Essays in Development Economics

by

Arkadev Ghosh

MA (Hons)., The University of Edinburgh, 2015
M.Sc., The London School of Economics and Political Science, 2016

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 2022

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 Development Economics

submitted by **Arkadev Ghosh**

in partial fulfillment of the requirements for the degree of **Doctor of Philosophy** in **Economics**.

Examining Committee:

Siwan Anderson, Professor, Economics, UBC

Co-supervisor

Munir Squires, Assistant Professor, Economics, UBC

Co-supervisor

Claudio Ferraz, Professor, Economics, UBC

University Examiner

Arjun Chowdhury, Assistant Professor, Political Science, UBC

University Examiner

Farzana Afridi, Professor, Economics, Indian Statistical Institute, Delhi Center

External Examiner

Additional Supervisory Committee Members

Patrick Francois, Professor, Economics, UBC

Supervisory Committee Member

Jamie McCasland, Assistant Professor, Economics, UBC

Supervisory Committee Member

Matt Lowe, Assistant Professor, Economics, UBC

Supervisory Committee Member

Abstract

The first chapter implements a field experiment in India to understand whether the effects of religious diversity on team productivity and worker attitudes depend on a firm's production technology. I randomly assigned Hindu and Muslim workers at a manufacturing plant in West Bengal to religiously mixed or homogeneous teams. Production tasks are categorized as high- or low-dependency based on the degree of continuous coordination required for production. I find that mixed teams are less productive than homogeneous teams in high-dependency tasks, but this effect attenuates completely in four months. In low-dependency tasks, diversity does not affect productivity. Despite lowering short-run productivity, mixing improves out-group attitudes for Hindu workers in high-dependency tasks – but there are little or no effects in low-dependency tasks. Overall, this pattern of results suggests that technology that incentivizes individuals to learn to work together is important in overcoming existing intergroup differences – and leads to improved relations and team performance.

The second chapter shows that close-kin marriage, by sustaining tightly-knit family structures, impedes development. We use US state-level bans on cousin marriage for identification. Our measure of cousin marriage comes from the excess frequency of same-surname marriages, a method borrowed from population genetics that we apply to millions of marriage records from 1800 to 1940. We show that state bans on first-cousin marriage did reduce rates of in-marriage, and that affected descendants therefore have higher incomes and more schooling. Our results are consistent with this effect being driven by weakening family ties rather than a genetic channel.

The third chapter studies mining activity in Indian states and districts between 1960-2015, and finds that mining intensity gradually decreases as elections approach. This pattern is manifested in output, mining accidents, and mineral licensing. The magnitude of these cycles are determined primarily by two factors: electoral competition and the intensity of Naxalite conflict, an ongoing left-wing insurgency

against the Indian government. While mining fatalities are costly during elections, I show that cycles in conflict prone areas are exacerbated in order to minimize the tax base of rebel groups, who thrive on extortion of mining revenues and target elections with violence.

Lay Summary

My dissertation consists of three distinct chapters in development economics. The first chapter implements a field experiment to understand whether the effects of religious diversity on team production and worker attitudes depend on a firm's production technology. I find that in high-coordination tasks, diversity initially leads to lower productivity. But this effect dissipates over time and contact in these tasks also leads to positive attitude change towards non-coreligionists. These effects are not present in low-coordination tasks. The second chapter studies the effects of close-kin marriage on economic development outcomes such as income, schooling and female labor force participation. Using state bans on cousin marriage in the U.S., we show that a reduction in first cousin marriages led to an improvement in these outcomes. The third chapter documents political business cycles in mining activity in India and explores why in contrast to other economic activity, mineral extraction is minimized in election years.

Preface

Chapters 2 and 4 are pieces of original, unpublished and independent work. Chapter 2 involves human participants. The protocol for the study was approved by UBC BREB with approval certificate number H19-00729. Chapter 3 is joint work with Professor Munir Squires (UBC) and Professor Sam Hwang (UBC). I have been involved throughout each stage of research: collecting data, conceptualizing the research design as well as conducting the empirical analysis.

Table of Contents

Appendices

List of Tables

List of Figures

Acknowledgements

I would like to express my deepest gratitude to Siwan Anderson, Patrick Francois, Matt Lowe, Jamie McCasland and Munir Squires for their incredible support, advice and patience throughout my time at UBC. I am greatly indebted to Patrick Baylis, Victor Couture, Claudio Ferraz, David Green, Sam Hwang, Ashok Kotwal, Thomas Lemieux, Nathan Nunn, Thorsten Rogall and other participants at the empirical and development brown-bag seminars at the Vancouver School of Economics for their feedback.

I have been extremely fortunate to share my PhD journey with some wonderful colleagues at the VSE who have helped me more than they can imagine. I would like to especially thank Spreeha Aggarwal, Anand Chopra, Sudipta Ghosh, Nadhanael GV, Leo Ma, Aruni Mitra, Federico Guzman, Anubhav Jha, Dashleen Kaur, Ronit Mukherjee, Clemens Possnig, Catherine van der List and Dongxiao Zhang.

I want to thank my friends outside of the PhD especially, Aritra Chowdhury, Abhirup Majumder and Devjyoti Paul for their wholesome support and companionship.

I thank my parents, sister and the memories of my grandparents for their unconditional love, support and sacrifice – none of this would be possible without them. Finally, I would like to thank my partner Aaheli for bearing with me and absorbing a lot of the burden that came with this journey. This achievement is as much yours as mine.

Chapter 1

Introduction

This thesis consists of three distinct chapters in development economics. The *first* chapter implements a field experiment in India to understand whether the effects of religious diversity on team production and worker attitudes depend on a firm's production technology. The *second* chapter examines the effects of close-kin marriage on a range economic development outcomes. In particular, it uses 19th and 20th century state bans on cousin marriage in the U.S. to identify the causal effect of weakening family ties on income, education and female labor force participation. Broadly speaking, the first two chapters explore the common theme of understanding how social integration (as well as weaker in-group ties) affects economic outcomes. By contrast, the *third* chapter focuses on political business cycles in a developing country context. It studies election cycles in several aspects of mining activity (output, licensing and fatalities) in India and aims to understand why in contrast to other economic activity, mineral extraction is minimized in election years.

Ethnic diversity in manufacturing firms is often associated with lower output due to poor social ties and taste based among workers (Becker, 1957; Lazear, 1998; Hjort, 2014). However, there is very little evidence on how these effects are determined by the nature of production or about the long-run effects of diversity in firms. The first chapter implements a field experiment in a modern factory in West Bengal to estimate the short- and long-run effects of religious mixing on team productivity and inter-group relations under different production technologies.

I partnered with a large processed food manufacturing plant in West Bengal India, that employs both Hindus and Muslims. There are multiple production tasks at this factory. With time-use data on the nature of contact amongst workers, I classify these tasks into two broad categories: High-Dependency (HD) and Low-Dependency (LD). This classification is based on the degree of continuous coordination required amongst workers performing a task to ensure uninterrupted production, and the dependence on teammates for breaks. Worker effort choices have a higher degree of complementarity in HD tasks than in LD tasks, where workers are required to coordinate intermittently.

Two important features of my experimental design are important for identification. The first is that I randomly assign nearly 600 workers to religiously mixed or homogeneous teams. The second is that the firm follows a quasi-random method of allocating workers to tasks. Together, they allow me to attribute different effects of religious mixing in HD and LD tasks to production technology differences rather than differences in characteristics of workers in these tasks. I kept the worker teams intact for a period of four months in order to estimate dynamic effects of mixing.

The study uncovers three key findings. The first is that religious diversity negatively affects team output, but only in HD tasks. In LD tasks, religious diversity is costless. The second key finding is that the difference in output between HD-Mixed and HD Non-mixed teams attenuate completely by the end of the fourth month. The third key finding is that, at endline, there is a reduction in negative out-group attitudes for Hindu workers, which is substantially larger from mixing in HD teams compared to LD teams. This is despite the fact that mixed HD teams suffered negative output shocks. In LD teams, mixing has little or no effects on attitudes of Hindus.

There could be multiple plausible explanations for these core findings. First, I rule out that these results are driven simply by average differences in productivity between Hindus and Muslims – I show that Hindus and Muslims are equally productive in this context. Second, the fact that there are no effects from religious mixing in LD tasks further rules out other explanations based on social reputation concerns around in-group members (Afridi et al., 2020) or distaste for out-group members (Hjort, 2014). Teams in LD tasks are still required to coordinate on many aspects of production even though worker efforts have lower complementarity. Instead, I argue that the effects are driven by the majority group (Hindus) having negative stereotypes about the minority group (Muslims), which leads to the former exerting low effort in high coordination tasks. While it is statically optimal for Muslim workers to also exert low effort in this scenario, given a long enough interaction period, Muslims exert high effort to change stereotypes against them. Hindu workers gradually update their priors, bringing about positive production and attitudinal changes over time. Muslim workers gain from a high-output equilibrium in the long-run.

This chapter contributes to work on ethnic diversity and firm production. Several papers document negative effects of ethnic diversity on productivity (Hjort, 2014; Afridi et al., 2020; Parrotta et al., 2012; Hamilton et al., 2012; Churchill et al., 2017). Hjort (2014) exploits quasi-random variation in the ethnic composition of teams in a Kenyan flower plant and finds that ethnically mixed teams have lower productivity due to taste-based discrimination. Afridi et al. (2020) exploit variation in the caste composition of teams caused by worker absenteeism in Indian garment factories and show that caste homogeneity boosts productivity. Going beyond these papers, I show how differences in the incentives to interact with co-workers due to production function differences, affect team productivity. Second, I estimate the *dynamic* effects of repeated inter-group contact on team production and social preferences in the same setting. Past studies exploit frequent team switching for identification and thus are not able to identify such effects. My results emphasize the need for intergroup contact to occur for a sufficiently long period of time because the minority group does not have the incentive to invest in shifting priors of the majority group otherwise. The disincentive to invest in out-group members in short-term interactions could explain why in firms where teams are frequently switched, a history of being in mixed teams does not reduce prejudice and discrimination (Hjort, 2014).

I also add to work on social preferences at the workplace (Bandiera et al., 2010, 2013; Mas and Moretti, 2009; Carpenter and Seki, 2011; Hjort, 2014; Ashraf and Bandiera, 2018). I show that in the Indian context, factory workers discriminate against non-coreligionists leading to output losses. The plant I study offers a flat monthly wage to its employees. The wage-level is based on seniority and experience at the firm. This is different from the setting in the majority of other papers on this topic, which study team productivity under group versus individual pay structures. I show that even without explicit daily pay incentives, social preferences at the workplace can have large effects on team productivity.

This chapter also relates closely to the literature on how social preferences are formed (Fershtman and Gneezy, 2001; Boisjoly et al., 2006; Jakiela et al., 2011; Mousa, 2018; Rao, 2019) and how the effects of intergroup contact depend upon the *type and nature* of contact (Allport et al., 1954; Pettigrew et al., 2011; Bazzi et al., 2017; Paluck et al., 2019). Lowe (2021) shows experimentally that intergroup contact has different effects based on the type of contact. While Lowe (2021) creates two types of contact (collaborative and adversarial) in a sport setting, I use naturally occurring variation in contact driven by production function differences in a firm setting.

Finally, the literature on employer learning in the U.S. (Farber and Gibbons, 1996; Altonji and Pierret, 1998; Altonji and Pierret, 2001; Lange, 2007) argues that if firms discriminate amongst workers based on easily observable characteristics (such as race), then as employers begin to observe (noisy) indicators of workers' performances, the initial information should gradually become redundant. I show that this holds true for co-workers in a team production setting.

The *second* chapter of the thesis focuses on the role of weakening kinship ties on development. Loose kinship ties have been linked to greater urbanization, economic growth and is thought to have been key to the historical development of Europe (Enke, 2019; Henrich, 2020). Alesina and Giuliano (2014) have linked strong family ties to lower contemporary growth rates through their negative effect on generalized trust, mobility and female labor force participation. However, direct casual evidence underlying this link is largely missing.

We use an exogenous decline in the rate of marriage between first cousins to estimate the effect of weakening family ties on range of socio-economic outcomes. We use 19th and 20th century data from the US, where state-level bans on first-cousin marriage allow us to causally estimate the effect of cousin marriage. While now rare, we estimate that 5% of marriages were between first cousins in the US between 1800-1850. Thirty-two US states have banned first-cousin marriage, starting with Kansas in 1858. We use the timing of the these bans, and the resulting decline in cousin marriage due to their imposition to establish causality. Families with high initial rates of cousin marriage are more likely to have been exposed to these bans than those with low rates of cousin marriage. Indeed, we show that families with high initial rates of cousin marriage see larger drops in these rates in states with early bans. We exploit this for identification.

We measure cousin marriage rates at the surname-level using a method from population genetics which we apply to 18 million marriage records between 1800 to 1940. This method relies on the excess frequency of marriages where spouses share a surname. While it is widely used to estimate cousin marriage rates in a population (Crow and Mange, 1965), this is the first paper to use the method in economics and we apply it to a far larger set of marriages than, to the best of our knowledge, has been done in other disciplines. We track cousin marriage rates over time by surname and link these surname-level rates of cousin marriage to individual economic outcomes in the full count 1850 to 1940 Censuses that include full names.

Our treatment variable is the interaction of surname-level variation in cousin marriage with state-level variation in the duration of bans on such marriages. Specifically we interact (1) the cousin marriage rate for a surname in the pre-period (1800-1858, before the first ban was introduced), and (2) how long a state had banned cousin marriage. This allows us to compare a targeted population of high-cousin-marriage families across states with differential exposure to the timing of bans, rather than simply analyzing the effect on state-wide outcomes. We control for state-specific fixed effects to account for any confounding variation at the state-level. The key assumption for causal interpretation of our coefficients is the following: the timing of bans on cousin marriage should *not* be correlated with factors that affect the relative outcomes of surnames with initially high versus low rates of cousin marriage. We discuss and address possible threats to this identification strategy in the introduction to the chapter.

Our first result is to show that the state bans on cousin marriage were indeed effective. We find that they reduced cousin marriages by about 50% (on average over the entire post-period i.e. between 1859-1940), with the effects being larger in states where bans were introduced earlier. We confirm this result using a separate dataset drawn from genealogical records which allows us to identify 'true' cousin marriage rates, rather than infer them from the frequency of same-surname marriages. We find similar estimates using this alternative dataset which suggests that our primary measure, while noisy, is correct on average.

We use variation in the extent to which a surname was exposed to a state ban and find that greater exposure led to higher incomes i.e. surnames with high initial cousin marriage rates experienced disproportionately larger increases in income. We also find positive effects on schooling. Importantly, we find that the gap in income levels in 1940 between individuals with differential exposure to the bans was absent in 1850, prior to the first ban. This rules out that our results are driven by pre-existing differences. To complement this, we further show that relative gains in income for high-cousin marriage surnames appear only a few generations after the bans start being enacted.

We explore the relationship between consanguinity and congenital health problems as a potential explanation for our findings. However, we do not find that cousin marriage bans affected rates of institutionalization due to physical or mental health issues. Rather, we find large increases in rural-urban migration as well as increases in female labor supply caused by the bans. We thus gravitate to an explanation based on the weakening of tight kinship, since past literature has pointed to these outcomes as markers of weaker family ties.

Our findings add to the literature on the effect of kinship on economic and political outcomes. Our causal micro evidence supports the finding of this literature that tight kinship hinders political and economic modernization. This work typically uses pre-modern measures of kinship tightness from Murdock (1949)'s Ethnographic Atlas and links them to contemporary outcomes (Lowes, 2020; Akbari et al., 2019; Bau, 2021; Schulz et al., 2019; Moscona et al., 2020). Notably, Enke (2019) finds that cultures with higher kinship tightness exhibit more in-group favoritism and hold communal, rather than universal, moral values. Further, he shows that with the onset of the industrial revolution societies with loose kinship experienced faster economic development. Complementary work uses survey measures of the strength of family ties and links these to rich individual-level data on household composition, political participation and economic outcomes (Alesina and Giuliano, 2010, 2014; Ermisch and Gambetta, 2010). Our 19th and early-20th century US setting offers a window into a society undergoing a substantial shift in marriage practices while providing individual-level, population-scale data.

