned down” at the lower bound of the distribution (0 in this example). Likewise, a binding minimum wage that creates a sharp lower bound would generate the same phenomena. For example,

²⁰⁸The overall variance can be computed using the well-known analysis of variance formula $Var(Y) = \bar{U} \cdot V^U + (1 - \bar{U}) \cdot V^{NU} + \bar{U} \cdot (1 - \bar{U}) \cdot (\mu^U - \mu^{NU})^2$ where the mean and variance of wages in the union and non-union sectors are, $V^U = \frac{1}{12 \cdot 25}$, $V^{NU} = \frac{1}{12}$, $\mu^U = .7$, and $\mu^{NU} = .5$, respectively.

if 10 percent of non-union workers are bunched at the minimum wage, the 0th to the 7th (8th) quantiles will be equal to the minimum wage when the unionization rate is 30% (20%). As a result, wage quantiles up to the 7th quantile won't change when the unionization rate increases, while quantiles slightly higher up will increase for the reason discussed above (overall distribution shrinking towards the upper middle).

As it turns out, other distributions like the normal distribution also yield the hump-shaped curve illustrated in the case of the uniform distribution. To see this, note that in Figure C.7c, vertical distance between the two CDFs (20% and 30% unionization rates) is a linear function of the wage. The horizontal distance is equal to the vertical distance divided by the slope of the CDF (the wage density $f(Y)$) evaluated at this point. Thus, the effect is increasing in Y as long as the derivative of $Y/f(Y)$ with respect to Y is positive. This trivially holds in the case of the uniform distribution since $f(Y)$ is a constant, and holds for more general distributions as long as $f(Y)$ is not growing "too fast" as a function of Y at the bottom end of the distribution.

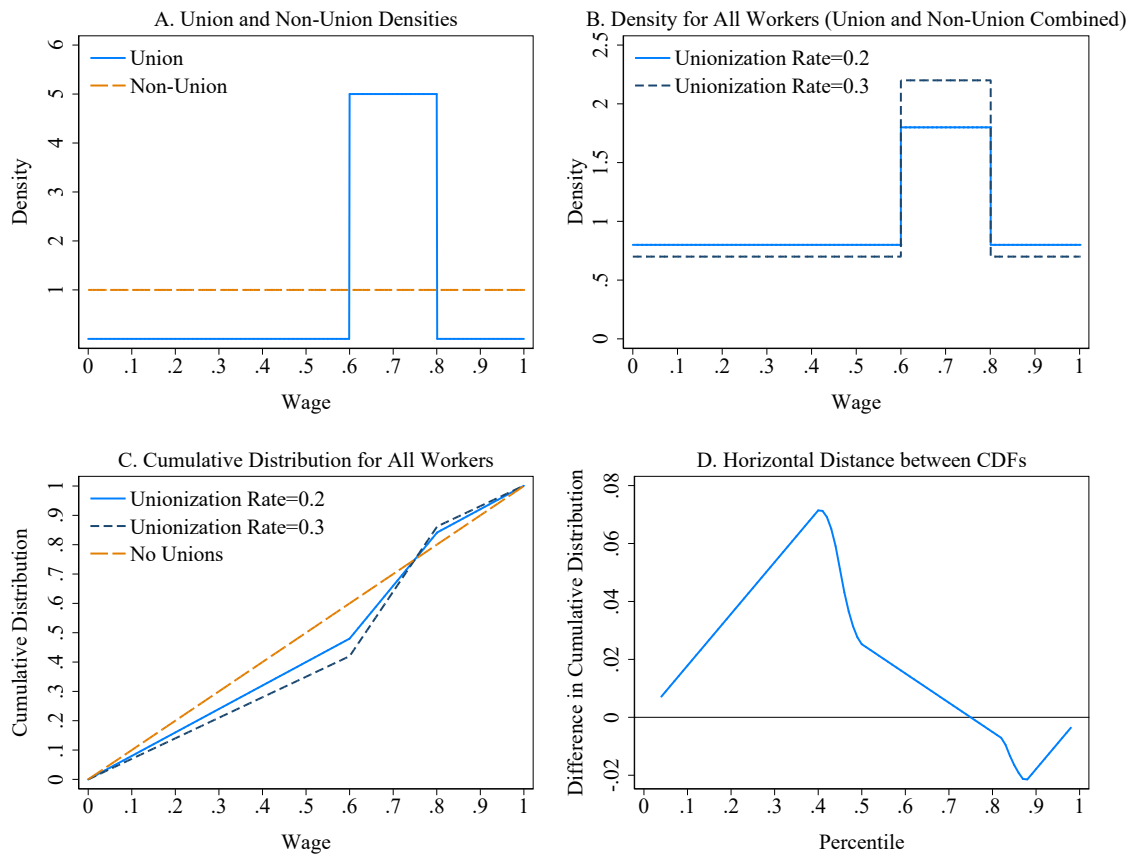


Figure C.7: Understanding the hump-shaped effects of unionization