The practice of cousin marriage in particular has been a focus of this literature. This has partly been driven by the influential idea that restrictions on unions between cousins loosened kinship bonds in Europe and led to the development of the modern world (Goody, 1983; Schulz et al., 2019; Henrich, 2020). Schulz (2019) and Akbari et al. (2019) find supportive evidence for this, showing that cousin marriage is linked to worse institutional outcomes and higher corruption. Research in contemporary societies has focused instead on the functional benefits of cousin marriage (Do et al., 2013; Mobarak et al., 2013; Edlund, 2018; Hotte and Marazyan, 2020).[1] The reasons they emphasize, dowry payments, inheritance, and the provision of kin-based insurance, may have been relatively unimportant in the 19th century US, leading the practice to eventually die out.[2] Another rationale for its disappearance in the US was

[1] These may explain the continued widespread practice of cousin marriage in many contemporary societies: Bittles (2001) estimates that about 10% of marriages worldwide are between first or second cousins. An alternative interpretation is the high degree of persistence in the custom of cousin marriage, as seen in Giuliano and Nunn (2020).

[2] Similarly, Munshi and Rosenzweig (2016) argue that kinship (caste) insurance networks reduce rates of rural-to-urban migration, which is consistent with our findings.

growing concern over its genetic consequences. However, recent surveys have concluded that the health consequences of cousin marriages are modest and do not justify legal restrictions (Bittles, 2012; Bennett et al., 2002). Mobarak et al. (2019) offer the best causal micro evidence available on this, using unmarried opposite-sex cousins as an instrument for cousin marriage. Their findings suggest that observational estimates of the negative consequences of cousin marriage on child health are exaggerated and that the true effects are small.

Our use of surnames to measure kinship and marital ties builds on work such as Cruz et al. (2017); Fafchamps and Labonne (2017) and Angelucci et al. (2010). Buonanno and Vanin (2017), in work conceptually related to our own, find that low surname diversity in Italian localities (evidence of in-marriage and limited migration) predicts higher tax evasion but lower crime rates. This is consistent with the idea that cousin marriage generates cohesion within the group at the expense of those outside of it.

The *third* chapter of this thesis studies mining activity between Indian state elections. A large body of literature has focused on understanding political business cycles i.e. how opportunistic politicians stimulate the economy before elections taking advantage of myopic voters (Cole, 2009; Bhattacharjee, 2014; Baskaran et al., 2015). These papers typically find a U-shape pattern in economic activity between elections with the level of activity being greater in election years – since this helps to create a positive perception of those running for office. I study several aspects of mining intensity (mining output, mineral licensing, accidents) in India and document exactly the opposite pattern – mining intensity increases after state elections and gradually dampens leading up to the next election.

I construct a new state- and district-level dataset from India and first document the existence of political cycles in mineral licensing, mining output as well as accidents at the state-level. I then use district-level data to understand how the magnitude of these cycles are affected by (1) electoral competition and (2) conflict (the Naxalite Insurgency in India). I focus on the Naxalite insurgency because extortion from mining companies is believed to be an important source of funding for the rebel groups. I show that both electoral competition and the propensity of conflict intensifies mining cycles. Politically competitive districts as well as conflict prone districts exhibit larger reduction in mining accidents in election years. While mining fatalities are costly during elections in general, I argue that the disproportionately larger reductions in mining activity in conflict prone districts is driven likely by politicians' objective of minimizing the rebels' resource base before elections. This is in the interests of politicians because these groups systematically target elections with violence. Finally, I study the intensity of conflict over the electoral cycle and find that while average levels of violence are not significantly different across mining and non-mining districts, electoral violence is relatively lower in mining districts. This suggests strategic behavior on the part of politicians. Politicians do not directly influence small businesses owners, contractors or even poor villagers who the rebels "tax" (in non-mining areas) (Kujur, 2009), but with their large scale involvement in the mineral industry in India (Asher and Novosad, 2018), they are able to manipulate activity in a way that suits their electoral agenda.

In addition to the main results, I find that mining cycles are larger in districts with greater Scheduled Tribe (ST) population and more diffused in states with higher literacy rates. A large section of mine workers belong to the ST community (Srivastava, 2005) in India. As a result, accidents are more likely to be electorally sensitive in districts having a larger voter base from the same community. Higher literacy rates are generally associated with greater demand for political accountability and lower propensity of exhibiting voter myopia, which could explain these patterns. Overall, this chapter provides new findings on electorally driven political behavior in an industry marred with severe corruption and conflict in India, analyzes the effect of both fixed and time-varying factors on such behavior, and also sheds light on consequences of such behavior on resource related conflict.

Models of political cycles were first developed by Nordhaus (1975) and Lindbeck (1976). The basis of their argument is that opportunistic politicians stimulate the economy before elections taking advantage of myopic voters. Thereafter, Rogoff and Sibert (1988) and Persson and Tabellini (1990) in a separate set of models argued that policy makers signal their ability by creating favourable economic conditions before elections, leading to the emergence of political cycles. The evidence from the empirical literature in developed countries is mixed. Berger and Woitek (1997) find elections to exert significant influence on economic output in both Germany and the United States. Veiga and Veiga (2007) document electoral manipulation in the provision of "visible" collective goods in Portuguese municipality elections. On the other hand, McCallum (1978) and Klein (1996) reject the hypothesis that macroeconomic outcomes are influenced by elections in the United States. Evidence from the developing world though more recent, is now growing. Gonzalez (2002) shows that the Mexican government systematically uses fiscal policy before elections to obtain votes. Drazen and Eslava (2010) find that in Colombia infrastructure spending increases before elections. In the Indian context, Cole (2009) shows that agricultural credit increases in the years running up to a state election and more significantly so in districts with a smaller win margin in the previous election. This cycle is generated only by public and not private banks. He however fails to find any impact of such credit expansions on agricultural output. Khemani (2004) also focuses on India and shows that fiscal instruments are targeted in election years to provide favours to pivotal voting groups. Similarly, Saez and Sinha (2010) find that public expenditure in health and education increases prior to elections in Indian states.

Evidence on political corruption associated with the mineral industry in India is limited, though there is plenty of anecdotal evidence documenting illegal practice. In a recent paper, Asher and Novosad (2018) show that global mining booms in the pre-election period result in criminal politicians running for political office and also winning with greater probability. In the post-election period such booms lead to politicians committing more violent crimes and accumulating greater wealth during their time in office.

There is a strong correlation between the presence of minerals and intensity of India's Naxalite conflict. Extortion of mining revenues by rebels is believed to fuel insurgency in India's "Red Corridor". Majority of the empirical work on Naxalite violence that study static predictors of conflict intensity, acknowledge mineral presence to be a strong determinant (see Ghatak and Eynde, 2017; Hoelscher et al., 2012). Vanden Eynde (2016) finds that negative labour income shocks (measured by deficient rainfall) intensify violence against government forces, but only in districts where the rebels' tax base is independent of local labour productivity (mining districts). When hit by a negative income shock, villagers are tempted to join the rebellion but only in mining districts where rebel groups are able to match their reservation wage, since their resource base is not dependent on local labour productivity. In non-mining districts, civilians are tempted to become police informants instead, leading to higher violence against civilians.

This chapter, though similar in spirit to the literature on political cycles, documents a counter cyclical (inverted U-shape) pattern in mining activity between Indian state elections. Minimizing industrial fatalities which are electorally unpopular, is an important factor driving this, though ex-ante deaths and accidents may seem to be less subject to myopic voting behavior. These results shed light on the dynamics of political behavior over the electoral cycle in the mining industry. Furthermore, the mere existence of political cycles in an industry with both private and public organizations suggests collusive actions between and politicians and private firms. Asher and Novosad (2018) provide evidence of the involvement of criminal politicians in the Indian mining sector. They observe net worth of politicians at the beginning and at the end of their term, and find that mining booms result in larger positive wealth changes. However, they do not analyze behavior over the political term. I address it to some extent in this chapter. This chapter also provides important new findings on the dynamic nature of the Naxalite conflict, and perhaps most critically evidence of political behavior in dealing simultaneously with electoral competition and the Naxalite threat.

Chapter 2

Religious Divisions and Production Technology: Experimental Evidence from India

2.1 Introduction

Evidence suggests that ethnic diversity can lower firm output due to poor social ties and taste-based discrimination among workers (Becker, 1957; Lazear, 1998; Hjort, 2014).[3] However, we know very little about how these effects depend on the nature of production or about the long-run effects of diversity in firms. It is important to develop knowledge of these issues to understand how firms respond to the costs of diversity. If managing a diverse workforce imposes large costs, firms may limit hiring to minimize inter-ethnic interactions, or segregate workers perpetuating discrimination. But these market distortions could be avoided if the negative effects of diversity are mitigated in the long-run through repeated intergroup contact and/or through the adoption of appropriate production technology.

This chapter contributes to our understanding of these issues by implementing a field experiment to estimate the short- and long-run effects of religious diversity on team productivity and intergroup relations under different production technologies. To this end, I partnered with a processed food manufacturing plant in West Bengal, India that employs both Hindus and Muslims — the two main religious groups who have a long-standing history of conflict in India (Pillalamarri, 2019). Production tasks at

[3] There is a large literature on the negative effects of ethnic diversity in decision making in the public sphere as well (Easterly and Levine, 1997; Alesina and Spolaore, 1997; Miguel, 2004). At the same time, diversity has been shown to have positive economic outcomes too – due to strategic complementarities in interacting with out-group individuals (Artiles, 2020; Montalvo and Reynal-Querol, 2017; Jha, 2013) and/or under certain specific requirements of ethnic interaction imposed by authority (Bhalotra et al., 2018; Marx et al., 2021).

the firm can be categorized into the following two types depending on the nature of contact between workers: High-Dependency (HD) and Low-Dependency (LD). This classification is based on the degree of continuous coordination required amongst workers performing a task to ensure uninterrupted production, and the dependence on teammates for breaks. Worker effort choices have a higher degree of complementarity in HD tasks than in LD tasks, where workers are required to coordinate intermittently.[4]

There are two key features of my research design that are important for identification. The first is that I randomly assign nearly 600 workers to religiously mixed or Hindu-only production teams. The second is that the firm follows a quasi-random method of assignment of workers to production tasks.[5] Taken together, they allow me to attribute potentially different effects of religious mixing in HD and LD tasks to production function differences, as opposed to differences in worker types in these tasks. Each production line at the factory comprises a series of (HD and LD) tasks. I designed the experiment to estimate the effects of religious mixing on line-level output, as well as on individual task-level team performance. With line-level output, I identify the *difference* in the effect of mixing in HD versus LD tasks, whereas with task-level performance, I identify the *level* effect of mixing in HD and LD tasks. I kept the randomized teams intact for a period of four months in order to estimate dynamic effects of mixing.

The experiment uncovers three key findings. The first is that religious diversity negatively affects team output, but only in HD tasks. Production lines with mixed teams in HD tasks (HD-Mixed lines) produce 5% lower output than lines with mixed teams in LD tasks (LD-Mixed lines). An analysis of performance measures at the task-level reveals that this loss is entirely attributable to mixed teams in HD tasks. In LD tasks, religious diversity is costless. The second key finding is that the difference in output between HD-Mixed and LD-Mixed lines attenuates significantly over the treatment period – from greater than 20% at the beginning of the experiment, the effect reduced to less than 1% by the end of the fourth month. This is driven entirely by output gains in mixed HD teams. The third key finding is that,

[4]An example of a HD task is work on a fast moving conveyor belt where each worker is responsible for collecting every second or third piece of a product on the belt. Even if only one of them cannot keep up, the machine speed needs to be reduced affecting the productivity of all workers. An example of a LD task is work in a mixing room. Workers typically have well-defined individual duties: for example, one worker is responsible for ensuring that raw materials are weighed properly, another one is entrusted with arranging flour buckets while a third worker mixes the raw materials. The workers need to coordinate intermittently and the productivity of one worker does not directly or immediately influence other workers. A detailed description of HD and LD tasks follows in section 2.2.

[5]The HR manager keeps a pool of job applicants who are assigned to tasks on a first-come-first-served basis when vacancies become available — workers do not get to choose their task when they join or over their tenure. A detailed description of this process and tests to check its validity are presented in section 2.3.4 and Appendix A.1.2.

at endline, there is a reduction in negative out-group attitudes for Hindu workers, which is substantially (23%-56%) larger from mixing in HD teams compared to LD teams. This is despite the fact that mixed HD teams suffered negative output shocks. In LD teams, mixing has little or no effects on attitudes of Hindus.

There are several plausible explanations for these core findings. Since there are no Muslim-only teams in this study, one might worry that these results are driven by productivity differences between Hindus and Muslims. In particular, if Muslims have lower productivity, the treatment effects could simply reflect differences in average productivity between mixed and Hindu-only teams. A number of results and additional tests suggest that this is unlikely. First, if Muslims were less productive overall, we would expect mixing to reduce productivity in LD tasks too. Second, the fall over time in the treatment effect of mixing in HD tasks is unlikely if Muslims were simply unproductive at these tasks. Third, I test for heterogeneity in this attenuation: I find that teams in which Hindus have had greater past contact with Muslims suffer smaller losses initially relative to teams in which Hindus have had little or no contact. The effects completely dissipate for the former group, but remain negative and statistically significant for the latter by the end of the intervention. These dynamics are also inconsistent with Muslims being less productive. Lastly, I show that at baseline Hindus and Muslims were equally likely to be promoted. This suggests that the firm does not perceive them to be differentially productive either. The null effect of religious diversity on productivity in LD sections[6] further rules out other explanations based on social reputation concerns around in-group members (Afridi et al., 2020) or distaste for out-group members (Hjort, 2014). Even though worker efforts have a lower degree of complementarity in LD tasks, teams are still required to coordinate on many aspects.

I develop a conceptual framework and instead argue that the most plausible explanation for the findings here is that Hindus have lower priors regarding how hardworking their Muslim co-workers are, relative to in-group Hindu co-workers. But Muslim workers do not make this distinction. This is because of the asymmetry between Hindus and Muslims in their exposure to non-coreligionists at baseline. Con-

[6]During a period of religious tensions in West Bengal following the passing of the Citizenship Amendment Act (CAA) and subsequent riots in New Delhi, I find religious diversity to have negative effects in LD tasks too. This rules out that mixing in LD tasks is simply a placebo treatment where there are no interaction among workers. Instead, the production technology is such that output is less sensitive to frictions amongst workers. However, extreme events can lead to workers sabotaging out-group members. These results are presented in Table A.22.

sistent with majority-minority relations, Muslims are always in mixed teams with Hindus, while a large section of Hindu workers in the firm do not work with Muslims.[7] This leads to Muslims having accurate priors about Hindus, but Hindus (depending on past exposure) not necessarily having accurate priors about Muslims. In HD tasks with complementary worker efforts, Hindus optimally choose low effort based on the low initial prior about their Muslim co-workers, leading to low team output.[8] Hindu workers do update their beliefs about Muslims and forward-looking Muslim workers internalize this behaviour of Hindu workers. Given a long enough interaction period, Muslims exert high effort despite the fact that Hindus initially exert low effort. This follows because Muslims can persuade Hindus to eventually exert high effort as the latter begin to observe *greater* realizations of high output days than expected under low effort from their Muslim teammates, and as a result gradually update their beliefs. By bearing this short-run cost, Muslim workers benefit from a high-output equilibrium in the long-run.[9] Consistent with this mechanism, I find that during the intervention period, Hindu workers are more likely to blame low output on Muslims (than other Hindus), while Muslims show greater willingness to sacrifice their scheduled break time for Hindus (than other Muslims).

The policy implications of my findings hinge crucially on whether firms are aware of the costs of religious diversity, and how they depend on the production technology. To explore this, I surveyed more than one hundred production supervisors across five different firms that produce similar products. I asked them to predict the results of my experiment and about ways to mitigate possible negative effects of religious divisions. They correctly predicted that religious mixing would be more costly in HD tasks than in LD. But despite the possibility of losses, the majority of supervisors reported to be averse to segregation of workers by religion.[10] About a quarter of the supervisors correctly cited negative effects of

[7] In factories and other formal workplaces across India, Muslims are generally used to working alongside Hindus, while a large share of Hindus are not used to working with Muslims. In this firm, roughly 50% of the Hindu workers worked in homogeneous teams at baseline, while all Muslim workers worked alongside Hindus. Similarly, 43% of Hindus reported to have no contact with Muslims outside of work, whereas only 9% of Muslims reported the same about Hindus. Based on this, together with evidence on discrimination against Muslims in access to education and labor markets in India (Kalpagam et al., 2010; Basant, 2007), I assume Hindus on average (mistakenly) have lower priors regarding how hardworking their Muslim co-workers are, relative to in-group Hindu co-workers. Of course, I show evidence that Hindus and Muslims are not differentially productive in section 2.6.

[8] In LD tasks, worker efforts are non-complements whereby the effort levels of Hindu workers are not dependent on their priors about Muslims. As a result, team output is not affected by diversity.

[9] Note that if the interaction period is not sufficiently long, then the minority group (Muslims) does not invest in the majority group. This is because there will not be enough periods of high-output payoff to recover the loss that the minority group suffers initially by exerting high effort, even as the majority group exerts low effort.

[10] Note that having religiously mixed and Hindu-only teams at the individual task-level (as in the experiment) is natural in

diversity dissipating with repeated intergroup contact, but the first-order concern was about such segregation potentially causing tensions. These findings suggest that effective policy design in this context must look beyond just the direct effects of diversity on production and also trade-off potential short-run costs for long-run benefits of integration.

The rest of the chapter is organized in the following manner. Section 2.2 describes the context: Hindu-Muslim relations in India (in brief) and the study firm: its workers, as well as high- and low-dependency tasks. I discuss the research design and data, and present balance checks in Section 2.3. Section 2.4 presents the econometric specifications used. The results and robustness checks are presented in Section 2.5. In Section 2.6, I discuss plausible mechanisms behind the core findings, and describe an outline of a conceptual framework (the model is presented in Appendix 2.10) for the favored mechanism, and provide some subsequent empirical support. Section 2.7 discusses some policy implications. Finally, section 2.8 concludes.

2.2 Context

2.2.1 Hindu-Muslim relations in India

Hindus form the majority of the Indian population (79.8%), while Muslims are the largest minority (14.2%) group (Census, 2011). Hindu-Muslim conflict has plagued India for centuries and has been a recurring phenomenon since partition and independence in 1947 when the country was divided on religious lines – an episode which itself was marked by large scale religious violence (Talbot and Singh, 2009). Muslims have since suffered greater discrimination and violence against them, as well as borne larger economic losses due to such tensions (Mitra and Ray, 2014). Across the country, Muslims continue to lag behind Hindus on various economic indicators including income and education (Asher et al., 2018), face social exclusion (Alam, 2010) as well as discrimination in the labor market (Kalpagam et al., 2010; Khan, 2019) due to their minority status. Hindu-Muslim relations have especially deteriorated in West Bengal recently as local state politics has seen significant polarization on religious lines (Nath and Chowdhury, 2019).

this context – because Hindus comprise 80% of the population and each task requires five to six workers on average (see Figure 2.1). Supervisors showed concerns about complete segregation of workers by religion on the production floor i.e., having only all-Hindu and all-Muslim teams.

The share of Muslim population varies greatly across states and districts in India. Muslims constitute roughly 25% of the population in the district where my partner factory is located: this is close to the share of Muslims in the factory itself, as well as in other manufacturing plants in the area. Therefore, in terms of representation of Muslims, the factory resembles the average manufacturing plant in the area.

2.2.2 The Factory: Production lines and worker characteristics

In this section, I describe the factory: the structure of production lines and sections, HD and LD tasks, as well as the operation of shifts. I also discuss the pay structure of workers and report characteristics of the workers by religion.

Production lines, sections and shifts

The factory produces packaged bakery products. There are six production lines in total, each of which produces a different product. Figure 2.1 illustrates the structure of the production lines.
Each line is sub-divided into sections (small blocks in the figure) based on the production task that is undertaken in that section. The numbers in parenthesis denote the count of workers in each of these sections.[11] Production occurs in three different shifts: morning, afternoon and night. There are three cohorts per production line, who as a team rotate shifts on a weekly basis.[12] As a result, workers have fixed teams at both the line-level and line-section-level i.e., their co-workers do not typically change, only their shift of work as a team changes weekly.[13]

Religious composition of production lines

Table 2.1 reports the proportion of Muslim workers in each line-level team across the three cohorts at baseline. Line 4 only has two cohorts while all the other lines have three cohorts each. While there is variation in the proportion of Muslims across teams, it is clear from this table that Hindus and Muslims are not segregated in particular lines or cohorts in the factory. On average, each line and cohort roughly have between 15%-25% Muslim workers, which is very close to the overall share of Muslims in the factory. This is formally shown in Figure A.7. I regress a dummy variable denoting a worker's religion on line and

[11] Some of the production lines can produce multiple products and these numbers can vary (though only very little) depending on the exact product being manufactured. Figure 2.1 is based on the number of people in each section during the baseline survey. The numbers during the intervention were slightly different for some sections.

[12] Teams move from morning to night to afternoon shifts.

[13] Occasionally workers are moved across shifts and lines. This is determined by worker absenteeism and turnover.

cohort fixed effects and show that balance in religious composition of production lines and/or cohorts cannot be rejected.

The fact that Muslims are in a minority, together with the structure of production lines that require small section-level worker teams within lines, means that a large section of Hindu workers have little or no contact with their Muslim counterparts. This can be observed in Figure 2.1, where the religious composition of production sections of all six lines is shown for *one particular cohort*. A large number of sections (close to 50%) have no Muslim workers at all. The share of Muslim workers in most of the other sections is between 0.1 and 0.3. The composition is similar across the other two cohorts as well. This is important for two reasons. First, the degree of inter-religious contact induced by the treatment (60% Hindus and 40% Muslims in mixed teams) represents a significant change from the baseline level of contact for Hindus. Second, the majority-minority asymmetry in exposure to non-coreligionists at baseline might mean that Hindus and Muslims behave differently when randomized into mixed teams.

Pay structure of workers

Workers at the factory are paid a flat monthly wage based on their experience and level of expertise (skill) on the job. Wages are not dependent on daily team productivity but performance is evaluated frequently; poor performance over a period of time can lead to workers being moved to a lower skill group. Alternatively, performing well can lead to promotion. Workers are categorized into unskilled, semi-skilled and operator groups. Approximately 80% of the workers are unskilled and the rest are semi-skilled or operators. Semi-skilled workers undertake the same tasks as unskilled workers, while operators are in charge of handling machines.

Characteristics of Hindu and Muslim workers

Summary statistics of worker characteristics are reported in Table A.9. It is apparent that workers are not sorted into HD and LD jobs based on their religious identity. There are however important differences between Hindus and Muslims. Muslim workers have lower schooling, as well as lower tenure at the factory. It has been documented in other studies as well that Muslims on average tend to have lower education relative to Hindus in India (Bhaumik and Chakrabarty, 2009). The difference in average tenure however might be surprising. This can be explained by the fact that in the district where the factory is located, Muslims have traditionally been tailors, which many families still continue to pursue as their business. Since families in this region are typically well-connected, this network allows Muslims to work in the informal tailoring sector, providing them with an outside option of employment. The management often cited this as a factor behind the larger turnover of Muslim workers.

Muslim workers report having much greater contact with Hindus outside of work (as well as at work), which is expected given that Hindus form the majority group in the study area and across India in general. Consistent with this, Muslims report to be more comfortable than Hindus when it comes to communicating with non-coreligionists. Surprisingly, both groups report to be equally uncomfortable taking orders at work from non-coreligionists. Finally, as shown in Table A.9, Hindus are much more likely to support the controversial National Registrar of Citizens (NRC), a bill which is often criticized for discriminating against Muslims.[14]

2.2.3 Direct Dependency as a measure of production technology

Direct Dependency is defined as the degree of continuous coordination required amongst workers performing a task to ensure uninterrupted production. I study it as the key aspect of production technology for two main reasons. First, a key distinction between high- and low-dependency tasks relates to a core idea in economics: the degree of complementary of labor inputs. Worker efforts have a high degree of complementary in HD tasks, while they have a lower degree of complementarity or are non-complements in LD tasks. Second, the degree of complementarity in labor inputs affect incentives to interact, suggesting that this might matter for the effects of religious divisions. Some key characteristics of high- and low-dependency sections (or tasks) are listed in Table 2.2. Figure A.8 provides a visual illustration of HD and LD tasks.

Task coordination

The first key distinction between high- and low-dependency tasks is in the amount of continuous coordination required amongst co-workers. A high degree of continuous coordination is required in HD sections, whereas it is only intermittent in LD ones. I quantify this with time-use data. Research assistants recorded minutes (out of 10) of continuous coordination required amongst workers for production to continue without interruption in each section. HD sections typically require workers to coordinate

[14]The NRC is a list of people who can prove that they came to India before 24th March, 1971. It is a widely held view that together with the Citizenship Amendment Act (CAA), the NRC could be discriminating against Muslims (Chapparban, 2020).

continuously for 9-10 minutes (out of 10), whereas the average in LD sections is only 2 minutes. Sections above the median value (≥ 9) on this scale are classified as HD sections and the rest as LD sections.

The distribution of Direct Dependency is shown in Figure 2.2. Most tasks require either high continuous coordination (9 or 10 minutes out of 10) or less than 2 minutes of continuous coordination – this leads to the bi-modal distribution in the figure. This allows easy classification of tasks into HD and LD types, an important (third) reason to pick this measure over others.

Control over breaks/relief time

The second key distinction between HD and LD tasks is about control over breaks during the production process. Due to dependence on co-workers every minute of the production process in HD tasks, each worker individually has little control over when they can take a break. Sub-groups of workers need to provide “relief" to other workers in the same HD section, a concept known as “relief time". There are often disagreements amongst workers regarding how to schedule these as well as arguments when some workers take more time than allocated. Supervisors reported such disruptions to be a common cause for lower productivity. By contrast, in LD sections each worker has much greater control over scheduling breaks.

Physical mobility

Physical mobility is restricted in HD sections. For example, workers are typically required to stand close to each other on conveyor belts and pick products up as they move on the belt. Coordination with others doing the same is therefore key. In LD sections, greater individual control over the production process allows workers greater physical mobility.

Repetitive monotony

Repetitive monotony is higher in HD sections compared to LD sections since work cycles are shorter. The machine speed set by the supervisor often determines the speed of work, allowing workers little control over the process. If workers do not perform up to the mark, supervisors may need to reduce machine speed causing loss in output. Informal interviews with the supervisors made it clear that it is not uncommon for them to vary machine speed in these areas. This could happen due to worker absenteeism leading to changes in teams, as well as due to workers simply not coordinating as expected on certain days of production. In LD sections, workers typically have more control over process speed, and can re-allocate their time across different sub-tasks to a greater extent.[15]

Direct Dependency and other task-level characteristics

In Figure A.6, I show all six production lines at the factory broken down into HD and LD sections. In Table A.10, summary statistics of various aspects of the physical environment of HD and LD sections are presented. I focus on factors which could act as potential confounders to the main mechanism in this paper. I measured the degree of non-work interaction (time workers spend chatting) and noise levels in each section of each production line and rule out that HD and LD sections are systematically different on these aspects of the physical work environment. The only statistically significant difference between HD and LD sections is in the average temperature; HD sections tend to be warmer by two degree Celsius. This difference is primarily due to a few colder LD sections in one particular production line. One could worry that hotter temperatures might intensify the negative effects of religious divisions, driving part of the effects that I find. This is not the case – all my results are robust to dropping this production line/sections from the analysis.[16]

2.3 Research design

This section discusses the research design. I first go through the randomization process and then present balance checks over a range of worker characteristics across the different treatment arms. Before the intervention, workers were informed that their teams would be changed in order to assess the effect of team-switching on productivity. The new team lists (post randomization) were printed and posted on the production floor. No additional information was provided. Religion of teammates can be directly inferred from their names in this context.[17]

[15]From the description of HD and LD tasks it might seem LD tasks are unequivocally better, but that is not the case. There are various aspects of LD tasks, such as heavy lifting in certain sections or working in mixing rooms that have unpleasant smell which workers reported to dislike.

[16]These results are available upon request.

[17]Whether a person is a Hindu or a Muslim can be determined from their first name itself in the Indian context. In very few cases where the first name maybe ambiguous, the last name would certainly reveal one's religion. My sample consists of only Hindus and Muslims.

2.3.1 Treatment and randomization

As mentioned earlier, the factory operates in three shifts (morning, afternoon, night) and an entire cohort of workers move from one shift to the next on a weekly basis. A new set of workers come to work in each shift on a particular day. Therefore, each line has three different cohorts working on it each day of the week. For the purpose of randomization, I moved workers across cohorts, holding their production line and section of work fixed.[18]

Individual workers were randomized into line-section-level teams in order to achieve two distinct types of teams (treatments) at the line-level. The first type comprised of line-level teams with religiously mixed groups only in HD sections (HD-Mixed lines), while the second type had religiously mixed groups only in LD sections (LD-Mixed lines). Two of the randomized cohorts within each line were of one team type while the third cohort was of the other type. Figure 2.3 below provides a visual illustration of the two types. I use Line 2 from Figure 2.1 for this illustration.

Individual section names are replaced by HD and LD labels to denote section (task) type. The first type of line-level team has all its HD sections mixed (partly shaded in grey) while its LD sections are comprised of only Hindu workers (HD-Mixed line). The structure in the second type is exactly the opposite – LD sections have religiously mixed teams while HD sections have only Hindu workers (non-shaded) (LD-Mixed line).[19] This leads to four different types of line-section-level teams: 1. HD Mixed 2. HD Non-Mixed 3. LD Mixed and 4. LD Non-Mixed. Whether a production line would have two cohorts of HD-Mixed lines (and one LD-Mixed) or the other way round was determined by the overall number of Hindus and Muslims in the line at baseline. Production data are available for both line-level as well as line-section-level teams. Therefore, any differences in overall line-level performance between teams can be disaggregated to line-section-level performance.

Randomization was constrained by one key limitation – the number of workers switching their section of work (their task) had to be minimized. Even though the induction of workers to specific tasks (unless as an operator) takes only between one to two days, it is impossible to train all workers in new

[18]For a small share (7.9%) of workers this was not the case. Some workers had to be (randomly) moved from their tasks at baseline to achieve the desired line-level team types. However, such task-shifting is not correlated with treatment status. Section 2.5.3 includes a discussion on this.

[19]At the line-section-level, religiously mixed and Hindu-only teams are the ones that are naturally formed at baseline (recall Figure 2.1).

tasks simultaneously – this would lead to substantial interruptions and breakdown in production. The management was not willing to do this. As a result, the randomization process was designed such that did it not require the majority of workers to change their section of work and hence the *dependency* of their task at baseline. I address concerns with respect to selection of workers into HD and LD jobs subsequently.

The first step in the randomization process involved determining the final (target) number of Hindus and Muslims in each section of each production line (across all cohorts). Since workers were not moved across production lines, this was typically constrained by the overall number of Hindus and Muslims in a line across the three cohorts at baseline. The share of Muslims in each production line at baseline was close to the overall share of Muslims in the plant (approximately 18%). After randomization, the share of Muslim workers in mixed sections (both HD and LD) of all six lines was typically between 35%-40% (this was of course balanced between HD and LD sections).[20]

The second step in the process involved sorting workers by section × religion × skill[21] (across all 3 cohorts in a line) and shifting workers across sections (tasks) in order to ensure that each section of each line had enough Muslim workers (summing across cohorts) required for randomization (as determined in the first step). This had to be done at baseline because not all sections of all lines had enough Muslim workers (sometimes none) such that the desired line-level team structures in Figure 2.3 could be achieved. For example, the *Injector* section in Line 3 had no Muslim workers at all across the three cohorts. In such cases some randomly chosen Hindu workers in that section were shifted out and replaced with randomly chosen Muslim workers from another similar section with enough Muslims. This process meant that at the end of step 2, all sections of all lines had both Hindu and Muslim workers[22] who would then be randomly allocated to line-section-level teams. This also satisfied the management's requirement of minimum section (task)-shifting.

[20]Note that the religious composition of a particular section in a line would be exactly the same across all cohorts if they belonged to the same line-level team type. In other words, if cohorts A and B in Line 1 were such that all their HD sections were mixed and LD sections were non-mixed, then each of their HD sections would have exactly the same ratio of Hindu to Muslim workers i.e. *Packing* in cohort A would have exactly the same number of Hindus and Muslims as *Packing* in cohort B. Non-mixed teams of course only have Hindu workers.

[21]Workers are classified into three skill levels: unskilled, semi-skilled and operator. Each section typically has an operator or a semi-skilled worker (depending on the type of work), and the rest are all unskilled workers. The randomization process did not alter this structure.

[22]This is required because for each section of each line there would at least be one line-level team where that section would have to have a mixed group.

Lastly in the third and final step, workers were sorted by their new section (post step 2) × religion × skill and randomly allocated into line-section-level teams in order to achieve the line-level team structures shown in Figure 2.3. Line-level teams were then randomly allocated to one of the three shifts. A detailed description of each step involved in the randomization process is presented in Appendix A.1. Figure A.1 provides a visual illustration of the same, especially focusing on how section-shifting allows formation of the desired line-level team structures.

2.3.2 Data collection, experiment timeline and attrition

Data used for the analysis in this paper come from two main sources. I use administrative records of production obtained directly from the firm's management to estimate treatment effects of diversity on line-level output. The firm records total output at the line-level line in each shift; this measure is tied directly to the revenues of the firm. Before the intervention, supervisors were also trained by the production manager to rate the performance of each line-section-level team independent of the performance of the entire line or other sections in the line. These ratings are used to directly estimate the effect of diversity on output in HD tasks separately from LD tasks.[23]

Workers participated in an in-person survey at baseline but only a phone survey could be conducted at endline due to COVID-19 related restrictions in India. The baseline survey included a wide set of questions ranging from employment related ones such as tenure, history of past teams, attitudes towards taking orders from and interacting with non-coreligionists, to objective worker characteristics such as age and schooling. I also asked workers about their political preferences, focusing on factors that could capture taste discrimination towards religious groups. These include preference for political parties that are associated with favoring a particular religious group and support for bills that are widely criticized for discriminating against Muslims.

The focus of the endline survey was primarily on interactions (accusations, blame, providing relief time etc.) that happened during the intervention and on worker attitudes that could capture the effects of inter-religious contact in HD and LD environments on inter-group relations. Summary statistics of key

[23] It is nevertheless possible that these ratings do not appropriately take into account spillover effects from upstream to downstream sections. In section A.3.3 (Appendix), I restrict attention to sub-samples for which spillovers are likely to be less of a concern and show that my main results are replicated.

variables are presented in Table A.9; differences in characteristics of Hindus and Muslims have already been discussed in section 2.2.2. Figure 2.4 presents the timeline of the intervention and sample size by treatment arm. There are 15 line-level teams[24] (7 HD-Mixed Lines and 8 LD-Mixed Lines) and 113 line-section-level teams (23 HD-Mixed, 33 LD-Mixed, 29 HD Non-Mixed and 28 LD-Mixed). A total of 586 workers were part of the intervention distributed in the following way in line-section-level teams: 175 in HD-Mixed, 117 in LD-Mixed, 196 in HD Non-Mixed and 98 in LD Non-Mixed. A total of 546 workers could be reached at endline for the phone survey (attrition rate 6.8%).[25]

2.3.3 Randomization check

Balance checks in Table 2.3 show that randomization was successful. Outcomes are divided into two broad categories - (1) those that are relevant at work (Panel A) and (2) general characteristics and attributes (Panel B). The unit of analysis here is an individual. The main regressors are the interaction terms Mixed × LD and Mixed × HD which denote the type of line-section-level team and hence the treatment status of an individual. Line × Section fixed effects are included in these specifications, whereby the main effect of HD versus LD is not separately identified. The omitted group is therefore all workers assigned to non-mixed teams.[26] Across a range of characteristics that include factors that are relevant at the workplace (such as tenure and past contact with non-coreligionists), as well as general attributes (such generalized trust, altruism and contact outside work), workers are similar across the treatment arms.

Finally, it is also important to show that the proportion of Muslim workers is balanced across mixed HD and LD teams, to rule out that the treatment effects are driven by different "degrees" of religious mixing across the two types of tasks. This is shown in Table A.4.

[24]Note that at full capacity the firm would have 17 line-level teams as shown in Table 2.1. However, in the experiment there are 15 line-level teams only. This is because during the period of the intervention, the firm decided to operate at lower capacity due to low product demand compared to previous years (even though the experiment was timed to coincide with the period when, in terms of seasonality, the firm usually experiences the highest demand). As a result, production lines 1-3 had three cohorts each whereas lines 4-6 only had two cohorts each (Figure 2.1). This change occurred before the randomization began, so the experiment was not affected by it.

[25]In Table A.25, I show that attrition is balanced across treatment arms.

[26]I use this particular specification for balance checks because the same specification is used to estimate treatment effects at the line-section-level on team production, as well as on individual-level survey outcomes. As a robustness check, I use Line fixed effects instead of Line × Section fixed effects in Table A.5 (whereby the main effect of HD versus LD is identified) and show that worker characteristics are balanced across HD and LD sections. I also show balance in individual characteristics across line-level teams (i.e. HD-Mixed lines versus LD-Mixed lines overall) in Table A.6.

2.3.4 Quasi-random allocation of workers to tasks at baseline

Since the majority of workers continued to work in their original tasks (i.e. the area of work was not randomized), one might worry about distinguishing between the effects of task types versus worker types (on team productivity) in religiously mixed teams. This is particularly important if workers are able to self-select into high- or low-dependency sections. The randomization check already rules out such systematic sorting. Nevertheless, I address this concern in more detail in Appendix A.1.2. I argue that worker characteristics are balanced across HD and LD tasks due to the firm's hiring and worker allocation policy and not simply by chance. The HR manager always has a pool of job applicants who are called upon on a first-come-first-served basis, when vacancies become available. As a result, workers do not have the option to choose their area of work when they join. However, workers may quit at different rates across the two types of tasks, leading to possible selection bias. If that were the case, this would be reflected in the average tenure of workers in HD and LD sections. As shown in Table A.5, this is not the case – tenure is balanced between workers in HD and LD sections. I then show that only a handful of workers (15.9%) have switched their area of work from when they first joined the firm. Finally, I show that these switches are not correlated with observable characteristics of the workers and have happened purely due to organizational requirements at the firm.

2.4 Econometric specification

Outcomes in this paper are measured at three levels: 1. Production line-level, 2. Production line-section-level, and 3. Individual-level. Line-level real output data are linked to the firm's revenues. Line-section-level ratings were recorded by production supervisors daily during the period of the experiment only. These data help investigate the source of line-level differences in real output. Survey measures at baseline and endline are at the individual worker level. I use these to study worker interactions during production as well as treatment effects on attitudes.

Line-Level specification

I compare line-level output between HD-Mixed and LD-Mixed lines as shown in Figure 2.3. The specification used is:

$$Y_{klst} = \beta_1 T_k + \alpha_l + \alpha_s + \alpha_t + \epsilon_{klst}, \tag{2.1}$$

where Y_{klst} is output from line-level team k, in line l, in shift s on day t. T_k denotes the treatment status (1 if HD-Mixed line and 0 if LD-Mixed line). The coefficient β_1 denotes the line-level treatment effect. α_l, α_s and α_t are line, shift and day fixed effects respectively. I include production line fixed effects to control for product type, shift fixed effects to account for differences in worker productivity at different times of the day (morning, afternoon, night) and day fixed effects to control for factory-wide shocks to demand. Standard errors are clustered at the line-cohort-level (or in other words at the line-level team). Since there are only 15 clusters at the line-level, I also present wild cluster bootstrap standard errors (Cameron et al., 2008) for these regressions.

Line-Section-Level specification

Supervisors assigned a daily rating (out of 5) to each line-section-level team, independent of the performance of other sections in the line. I use this data to evaluate the source of line-level differences in output. The following baseline specification is used:

$$Y_{mklst} = \beta_1 Mixed_{mkl} \times LD_{ml} + \beta_2 Mixed_{mkl} \times HD_{ml} + X_{mkl} + \alpha_{ml} + \alpha_s + \alpha_t + \epsilon_{mklst}, \tag{2.2}$$

where Y_{mklst} is the performance rating of section m of team k in line l in shift s on day t. $Mixed_{mkl}$ denotes whether the section has a religiously mixed or homogeneous team (which is determined by line-level team type k). LD_{ml} and HD_{ml} are dummies coded 1 if the section is classified as HD and LD respectively (this is defined by line l and section m only). I use the interaction terms $Mixed_{mkl} \times HD_{ml}$ and $Mixed_{mkl} \times LD_{ml}$ to identify effects of having mixed teams in HD and LD sections respectively (given by the coefficients β_1 and β_2). Since *line* × *section* effects α_{ml} are included in these regressions, the dummies HD_{ml} and LD_{ml} are not separately introduced. X_{mkl} is a vector of line-section-level controls. α_s and α_t are shift and day fixed effects respectively.

Individual-level specification

I surveyed workers both at baseline and endline. I use the baseline data for randomization checks as shown in section 2.3 and also for heterogeneous treatment effects which follow in section 2.5. As mentioned earlier, the endline data is used to evaluate treatment effects on worker attitudes and interactions between teammates during production. The main specification is:

$$Y_{imkl} = \beta_1 Mixed_{mkl} \times LD_{ml} + \beta_2 Mixed_{mkl} \times HD_{ml} + X_{imkl} + \alpha_{ml} + \epsilon_{imkl}, \tag{2.3}$$

where Y_{imkl} is the outcome of interest for individual worker i of section m of team k in line l. X_{imkl} is a vector of individual-level controls. All other variables are described exactly as before. The treatment effects are estimated by coefficients just as in the line-section level specification described above.

2.5 Results

2.5.1 Production data

This section begins by showing that HD-Mixed lines produce lower output than LD-Mixed lines, but this effect attenuates over time. I then proceed to the line-section-level analysis and show that line-level differences in output are driven entirely by losses from religious mixing in HD sections, while mixing has no effect in LD sections.

Line-Level

Production supervisors record total output from each production line at the end of each shift. Table 2.4 shows that HD-Mixed lines produced lower output compared to LD-Mixed lines. Observations in this regression are at the line-cohort-day-level. The outcome variable in Column (1) is the log of total output (in pieces) produced by a line-level team in a particular shift of a day. Column (1) shows that HD-Mixed lines on average produced 5% lower output compared to LD-Mixed lines over the period of the intervention. This effect is economically large. Given average output per shift of 450,000 pieces (across all lines) and the average product priced at Rs 10 ($ 0.13), the results suggest that the firm's revenue would increase by up to Rs.225,000 ($3100) per shift, from having only LD-Mixed lines relative to having only HD-Mixed lines.

The firm also records total output using the number of boxes with final products that are packed at the end of a shift. These boxes are used to ship products to the market and each box typically includes multiple pieces of a product. The effects are robust to using this variable as the outcome instead (Column 2). Since each production line can manufacture more than one variant of the same product, I show robustness to the inclusion of *line* × *variety* fixed effects in Table A.11. Finally in Table A.12, I include *line* × *day* fixed effects – the results remain robust.[27]

Over the entire period of the intervention HD-Mixed lines produced 5% lower output than LD-Mixed lines, but how did the treatment effect evolve over time? This would inform us whether repeated interaction with the same set of non-coreligionist co-workers can help ameliorate some of the negative effects of mixing on output. In Figure 2.5, I present a event study plot of output (logged) produced over the period of the intervention, by team type. These are from binned regressions using the same specification as in section 2.4, with the treatment period split into five equal sized bins. The difference in output produced by HD-Mixed and LD-Mixed lines was the largest at the beginning of the intervention and it gradually attenuated over time.[28] Interestingly, output from both HD-Mixed as well as LD-Mixed lines followed an upward trajectory throughout the four months of the intervention. This might be because of two reasons: 1. The firm was itself adjusting to new teams and therefore only gradually increased production targets as workers became more comfortable with each other and 2. The experiment was timed to coincide with the period during which the factory faces high demand for its products; so that production remains uninterrupted, absenteeism is low and teams don't disintegrate. This could have also led to the firm setting higher output targets in each subsequent month of the intervention.

Overall, these results imply that religious diversity is *relatively* more costly in HD tasks than in LD tasks. But the overall line-level differences (between HD-Mixed and LD-Mixed lines) could be driven by religious mixing lowering output in both types of tasks but more in HD, or mixing increasing output in both types of tasks but more in LD. Another possibility is that it affects output negatively (or has no effect) in HD tasks, but positively in LD tasks. Finally, it is also possible that mixing only (negatively)

[27]Based on raw material usage, supervisors at the firm record *standard output* against actual output produced. Negative deviations from standard output imply greater raw material wastage. In Appendix A.2, I show that wastage or "Output Gap" is larger in HD-Mixed lines despite raw materials being allocated equally among HD-Mixed- and LD-Mixed lines. Furthermore, the variance of Output Gap is also greater in HD-Mixed lines, suggesting that diversity in HD tasks lead to relatively greater uncertainty in terms of achieving daily output targets. Speculatively, this might mean that team output is more susceptible to idiosyncratic shocks such as religious events or conflict from religious mixing in HD work environments.

[28]The difference in *standard output* between the teams is much smaller (and statistically indistinguishable from 0) as observed in Figure A.9. This would be expected if the firm did not react to these differences across teams by redistributing planned production away from low productive lines to high productive ones.

affects output in HD tasks but not in LD tasks. I cannot distinguish between these possibilities using line-level data as there are no homogeneous line-level teams by design. I take this up next in the line-section-level analysis, where such comparisons are possible due to the presence of teams composed of only Hindu workers. In other words, the *level* effect of diversity in HD and LD tasks can individually be identified at the line-section-level, whereas at the line-level only the *difference* from religious mixing in HD versus LD tasks could be identified.

Line-Section-Level

I now present treatment effects on line-section-level performance ratings. Recall that there are four different types of teams at this level: HD Mixed, HD Non-Mixed, LD Mixed and LD Non-Mixed. The performance of each section was rated (between 0 to 5) daily by production supervisors. These ratings were based on a benchmark measure of time-use efficiency. The benchmarks were different across tasks. For example, *Mixing* sections were rated on the number of batches mixed per hour, while most other sections downstream until *Packing* were rated on the number of trays with unfinished products that were sent onto the following section every hour, accounting for the number of trays received from the previous section. This ensured that no section was penalized for the actions of sections upstream. *Packing* sections were rated on the number of boxes packed with final goods as well as on packaging material wastage.

Table 2.5 presents the core results from the line-section-level analysis. In column (1), I regress raw ratings on a dummy variable that denotes whether a line-section-level team is religiously mixed or not (Mixed). The coefficient on Mixed is negative and marginally significant suggesting that mixed teams perform worse overall. Note that line × section effects are included in all specifications in the line-section-level analysis, whereby the identifying variation comes from within the same line-section across different treatment cohorts (teams). These are important to include because of the different benchmarks used to rate each section.

All regressions also include average tenure and schooling of workers in the section as controls, to account for differences between Hindus and Muslims on these dimensions.[29] In column (2), I introduce

[29]The results are robust to the exclusion of these controls as reported in Table A.14 in the Appendix.

the interaction terms (Mixed × HD) and (Mixed × LD) to estimate the effect of having a mixed team in a HD section separately from a LD section. The coefficient on Mixed × LD in column (2) is small and not statistically significant while that on Mixed × HD is negative and statistically significant at the 5% level. This suggests that having mixed teams lead to lower ratings in HD sections but not in LD ones. In columns (3) and (4), the outcome variable is coded 1 if the the rating received is above median and 0 if lower.[30] The effects with a binary dependent variable are similar to those with raw ratings and more precisely estimated. In summary, this is direct evidence that lower output in HD-Mixed lines (relative to LD-Mixed) is caused entirely by lower output in religiously mixed HD sections, while in LD sections, mixing is costless. In Table A.13, I run separate regressions for HD and LD sections and find similar effects.

I next examine whether there is convergence in line-section-level performance over time between mixed and non-mixed HD teams. This is likely given that line-level output differences between HD-Mixed and LD-Mixed lines attenuated over time (recall Figure 2.5). I split the intervention period into five equal sized bins (exactly as in the line-level analysis), and show that this is indeed the case. The results are presented in Table A.15.

The baseline effect is reported in column (1), which shows a large, negative and statistically significant effect of having a mixed HD team. In column (2), I introduce interaction effects with the event bins. Coefficients on earlier bins are larger (negative) and they gradually reduce in magnitude. This suggests that the largest negative effects of religious mixing on HD section output occurred at the beginning of the experiment when the new teams were first formed; and performance ratings of mixed and non-mixed teams gradually converged over time. The baseline effect and interactions with the event bins are presented for LD sections in columns (3) and (4) respectively. The baseline effect is small and not statistically significant, while the interactions are noisy with no clear dynamic pattern. Overall, these results are re-assuring in that they line up closely with the line-level event-study analysis, but using granular production data at the line-section-level.

[30]A large fraction of ratings is concentrated between 4 and 5 (see Figure A.10), making a binary dependent variable also appropriate for this specification. I aggregate up line-section-level ratings to the line-level (averaging across all sections) and run specification 2.1. The results are presented in Table A.23, which show, similar to Table 2.4, that HD-Mixed lines perform worse than LD-Mixed lines, although the effects are less precisely estimated due to the smaller variation in section ratings compared to actual line-level output.

2.5.2 Endline phone survey

The endline survey focused on two main sets of outcomes: 1. Those that capture actual interactions between workers during production and 2. Attitudes towards non-coreligionists co-workers. Only a phone survey could be conducted at endline because of restrictions related to COVID-19. As a result, a large set of outcomes that I was interested in, including political preferences that respondents maybe uncomfortable discussing over the phone, could not be recorded.[31] I take up each of the two sets of survey outcomes in turn.

Worker interactions

In Table 2.6, I focus on the first set of factors. These collectively proxy for the degree of cohesion and coordination in a line-section-level team. There are three main outcomes variables. The first question asked respondents to identify co-workers who they thought did not contribute sufficient effort at any point during the intervention ("Identified teammate as contributing low effort"). If a worker identifies his teammate to have not contributed as much effort as other workers did, or to the extent that is expected, then this outcome is coded 1 for a worker-teammate pair. I then asked workers to identify teammates who have blamed them in the past for not performing up to the mark ("Blamed by teammate"). The outcome variable is coded 1 for teammates who have blamed the respondent at least once during the intervention period. The final question asked workers to pick teammates who they would give up their relief time for, if asked or already have in the past. Relief time refers to breaks that each worker is entitled to at regular intervals during their shift. In HD sections, workers typically need to coordinate on breaks to a greater degree than in LD sections. The outcome variable is coded 1 for teammates that workers are *not* willing to give up their relief time for ("Unwilling to give up relief time"). Note that these questions were asked retrospectively in lieu of more high frequency data, since many workers reported to have had

[31] In addition, one might be worried about social desirability bias in the responses, since the outcomes I study are self-reported (even though both the baseline and endline surveys were conducted one-to-one with the respondents and anonymity and confidentiality were emphasized). To deal with this, I correlate baseline responses to survey questions (that were asked again at endline and used as outcomes in Table 2.7) with scores from an Implicit Association Test (IAT) that the workers took. The test involved associating Hindu and Muslim names with positions in the firm hierarchy (worker, operator, supervisor, production manager etc). A positive score on this test denotes a bias towards having Hindus in higher positions, while a negative score shows preference towards Muslims. I correlate these scores with workers' reported attitudes towards taking orders from non-coreligionsts as well communicating with non-coreligionists (Figure A.12). Hindu workers with a larger positive score are less likely to say they are comfortable taking orders from and communicating with Muslims. Similarly, Muslim workers with a larger negative score are less likely to say they are comfortable taking orders from and communicating with Hindus. This suggests that workers' responses are correlated strongly with their actual preferences and helps provide confidence in the self-reported survey outcomes.

problems with their teammates in the past but also mentioned that they subsided over time.[32]

Observations in Table 2.6 are at the worker-teammate level for line-section-level teams. In other words, there are $(N-1)$ observations for each worker, where N denotes the total number of workers in the line-section-level team. I include line × section fixed effects and therefore compare similar size teams doing the same task. Columns (1), (3) and (5) show that mixed teams perform worse on all of these measures. Workers in mixed teams are 4.2 percentage points (30%) more likely to identify a teammate as contributing low effort, 4 percentage points (50%) more likely to have been blamed by a teammate and 6.4 percentage points (25.6%) less likely to give up their relief time for a teammate. In columns (2), (4) and (6), I introduce the interaction terms Mixed × HD and Mixed × LD to test for differential effects by task type. Clearly, having mixed teams in HD sections lead to greater frictions.

However, I find that workers in mixed LD sections report to have been blamed more by co-workers than those in mixed HD sections. Individual mistakes are more easily identifiable in LD tasks compared to HD tasks, which is perhaps why this pattern is observed.[33] Note that both of these effects are statistically significant on their own.

More generally, it can be observed that mixed teams in LD sections also suffer from these frictions to a greater extent than homogeneous teams – the effects on the interactions Mixed × LD are positive and meaningful in magnitude though not precisely estimated. In fact, one cannot statistically reject that the effects in LD sections are different from those in HD, though the effects in HD sections are much larger. Importantly however, these do not translate into mixed teams performing any worse than non-mixed teams in LD sections, which is the case in HD sections, as shown in Table 2.5. The sample is restricted to only Hindu respondents in Table A.16 and similar patterns are observed.

These results are consistent with the treatment effects on output and inform us of actual interactions between workers that led to those effects. Coordinating closely as a team on a wide set of issues is important in HD tasks and lower team cohesion caused by these frictions can reduce team output. While mixing in LD tasks also leads to some frictions, the production technology is such that team output is less likely to be sensitive to these problems, which explains the null effects in LD tasks.[34] But despite these

[32]For example, workers were asked if they have been blamed by a teammate at least once in the past, or asked to identify workers who they thought were not contributing effort at any point during the intervention.

[33]It is also plausible that by endline these frictions had subsided more in HD sections than in LD ones.

[34]Of course, extreme events such as religious violence after the passing of the Citizenship Amendment Act (CAA) did affect

frictions, output differences between HD-Mixed and HD Non-Mixed sections attenuated over time. The next set of results study treatment effects on attitudes of workers towards non-coreligionists at endline, and formally tests whether the attenuating output effects are accompanied by improved inter-group relations.

Attitudes at endline

For attitudes, treatment effects are restricted to Hindu workers only, as Muslim workers are always in mixed teams. I use three main outcome variables, two of which are questions also asked at baseline. Workers were asked if they are equally comfortable taking orders from non-coreligionists ("Taking Orders"), whether they find communicating with non-coreligionists (in general) as comfortable as co-religionists ("Communicating") and finally if they prefer to be in mixed or all-Hindu groups if teams were to change again in the future ("Co-working"). While the first two questions were unincentivized, for the third question, surveyors mentioned to the workers that their responses would be recorded and kept in mind for future team changes.

I first report the main effect of being randomized into a mixed team. Outcomes in Table 2.7 are at the individual worker level. All outcomes show positive effects from mixing. Relative to those in homogeneous teams, Hindu workers in mixed teams are 16.8% more likely to report that they are comfortable taking orders from Muslims (Column 1) and 20% more likely to be comfortable communicating with Muslims (Column 2). Finally, Column 3 shows that they are 18.7% more likely to *not* express preference for being in a Hindu-only team. These effects are economically significant in magnitude and suggest large gains for Hindu workers from repeated contact with Muslim colleagues. In Columns (2), (4) and (6), I introduce the interaction terms Mixed × HD and Mixed × LD. The effects are entirely driven by contact in HD sections.

The coefficients on Mixed × HD are economically large in magnitude and statistically significant at the 1% level. The coefficients on Mixed × LD are small and not statistically significant, suggesting a null effect in LD sections. The differences between the effects in HD and LD sections are large and statistically significant. These findings on positive attitude changes of Hindu workers towards Muslims (from mixing

output in mixed LD sections (see Table A.22). But overall, the lower sensitivity of team output to these frictions is perhaps also why there is little incentive for workers to try to overcome their differences. This is reflected in the next set of results where I show reductions in negative out-group attitudes from mixing for Hindu workers, but only in HD tasks.

in HD tasks) are consistent with attenuating output differences between mixed HD and non-mixed HD teams (Table A.15), as well as with the overall convergence in line-level output between HD-Mixed and LD-Mixed lines (Figure 2.5).

Summarizing the main results

Overall, it is insightful and non-obvious that the largest positive effects of treatment on attitudes occurred in teams that also suffered the largest negative output shocks. This suggests that working in close quarters even with some frictions (in HD teams) leads to more positive effects on intergroup relations than working in LD teams. These results emphasize the importance of contact that forces people to learn to work together in overcoming existing differences leading to reduced prejudice. Purely from a profit maximizing point of view however, firms may have little incentive to mix workers in HD tasks if it leads to output loss. This unfortunately suggests that discrimination may persist in equilibrium and emphasizes the need for targeted management practices to mitigate them.

2.5.3 Robustness: Threats to identification

In this section, I discuss potential threats to the identification strategy and describe how they are dealt with. I discuss factors linked to the research design (such as the absence of Muslim-only teams) as well as those that randomization cannot directly account for (such as differences between mixed and homogeneous teams on demographic dimensions other than religion).

Religion and productivity

First, I address concerns regarding potential bias that could stem from religion simply proxying for differences in other dimensions (education, tenure, etc.) between mixed and non-mixed teams induced by randomization. One might be worried that it is not the interaction between religious mixing and the production technology itself that leads to productivity loss, but that this interaction simply proxies for these other differences between Hindus and Muslims (or HD and LD tasks).[35] For example, it might be the case that differences in schooling between Hindus and Muslims are not important in LD sections,

[35] A more fundamental worry may be that due to lower schooling and tenure, Muslims uniformly have lower productivity. If that were the case however, we should find religious mixing in LD tasks to lower output as well – but we do not find that. I nevertheless control for these factors in the line-section-level analysis, though the results remain robust to excluding them instead.

but might be a problem in HD sections, given the nature of contact. In other words, it is differences in education between Hindus and Muslims that matter in some tasks and not others, as opposed to the production technology being the important factor.

To deal with this, I introduce interactions between the dummy variable Mixed and these variables as controls, in addition to the interaction terms Mixed × HD and Mixed × LD in the line-section-level specification. I specifically use three variables: group size, tenure of workers and schooling of workers. HD sections tend to have more workers, and one might be concerned about differences in responses of workers from being mixed in larger groups as opposed to smaller groups. For example, diversity might be costly when groups are larger because there is likely to be a wider set of issues that require coordination on. The other two are more obvious choices given the differences amongst Hindu and Muslim workers on these dimensions. The results are reported in Table A.17 – I introduce the interacted controls sequentially. Column 4 reports results from the specification with all the controls. Reassuringly, the interaction term Mixed × HD remains negative and significant after the inclusion of these controls. Note that the interactions Mixed × Schooling and Mixed × Tenure are both positive and meaningful in magnitude (though not statistically significant). This suggests that higher tenure and schooling can dampen some of the negative effects of diversity in HD sections. Note also in columns (3) and (4), the coefficients on the interaction term Mixed × LD are statistically significant suggesting that diversity might be costly in LD tasks as well, if workers have very low tenure or schooling.

By design, Muslims workers are only in mixed teams in this experiment. In other words, the treatment of being in a mixed team is perfectly collinear with the presence of Muslims. This was done for two main reasons. First, Muslims comprise of only 18% of all workers in the factory, whereby forming homogeneous Muslim teams would lead to significant loss of statistical power in estimating the effects of religious mixing. Second, at baseline, there were no homogeneous Muslim teams to begin with; therefore experimentally generating such teams could raise ethical concerns. The issue this raises is that Muslim workers may have lower productivity, and this could be driving my findings. However, the null effect of mixing in LD tasks suggests that this is extremely unlikely. Second, there is significant heterogeneity in how mixed teams perform at HD tasks. When Hindus have been in mixed teams with Muslims in the past, I find the negative effects of diversity to be muted significantly (see Table 2.8). If Muslims generally had lower productivity, and especially so at HD tasks, it is unlikely that the negative effects of mixing in these tasks would attenuate so significantly when analyzing heterogeneity *by characteristics of Hindu workers* in mixed teams. These results are discussed in more detail in the following section. Third, the large attenuation of the negative effects over time is also not consistent with Muslims having lower productivity at HD tasks. Finally, I find that Hindus and Muslims were equally like to be promoted as operators or semi-skilled workers at the factory at baseline (see Table A.9). Since the skill-designation of workers affect salary, this suggests that the firm does not perceive Hindus and Muslims to be differentially productive either.[36]

New versus old teammates

One might be concerned that the finding that religious diversity negatively affects productivity is driven in part by the difficulty of working alongside new co-workers, as opposed to the frictions that arise when working alongside non-coreligionists. This would be problematic if the share of new co-workers was not balanced between HD- and LD-Mixed teams, as well as between mixed teams (HD or LD) and Hindu-only teams.

I formally reject this possibility in the results reported in Table A.18. These are individual worker-level regressions where the outcome variable is the proportion of workers in one's current team (randomized team) that were also in their line-section-level team pre-randomization. The mean of the outcome variable is 0.34, which is expected since workers in each production line-section were randomized between three different cohorts – whereby roughly a third of the workers would be known to each other after new teams were formed. Importantly, as shown in Columns (1) and (2), the proportion of new workers is balanced across mixed and non-mixed line-section-level teams. Further, the interactions Mixed × HD and Mixed × LD are small in magnitude and not statistically significant. This suggests that the findings in this paper do not simply result from the inability of workers to coordinate with new colleagues, since workers on average had the same proportion of new teammates irrespective of treatment status.

[36]Further, the results in Table 2.7 (with only Hindus), do not suffer from this collinearity problem. They suggest that an environment that forces people to learn to work together is important to alleviate group-level differences. The positive effects on attitudes of Hindu workers would be unlikely if Muslims did not perform well at HD tasks.

Treatment status and section changes due to randomization

The randomization process involved moving 7.9% of the workers from their original sections (tasks) at baseline so that the line-level team structures in Figure 2.3 could be achieved. While this is a small share of workers, it is nevertheless important to show that treatment status is not correlated with the probability of section-switching. If that were the case one could argue that the treatment effects are potentially contaminated. For example, if mixed HD teams have a greater share of workers who changed their sections, it is possible that it is in fact the time required to adjust to new tasks that explains the results. To rule this out, in Table A.19, I regress a dummy denoting whether the section (task) of a worker was changed due to randomization, on the treatment dummies. In columns (1) and (2), I include only a dummy for whether the team is religiously mixed or not (Mixed) and then in columns (3) and (4) I include its interactions with section type (HD or LD). I include line × baseline section effects in columns (1) and (3) and line × section effects in columns (2) and (4). The coefficients across the different specifications are small and not statistically significant. Only in column (4), the coefficient on Mixed × HD is negative and marginally significant, suggesting that the probability a worker switched their baseline section is actually marginally *lower* for those in mixed HD sections. This exercise therefore rules out the possibility that the treatment effects are driven by differential rates of section-switching across treatment arms during the randomization process.

2.6 Mechanism

The three main findings of this experiment are: 1. Religious mixing leads to lower team output but only in HD tasks, 2. Output differences between mixed and non-mixed HD teams attenuate over time and 3. Attitudes of Hindus towards Muslims improve from mixing in HD tasks but not in LD. In this section, I first discuss plausible mechanisms behind these core findings and then my favored explanation through the brief outline of a model. The predictions of the model are borne out by the data and these empirical tests also help refute the other explanations.

2.6.1 Assortative (mis)matching in complementary tasks

If Muslims have lower productivity, positive assortative matching (only all-Hindu and all-Muslim teams) would be the output maximizing allocation of workers in HD tasks. While this can explain the static results of mixing, there must additionally be on the job learning or skill transfer from Hindus to Muslims in this framework to explain the dynamic results. However, evidence presented in section 2.5.3 already rules out that Hindus and Muslims are differentially productive.

2.6.2 Communication

Religious mixing could also lead to lower output in HD tasks due to pure communication problems amongst Hindus and Muslims. And over time, improving communication can cause production gains. An important strength of my setting is that there are no linguistic differences amongst religious groups — majority of the workers in my sample are born in the same district and speak the same language. It is therefore unlikely that the inability to communicate effectively with non-coreligionists is the key channel either.[37]

2.6.3 Favored mechanism: Minority-stereotyping and discrimination

Having ruled out Hindu-Muslim differences in productivity and/or communication breakdown as primary channels, I focus on stereotyping and discrimination as the potential main mechanism. I present a conceptual framework (the full model is presented in section 2.10 at the end of the chapter) of minority-stereotyping to rationalize the core results (by the majority) and present its empirical tests.

Outline of the conceptual framework

A key distinction is made between Hindu and Muslim workers in this framework based on the asymmetry in their exposure to non-coreligionists at baseline. A large section of Hindu workers at the factory have never worked with Muslims, while 100% of the Muslim workers have worked with Hindus (recall Figure 2.1). Based on this, together with evidence on discrimination against Muslims in access to education and labor markets in India (Kalpagam et al., 2010; Basant, 2007),[38] I assume that Hindus (mistakenly) believe

[37] Consistent with this, I find that baseline contact (self-reported) with Muslims outside of work (for Hindus) does not mitigate the negative effects of mixing in HD tasks (see Table A.24). The hypothesis here is that greater contact with non-coreligionists outside of work might make individuals more effective communicators with them.

[38] In fact, Muslims in my sample have significantly lower schooling than Hindus (Table A.9).

Muslims have lower productivity. Muslim workers do not make this distinction between in-group and out-group workers. This asymmetry in baseline priors leads to multiple equilibria in HD interactions due to complementarities in the production function.[39]

Workers interact in teams for a given length of time and can exert high or low effort, with high effort being more costly. Hindus (depending on past exposure) start off with the belief that Muslims may not be capable of high effort: in other words with some probability Hindus believe Muslims are a behavioural type who always exert low effort (stereotyping).[40] Hindus and Muslims are identical in all other aspects: both are capable of high effort and both face the same cost of effort. If Hindus assign a small probability to Muslims being capable, they optimally exert low effort in HD tasks, since the cost of high effort is too large given their belief. Hindu workers use Bayes rule to update their prior based on their own effort and realized output (teammate effort is not directly observed). Muslim workers understand that they are being underestimated and can "invest" in shifting their Hindu teammate's prior.

If Hindus exert low effort, it is a static best response for their Muslim teammates to also exert low effort in HD tasks. However, I show that there is an equilibrium where given a fixed remaining interaction length, Muslims invest in shifting beliefs of their Hindu teammates by exerting high effort (even as Hindus exert low effort) if and only if Hindu workers' priors are not below a certain threshold value. The intuition behind this is that if Muslims exert high effort, Hindus observe *greater* realizations of high output events than expected given their belief; and therefore they gradually update. Once beliefs of Hindus are high enough such that they find exerting high effort optimal, a high-output static equilibrium is coordinated on. Muslims only find their investment in the majority group worthwhile if initial Hindu beliefs are not too low; since transition to the high-output equilibrium must occur early enough such that Muslim workers' initial investment cost is compensated for by sufficient periods of high-output payoff.

Empirical tests

An important implication of the model is that Hindu workers with a high initial belief that Muslims also exert high effort are less likely to discriminate against them. Therefore, Hindu workers who in the past

[39] In HD tasks, the joint effort of all workers determines the likelihood of high (and low) output. In LD, total output is modelled as the sum of individual expected output (output is still a stochastic function of individual effort) and therefore the priors of Hindu workers are inconsequential in determining effort.

[40] Hindus think Muslims may be behaviourally disposed to exerting low effort or face infinite cost of high effort — in terms of the model these two are equivalent.

have had Muslim co-workers (for a sufficiently long period of time), should continue to optimally exert high effort based on their higher priors, when randomized into a mixed HD team. At baseline, I collected data on a range of different factors that can help directly test this hypothesis. I have details of the team each worker was in before the intervention which allows me to determine the degree of past contact that they have had. In addition, I collected data on political preferences of workers at baseline. These specifically relate to factors that could capture anti-Muslim sentiments.

I test how these factors affect the performance of mixed teams in HD tasks[41] in Table 2.8. In column (1), I first show that there is an overall attenuation of the (negative) effects in HD tasks over time. In columns (2) and (3), I split the sample into the following parts: those teams in which Hindus (on average) have had above median contact with Muslims at baseline and those where they have had below median contact. Consistent with the model, the negative effect of diversity on team output is concentrated in the second group, while the effect on the former is small and not statistically significant. In columns (6) and (7),[42] I split the sample based on measures of political preference of Hindus to capture a different set of stereotypes against Muslims. I use support for the (Hindu) majoritarian BJP party and the NRC (which together with the CAA has been widely regarded as discriminating against Muslims), proponents of which might hold the stereotype that Muslims are inherently less patriotic and committed to the cause of the nation and its growth (Banerjee, 1991).[43] The results follow the same pattern as for baseline exposure. The heterogeneity in the attenuation of the effects by characteristics of Hindus is also apparent in Columns (2)-(7). Mixed teams in which Hindus have either not worked with Muslims in the past, or have low tenure, or strongly support the BJP/NRC suffer larger losses initially from mixing and the effects do not completely dissipate, while the effects are smaller initially and dissipate entirely if Hindus do not have these characteristics/preferences. This is also consistent with the theory.

A key feature of the model is that the minority group "invests" in the majority group to ameliorate negative stereotypes about them. In Table 2.9, I restrict attention to only mixed teams (across HD and LD

[41] I also study how these factors affect performance of mixed teams in LD sections. The effects are small in magnitude and not statistically significant. These results are presented in Table A.20.

[42] In columns (4) and (5), I split the sample based on tenure at the firm and find the effects to be driven by workers with low tenure.

[43] In terms of beliefs about effort, Hindus with these preferences might hold the prior that Muslims do not care enough about effort at work. In general, in the baseline data I find that Hindus who have had greater contact with Muslims are less likely to support the NRC or report to favour the BJP. These correlations are consistent with the results in Table 2.8.

sections) and use dummies for the religion of the respondent, that of the person being referred to in the survey question, and their interaction as the main regressors. Columns (1), (3) and (5) show that while Muslim workers are more likely to be identified as not contributing effort, be blamed and have fewer co-workers willing to give up relief time for them (over the entire intervention), they themselves are less likely to criticize their co-workers.

The coefficients on the interaction terms introduced in Columns (2), (4) and (6) show that the criticism of Muslim workers come largely from their Hindu counterparts, while Muslim workers are actually willing to give up relief time for Hindu co-workers with a higher probability (than for Muslim co-workers). This decomposition lends support to the idea that it is indeed stereotyping of minorities by the majority group which results in lower team cohesion and output initially; while the minority group initiates the integration process.[44]

Taken together, these results support minority-stereotyping and discrimination as the primary mechanism behind the core results. They also provide evidence against the hypothesis that Muslims may have lower productivity, or that the negative effects of mixing in HD tasks result from communication issues. These latter explanations are inconsistent with the null effects in mixed teams in which Hindus have less stereotypical attitudes towards Muslims.

2.7 Policy discussion: Firm supervisor survey

Do firm supervisors understand the costs of diversity and how they depend on the production function? Can they predict the findings from this experiment, and if so, do they suggest integration of workers only in LD tasks or do they recommend other management practices to ameliorate possible negative effects of religious mixing in HD tasks? To analyze these policy relevant questions, I surveyed supervisors and operators (personnel with some leadership role) of five different processed food manufacturing plants in April, 2021.

Participants were first asked to denote which of the two tasks (HD or LD in Figure A.8): (1) requires greater continuous coordination and communication amongst co-workers and (2) is likely to cause more

[44] In Table A.21, I further decompose the findings of Table 2.9 into HD and LD sections and show that the effects discussed above are driven by HD sections and less so by LD sections.

frictions and arguments amongst workers. They picked the HD task more frequently for both of these questions (Figure A.13). Interestingly, while close to 80% of the supervisors chose the HD task for (2), a fair share of them also picked either the LD task (17%) or mentioned both HD and LD (35%) for (1). This reiterates an important point about the mechanism behind the core results in this paper: workers do not simply sabotage or undermine the efforts of their out-group members, which is possible to do in LD tasks as well. Rather, a negative perception of out-group members causes frictions which are costly when working in production environments that require workers to be significantly dependent on each other.

Participants were then asked to predict whether a religiously homogeneous or mixed team would be more productive in each task. They were informed that I have conducted an experiment to test this and that they would be rewarded with Rs 25 (about 30% of their hourly wage) if their answer matches with my findings— this was meant to reduce social desirability bias in the answers. Between 40%-45% of the supervisors mentioned that religiously mixed teams would be more productive in both tasks (Figure A.14). This could be because of social desirability bias or as I show next, supervisors actually think of issues beyond direct productivity arising from segregation of workers (by religion), which prompts them to answer in this manner. Nevertheless, a significantly higher share of respondents mentioned that a homogeneous team would be more productive at the HD task (30%) than the LD task (8%). Overall, about a third of the supervisors predicted correctly that homogeneous teams would be more productive in HD tasks and about half of them correctly mentioned that mixing would be inconsequential in LD tasks.

While it is possible that a large share of supervisors do not understand the costs of diversity and consequently do not segregate workers, it is also possible that there are additional costs that do not justify segregation. To understand this systematically, respondents were finally asked if they are willing to segregate workers by religion and/or age if workers do not perform well as a team because of these differences. I use age as a natural benchmark because in the Indian context age differences could be an important source of conflict amongst teammates. The supervisors generally seem to be averse to segregation on either dimension, but they are especially opposed to segregation by religion (Figure A.15), despite the potential for losses. About a quarter of the supervisors correctly (as I find) cite negative effects of diversity dissipating over time as their reason. However, the first order concern is about segregation actually raising tensions further. Informal conversations with supervisors suggest that some of the concerns they have in mind are with respect to such segregation creating a hostile environment in common areas of interaction (canteen, tea room), in addition to tensions on the production floor.

In sum, this survey shows that roughly between a third to a half of the supervisors correctly predicted the results of the intervention. However, it is clear that despite the possibility of losses, the majority of supervisors are averse to segregation of workers by religion. Many of them, being aware of the long-term gains from repeated contact are willing to trade off potential short-run costs of non-segregation to productivity. But they are also concerned about costs to segregation that are typically hard to identify as a researcher by simply analyzing production data. This is perhaps why previous studies find it difficult to reconcile productivity losses arising from diversity with non-segregation of workers in the firms they study (Hjort, 2014).

Speculatively, the short-run but significant cost associated with organizing HD production with an ethnically diverse workforce may impede firms (especially those with low working capital) from growing and adopting complex assembly-lines processes in developing countries (Kremer, 1993; Hsieh and Olken, 2014). Since the manufacturing sector is likely to have a larger concentration of HD tasks (relative to agriculture or services), this could be one potential explanation for why ethnically fractionalized countries have lower share of employment in manufacturing as well as firms with lower foreign technology take up – even after controlling for income (Figure 2.6).

The implications of the interaction between ethnic diversity and production technology for broader economic change can be an important avenue for future research.

2.8 Conclusion

My findings suggest that both the nature and duration of contact are important in understanding how religious diversity in firms may impact productivity. An environment that makes workers highly dependent on each other creates incentives for them to invest in building social capital with out-group members. This brings about positive changes in attitudes as well as productivity gains over time, but it might be unprofitable in the short-run through lost output. Overall, my results suggest a potential tension between the goal of maximizing short-run productivity and that of improving intergroup relations. More speculatively, they might help explain why in equilibrium there can be a lot of integration at work without intergroup relations improving – the integration might only occur in contexts where intergroup contact is socially ineffective.

Beyond conceptual contributions, this chapter has a few important implications for policy. First, firms with high-dependency production should minimize team switching in order to mediate possible negative effects of diversity. Second, in firms with low-dependency production, exposure to non-coreligionists might not necessarily reduce negative outgroup attitudes. While this might cost the firm little in terms of lost output, a less cohesive work culture can lead to problems outside of daily production. Such firms might benefit from additional measures to ensure a collaborative environment for workers to interact in. This can even be achieved outside the workplace, for example though sports teams (Lowe, 2021; Mousa, 2018). In general, an open question remains whether that can also lead to productivity gains at the workplace. If that is indeed the case, the cost to output from mixing workers in HD tasks to integrate them could be avoided.[45] However, if belief updating with respect to specifics about co-workers' effort levels at work is the driving factor (as suggested in the theoretical framework), contact outside the firm might not be able to entirely mitigate the negative effects of diversity.

As economies undergo structural transformation, the nature of economic production changes, which potentially influences the type of inter-ethnic interactions. My results suggest that the costs of diversity may increase as economies move away from traditional agriculture to manufacturing activity and then decrease with the transition to services. In traditional agricultural societies, land cultivators largely work in LD environments with limited contact with new people, but manufacturing activity involves a higher share of HD work (construction work, small firms etc.), as well as contact with new people on a regular basis, making diversity costly. In services, with a comparatively higher share of LD work and a regular set of colleagues, these costs might be low again. One important aspect of this is that identity diversity might act as a hindrance in transitioning to formal manufacturing work by perpetuating discrimination amongst groups (A.Churchill and Danquah, 2020).

Finally, the finding that minorities (Muslims) bear the cost of integration in this context is general-

[45] Of course arranging such organized collaborative contact might itself be costly for a firm in terms of time and resources.

izable to many other settings, especially the U.S. For example, the argument that African-Americans are rewarded less for their effort (relative to the average American), requiring them to work harder to achieve similar career goals (DeSante, 2013) or the finding that Asian immigrants in the U.S., being aware of their unequal racial status, *work twice as hard* as a normative path to success and assimilation by achieving model-minority status (Zhou, 2004 and Zhou and Xiong, 2005), relate closely to my results in the Indian context. Overall, this implies that minority (the oppressed) groups, despite being discriminated against, may play a crucial role in the process of nation-building through initiating economic and social integration in diverse societies. But the above statement must be caveated – we still have much left to understand with respect to if/how this would also translate into assimilation in the sense that they do not feel a divide between participation in mainstream institutions and cultural practices.

2.9 Tables and figures

2.9.1 Tables

Table 2.1: Proportion Muslim by line-level team and cohort (at baseline)

Line	Cohort 1	Cohort 2	Cohort 3	Average
Line 1	0.30	0.23	0.19	0.24
Line 2	0.11	0.07	0.27	0.15
Line 3	0.24	0.24	0.13	0.20
Line 4	0.22	0.26	-	0.23
Line 5	0.15	0.19	0.07	0.14
Line 6	0.14	0.07	0.23	0.15
Average	0.19	0.18	0.18	0.18

Note: Each production line (apart from Line 4) has a total of three cohorts working on it in each of the three shifts in a day. This table reports the share of Muslim workers in each line (across all sections) at baseline – for all three cohorts. Please note the total number of workers in each line-cohort is the same.

Table 2.2: Characteristics of High- and Low-Dependency tasks

Work condition	High-Dependency (HD)	Low-Dependency (LD)
Task coordination	High and Continuous	Low and Intermittent
Control over breaks	Low	High
Physical mobility	Restricted	Good
Repetitive monotony	High (Machine Speed)	Low (Occasionally paced by machine)

Note: This table lists some key differences between work characteristics of High- and Low-Dependency tasks.

Table 2.3: Randomization check

	Panel A: Outcomes relevant at work				Panel B: General characteristics and attributes				
	Tenure	Muslim co-workers *Hindus*	Taking Orders	Communicating	Age	Schooling	Trust	Altruism	Inter-religious contact outside work
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mixed × LD	0.0623	0.0277	0.0496	0.0857	1.5012	-0.2158	0.5354	0.0280	0.0371
	(0.3378)	(0.0219)	(0.0556)	(0.0525)	(1.4079)	(0.5013)	(0.3505)	(0.2203)	(0.0475)
Mixed × HD	-0.0042	0.0169	-0.0027	-0.0347	0.7232	0.3499	-0.0883	-0.0495	0.0088
	(0.3259)	(0.0172)	(0.0471)	(0.0481)	(0.8184)	(0.3619)	(0.3015)	(0.1649)	(0.0475)
p(Mixed × LD = Mixed × HD)	0.87	0.70	0.44	0.07	0.60	0.33	0.14	0.75	0.65
Mean Dep Var.	4.45	0.12	0.73	0.53	34.47	7.84	3.79	6.65	0.45
Line × Section F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Religion F.E.	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	586	478	586	586	586	586	586	586	586
Adj. R^2	0.122	0.029	0.012	0.045	0.076	0.070	-0.005	-0.013	0.099

* p<0.10, ** p<0.05, *** p<0.010. The unit of observation is an individual worker. Standard errors clustered at the line-section-level team. "Tenure" and "Schooling" are measured in years and as highest grade completed respectively. "Taking Orders" is a dummy variable coded 1 if the respondent reported to be always comfortable taking orders from non-coreligionists and 0 if they reported to be sometimes or always uncomfortable. "Communicating" is coded 1, 0.5 and 0 for the responses "Always comfortable", "Sometimes uncomfortable" and "Always uncomfortable" when asked about being comfortable communicating with non-coreligionists. Survey questions on "Trust" and "Altruism" are used from the World Value Survey (WVS). The dependent variable "Inter-religious contact" refers to the degree of cross-religion interaction that workers had at baseline, outside of work. The variable is coded 1, 0.5 and 0 if a worker mentioned that during the daily course of their life they: 1) interact with more than 5 non-coreligionists 2) interact with 1 to 5 non-coreligionists, or 3) do not interact with anyone outside their religion, respectively.

Table 2.4: Treatment effect on line-level output

	(1) **Log Output (Pieces)**	(2) **Log Output (Boxes)**
HD-Mixed vs LD-Mixed Line	-0.0487***	-0.0519**
	(0.0163)	(0.0220)
Bootstrap (Wild Cluster) C.I.	[-0.093, -0.013]	[-0.107, 0.007]
Day F.E.	Yes	Yes
Shift F.E.	Yes	Yes
Production Line F.E.	Yes	Yes
Mean Dep Var.	10.80	6.97
	(1.24)	(0.943)
N	1045	1045
Adj. R^2	0.722	0.644

* p<0.10, ** p<0.05, *** p<0.010. Observations are daily output produced by line-level teams. Standard errors clustered at the line-level team in parenthesis. Wild cluster bootstrap (Cameron et al., 2008) confidence intervals in square brackets. HD-Mixed Line is a dummy coded 1 for a line-level team with all HD sections religiously mixed and LD sections non-mixed, and 0 for exactly the opposite line-level structure (LD-Mixed Line).

Table 2.5: Treatment effect on section ratings

	Rating (Raw)		**Rating > Median**	
	(1)	(2)	(3)	(4)
Mixed	-0.0204*		-0.0254***	
	(0.0119)		(0.00899)	
Mixed × LD		-0.0067		-0.0047
		(0.0144)		(0.0121)
Mixed × HD		-0.0349**		-0.0474***
		(0.0185)		(0.0121)
p(Mixed × HD = Mixed × LD)		0.229		0.011
Mean Dep. Var.	3.82	3.82	0.44	0.44
	(0.83)	(0.83)	(0.50)	(0.50)
Education and Tenure Controls	Yes	Yes	Yes	Yes
Day F.E.	Yes	Yes	Yes	Yes
Shift F.E.	Yes	Yes	Yes	Yes
Line × Section F.E.	Yes	Yes	Yes	Yes
N	6909	6909	6909	6909
Adj. R^2	0.600	0.600	0.358	0.358

* p<0.10, ** p<0.05, *** p<0.010. Observations are daily ratings received by line-section-level teams. Standard errors clustered at the line-section-level team. "Mixed" is a dummy variable coded 1 if the line-section-level team is religiously mixed. Line × Section fixed effects are included in the all specifications; as a result the main effect of HD versus LD is not separately identified in columns (2) and (4). Education and tenure control for the mean of schooling and tenure of workers in the line-section-level team.

Table 2.6: Treatment effect on worker interactions

	Identified teammate as contributing low effort		**Blamed by teammate**		**Unwilling to give up relief time**	
	(1)	(2)	(3)	(4)	(5)	(6)
Mixed	0.0420***		0.0400**		0.0640*	
	(0.0137)		(0.0158)		(0.0365)	
Mixed × LD		0.0317		0.0817***		0.0339
		(0.0226)		(0.0228)		(0.0563)
Mixed × HD		0.0445***		0.0301*		0.0719*
		(0.0154)		(0.0175)		(0.0423)
p(Mixed X HD = Mixed X LD)		0.62		0.05		0.57
Mean Dep. Var	0.14	0.14	0.08	0.08	0.25	0.25
Worker Skill F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Religion F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Line × Section F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3696	3696	3684	3684	3727	3727
Adj. R^2	0.016	0.016	0.013	0.014	0.072	0.072

* p<0.10, ** p<0.05, *** p<0.010. Observations are at the worker-teammate level for line-section-level teams i.e. there are (N-1) observations per worker, where N denotes the number of workers in the section. Standard errors clustered at the line-section-level team. “Mixed" is a dummy variable coded 1 if the line-section-level team is religiously mixed. Line × Sections fixed effects are included in the all specifications; as a result the main effect of HD versus LD is not separately identified in columns (2), (4) and (6). Workers were asked to choose teammates who they: (1) think have not contributed sufficient effort at any point during the intervention (2) have been blamed by during the intervention and (3) would give (or already have) up their relief time for.

Table 2.7: Treatment effect on attitudes at endline: Hindus

	Attitudes towards Muslims					
Comfortable:	**Taking Orders**		**Communicating**		**Co-working**	
	(1)	(2)	(3)	(4)	(5)	(6)
Mixed	0.1249***		0.0985**		0.1145***	
	(0.0448)		(0.0403)		(0.0348)	
Mixed × LD		0.0180		-0.0781		0.0198
		(0.0778)		(0.0618)		(0.0626)
Mixed × HD		0.1866***		0.2004***		0.1691***
		(0.0555)		(0.0406)		(0.0439)
p(Mixed × HD = Mixed × LD)		0.086		0.000		0.076
Mean Dep. Var.	0.74	0.74	0.49	0.49	0.61	0.61
Sample	Baseline Mean		Baseline Mean		Endline Non-mixed Mean	
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Worker skill F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Line × Section F.E.	Yes	Yes	Yes	Yes	Yes	Yes
N	448	448	448	448	448	448
Adj. R^2	0.066	0.072	0.066	0.088	0.063	0.068

* p<0.10, ** p<0.05, *** p<0.010. The unit of observation is an individual worker. Standard errors clustered at the line-section-level team. "Mixed" is a dummy variable coded 1 if the line-section-level team is religiously mixed. The main effect of HD versus LD is not separately identified in columns (2), (4) and (6) because Line × Section fixed effects are included. "Taking Orders" is a dummy variable coded 1 if the respondent reported to be comfortable taking orders from Muslims, and 0 otherwise. "Communicating" is coded 1, 0.5 and 0 for the responses "Always comfortable", "Sometimes uncomfortable" and "Always uncomfortable" respectively, when asked about being comfortable communicating with Muslims. For "Co-working" the outcome is coded 1,0.5 and 0 for the responses "Mixed team", "Indifferent" and "Hindu-only team" when asked about respondents' preferred team type for future changes.

Table 2.8: Heterogeneous attenuation by characteristics of Hindus at baseline (HD section ratings)

Sample:	**Full**	**Contact at Baseline**		**Tenure at Baseline**		**Support for BJP and NRC**	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Above Median	Below Median	Above Median	Below Median	Above Median	Below Median
Mixed × 0-60 days	-0.0661*	-0.0458	-0.1816***	0.0202	-0.0953**	-0.1248***	-0.0496
	(0.0391)	(0.0498)	(0.0595)	(0.0800)	(0.0431)	(0.0438)	(0.0619)
Mixed × 61-120 days	-0.0355	-0.0195	-0.0802**	0.0497	-0.0319	-0.0792***	0.0177
	(0.0234)	(0.0429)	(0.0337)	(0.0500)	(0.0325)	(0.0242)	(0.0712)
Mean Dep. Var.	3.86	3.83	3.89	3.84	3.88	3.87	3.83
Education and Tenure Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Shift F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Line × Section F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3466	1884	1582	1462	2004	2384	1082
Adj. R^2	0.609	0.633	0.596	0.605	0.620	0.602	0.631

* $p<0.10$, ** $p<0.05$, *** $p<0.010$. Observations are daily ratings received by line-section-level teams. Standard errors clustered at the line-section-level team. “Mixed" is a dummy variable coded 1 if the line-section-level team is religiously mixed. In column (2), the sample consists of all line-section-level teams in which the share of Muslim teammates, that Hindus in that team had at baseline, is above median. In column (3), the sample consists of all line-section-level teams in which the share of Muslim teammates, that Hindus in that team had at baseline, is below median. In columns (4) and (5) teams are split by median tenure of Hindus at baseline. In column (6), the sample consists of all line-section-level teams with above median support for the BJP or the NRC (averaged across all Hindu workers in the team). In Column (6), the sample consists of all line-section-level teams with below median support for the BJP or the NRC (averaged across all Hindu workers in the team).

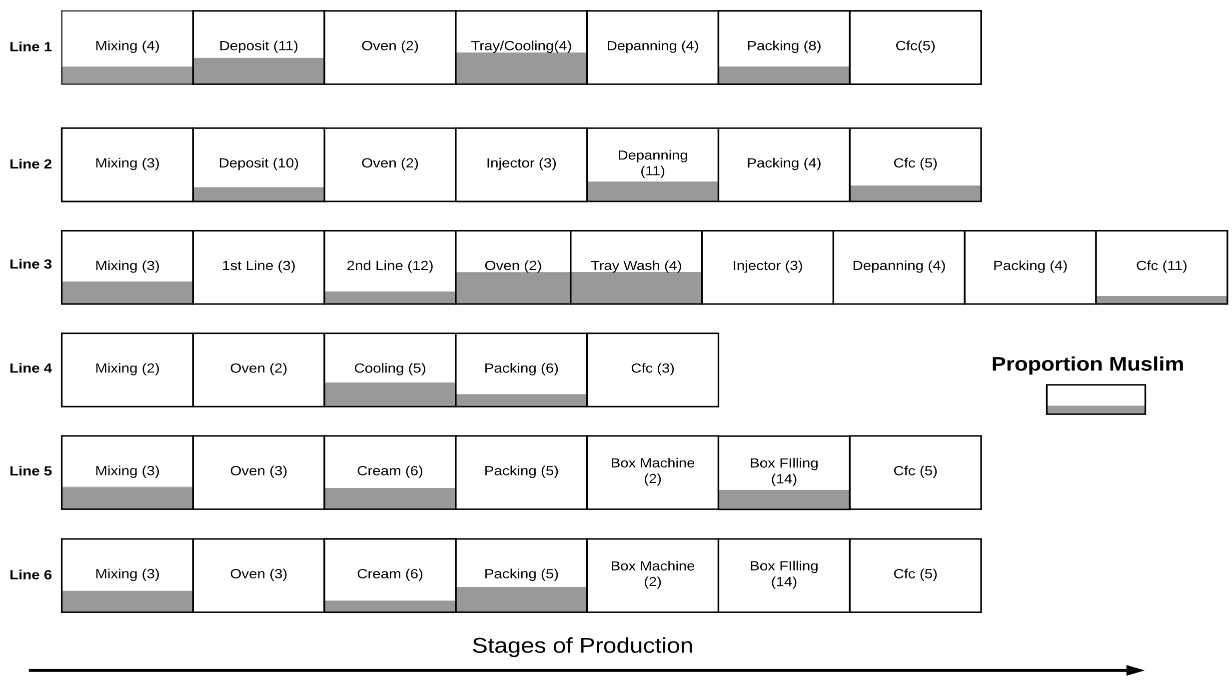
Table 2.9: Treatment effect on worker interactions: Decomposition (Mixed teams)

	Identified teammate as contributing low effort		**Blamed by teammate**		**Unwilling to give up relief time**	
	(1)	(2)	(3)	(4)	(5)	(6)
Target Muslim	0.0528***	0.0861***	-0.0159	0.0048	0.0474***	0.0090
	(0.0175)	(0.0225)	(0.0126)	(0.0213)	(0.0176)	(0.0332)
Respondent Muslim	-0.0172	0.0152	-0.0009	0.0192	-0.0450	-0.0830**
	(0.0221)	(0.0266)	(0.0199)	(0.0291)	(0.0316)	(0.0356)
Target Muslim ×		-0.0995**		-0.0612		0.1139*
Respondent Muslim		(0.0485)		(0.0406)		(0.0657)
Mean Dep. Var	0.15	0.15	0.10	0.10	0.28	0.28
Worker Skill F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Line × Section F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2033	2033	2029	2029	2035	2035
Adj. R^2	0.018	0.025	0.013	0.016	0.064	0.084

* p<0.10, ** p<0.05, *** p<0.010. Observations are at the worker-teammate level for line-section-level teams i.e. there are (N-1) observations per worker, where N denotes the number of workers in the section. Standard errors clustered at the line-section-level team. Workers were asked to choose teammates who they: (1) think have not contributed sufficient effort at any point during the intervention (2) have been blamed by during the intervention and (3) would give (or already have) up their relief time for.

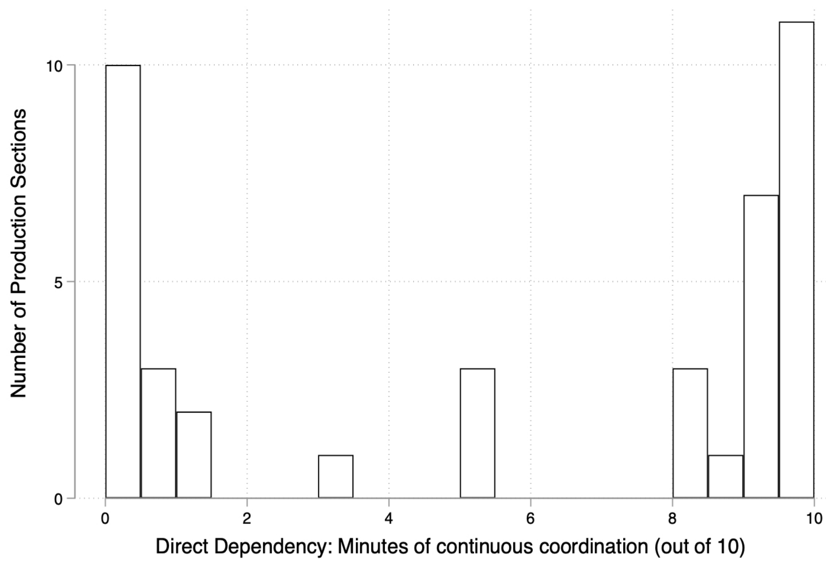
2.9.2 Figures

Figure 2.1: Structure of production lines



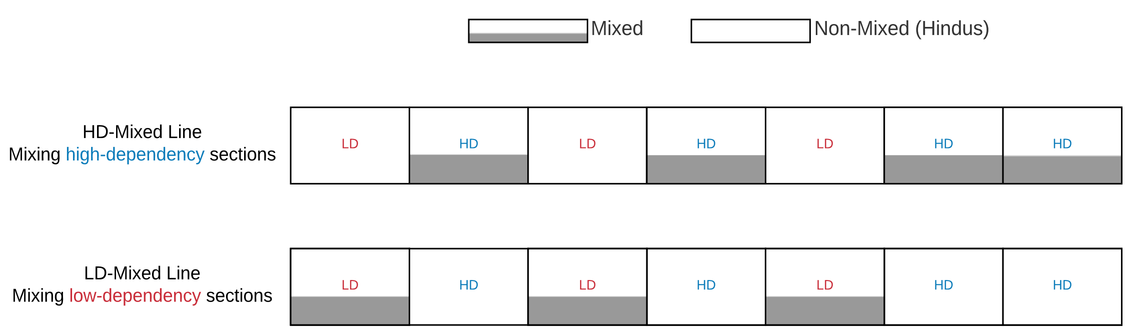
Note: This figure shows the structure of all six production lines in the factory. The numbers in parentheses denote the count of workers in each section per cohort. Each production line has three cohorts working on it in each of the three shifts in a day. The color shades denote the proportion of Muslim workers in each section in one particular cohort at baseline. Please refer to figure A.5 in the Appendix for this figure without the color shades.

Figure 2.2: Distribution of Direct Dependency



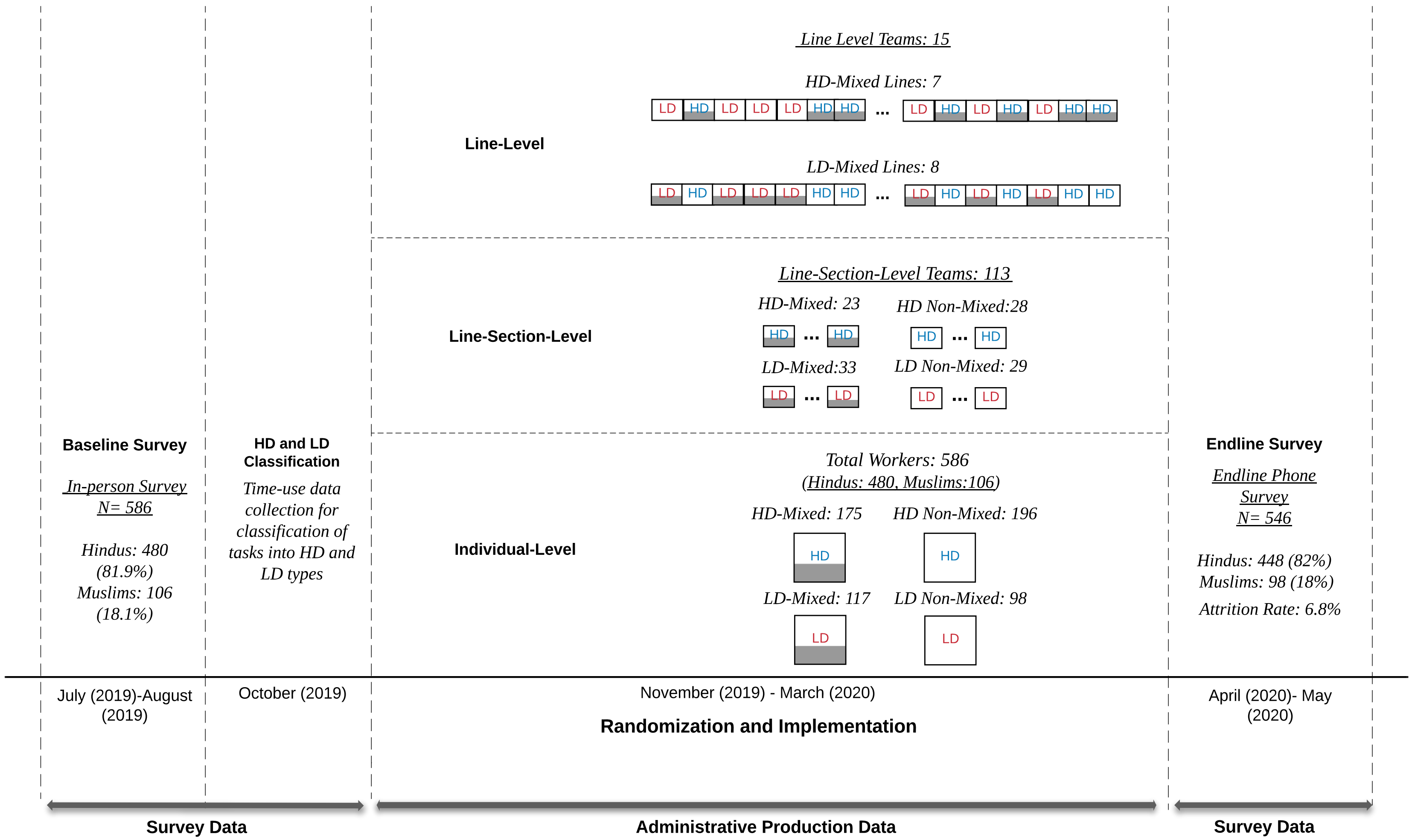
Note: This figure shows the distribution of Direct Dependency. Enumerators visited every section of each production line and took stopwatch measures of the number of minutes (out of 10) for which workers were continuously dependent on each other for production to occur. The figure is generated from these stopwatch records at the line-section-level.

Figure 2.3: Randomized team structure



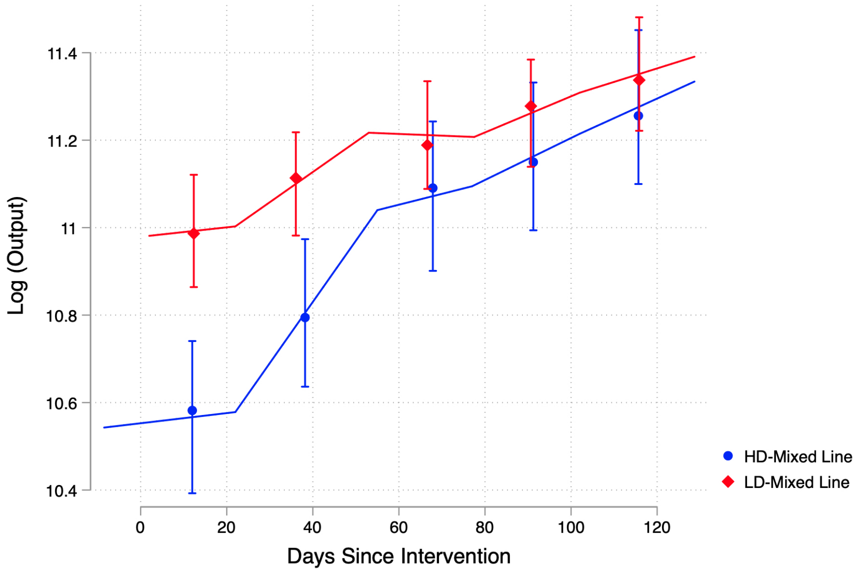
Note: This figure shows the two different types of line-level teams after randomization. Sections are partially shaded to denote mixed teams. HD-Mixed lines had all their HD sections mixed and LD sections non-mixed. The opposite is true for LD-Mixed lines.

Figure 2.4: Experimental design and timeline



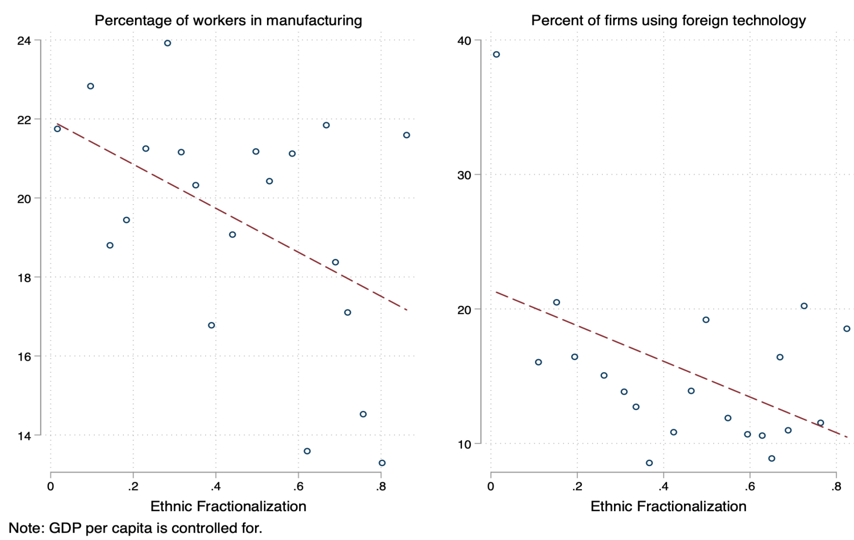
Note: Shaded boxes denote mixed teams. The share of Muslim workers in mixed teams is between 35%-40% (balanced across HD and LD sections). This diagram shows the timeline of the intervention. The baseline survey was completed between July and August in 2019. Time-use data in order to classify tasks into HD and LD types were collected in October 2019. The experiment was conducted between November 2019 and March 2020. Details of sample size by treatment arms are presented in the figure. A phone survey was conducted at endline in April and May of 2020 due to COVID-19 related restrictions.

Figure 2.5: Treatment effect on line-level output (Event study)



Note: This figure is generated from binned regressions (this plot is created using the STATA command binsreg, which implements binscatter estimation with robust inference proposed in Cattaneo et al. (2019)). using exactly the same controls variables as in Table 2.4. The treatment period is divided into 5 equal sized bins. The outcome variable is output produced in pieces (logged). Bars denote 95% confidence intervals.

Figure 2.6: Ethnic diversity and the manufacturing industry



Note: Data on ethnic fractionalization come from the Historical Index of Ethnic Fractionalization (HIEF) dataset. Data on share of workers in manufacturing and percentage of firms using foreign technology are obtained from the World Bank Enterprise Survey.

2.10 Full Model

This section presents the theoretical framework. The primary objective of the model is to rationalize the core empirical results, especially the mechanism behind the attenuation in output losses in HD-mixed sections over time. The model makes predictions specially with respect to heterogeneous treatments effects based on worker characteristics, which I subsequently test for in the data.

A key distinction is made between Hindu and Muslim workers in this framework. Consistent with majority-minority relations, a large section of Hindu workers at the factory have never worked with Muslims in the past, while 100% of the Muslim workers have worked with Hindus. Based on this asymmetry in exposure at baseline, together with the evidence on discrimination against Muslims in access to education and labor markets in India (Kalpagam et al., 2010; Basant, 2007),[46] I assume Hindus (mistakenly) on average believe that Muslims are not as hardworking as them.[47] Muslim workers do not make this distinction between in-group and out-group workers given that they have had much greater contact with Hindus. This asymmetry in baseline priors leads to multiple equilibria in HD interactions due to complementarities in the production function, while in LD this does not matter.

As mentioned, Muslims have accurate priors about Hindus, while Hindus might not necessarily have them, it will depend on past exposure. Muslims are aware that they are being stereotyped by Hindus but can "invest" in shifting the priors of Hindus.

2.10.1 Setup

Production is composed of two types of tasks, HD (high-dependency) and LD (low-dependency). I make the following assumptions about the production process:

1. There are two workers in each type of task (generalizes to multi-worker easily)

2. There are two types of output: High (O_H) and Low (O_L)

3. Worker effort is the only input in production and it is not observed directly by teammates

[46] In fact, Muslims in my sample have significantly lower schooling than Hindus (Table A.9).

[47] An implicit assumption here of course is that in reality Hindus and Muslims are equally productive. I present direct evidence for this in section 2.5.3.

4. There are two types of effort: High (e_H) or Low (e_L)

5. Output in each task is a noisy function of worker effort

6. Effort is costly: $c(e_H) > c(e_L) = 0$

Assumptions 1,2,4 and 6 are made for simplicity and easily generalizes to settings where output, effort and effort cost are continuous variables and there are multiple workers in each task. There are several factors that workers do not have direct control over which influence productivity and also make perfectly observing teammate's effort difficult. These include machine breakdowns, inadequate raw material planning and unanticipated production stoppages due to supply chain issues. Assumptions 3 and 5 are made based on these factors.

The production function for HD tasks can be written as:

$$y_{HD}(e_{k1}, e_{k2}) = p(e_{k1}, e_{k2})O_H + \{1 - p(e_{k1}, e_{k2})\}O_L \tag{2.4}$$

where e_{ki} denotes effort level $k = (H, L)$ for worker $i = (1, 2)$. $p(e_{k1}, e_{k2})$ denotes the probability of high output (O_H) conditional on effort. Clearly, the joint effort of both workers determines the probability of high output and the marginal value of effort is thus higher in teammate's effort level.

The probability of high output conditional on effort levels of both the workers are:

$$(e_H, e_H) = p_H \tag{2.5}$$

$$(e_H, e_L) = (e_L, e_H) = p_{HL} \tag{2.6}$$

$$(e_L, e_L) = p_L \tag{2.7}$$

where $p_H > p_{HL} > p_L$. The production function in LD tasks is linear in worker efforts and is written as:

$$y_{LD}(e_{k1}, e_{k2}) = \sum_{i=1}^{2}\{p(e_{ki})o_h + (1 - p(e_{ki}))o_l\} \tag{2.8}$$

where o_h and o_l denote high and low individual output levels respectively. In LD sections total output is therefore the sum of individual expected output. The probability of high and low output (conditional on e_H and e_L) are p_H and p_L respectively with $p_H > p_L$.

2.10.2 One shot production

In HD sections, high effort (e_H) is statically preferred by a worker if and only if their teammate also exerts high effort (e_H). In other words, there is no incentive to free-ride when teammate exerts e_H. Mathematically, this condition implies:

$$(p_H - p_{HL})(O_H - O_L) > c(e_H) > (p_{HL} - p_L)(O_H - O_L)^{48} \tag{2.9}$$

In LD sections, value of a worker's effort is not dependent on teammate's effort level. As a result we assume e_H is the dominant action, implied by:

$$(p_H - p_{HL})(O_H - O_L) > c(e_H)^{49} \tag{2.10}$$

2.10.3 Analysis of the model

Workers interact repeatedly for T periods in a task. HD sections are the interesting case here due to complementarity in worker efforts in the production function. I first solve the model for HD sections and subsequently discuss LD sections.

Hindu workers (majority group)

Hindu workers who are in mixed teams could have been in non-mixed ones, which they believe would be more productive. In other words, they assign probability π_t (in period t, initial belief is π_0) on their Muslim teammate exerting high effort, which in the case of coreligionists (other Hindus) is 1. Therefore, with probability (1-π_t) Hindus believe Muslims maybe be "lazy" – a behavioural type that always exerts low effort. This can be thought of as Hindus thinking Muslims have infinite cost of high effort or that they are simply behaviourally disposed to exerting low effort.

[48] This expression is obtained by re-writing: $p_H O^H + (1-p_H)O^L$ - $c(e_H) > p_L O^H + (1-p_L)O^L > p_{HL} O^H + (1-p_{HL})O^L$ - $c(e_H)$.

[49] The expression is $p_h o^h + (1-p_h)o^l$ - $c(e_H) > p_l o^h + (1-p_l)o^l$.

A Hindu worker's problem is then given by:

$$V = max_{(e_t)_{t=0}^{T}} \sum_{t=0}^{T} P_t(e_t, \pi_t) \tag{2.11}$$

where e_t denotes the action in period t (choice variable), π_t is the prior at t (state variable) and $P_t(e_t, \pi_t)$ is the expected (perceived) payoff at time t. Each period, given their current prior, their own action and realized output, the Hindu worker's belief about the effort level of their Muslim teammate is updated. The transition matrix at any period t, with current prior π_t, is:

Table 2.10: Bayesian updating (Hindu workers): Prob(Muslim worker exerts e_H)

Own Effort/Realized Output	O^H	O^L
e_H	$\frac{\pi_t p_H}{\pi_t p_H + (1-\pi_t) p_{HL}}$	$\frac{\pi_t (1-p_H)}{\pi_t (1-p_H) + (1-\pi_t)(1-p_{HL})}$
e_L	$\frac{\pi_t p_{HL}}{\pi_t p_{HL} + (1-\pi_t) p_L}$	$\frac{\pi_t (1-p_{HL})}{\pi_t (1-p_{HL}) + (1-\pi_t)(1-p_L)}$

Note: The prior of a Hindu worker in period t is denoted by π_t.

Muslim workers (minority group)

Unlike Hindu workers, Muslim workers have always been in mixed teams. They are used to being stereotyped in this manner. In other words they are aware that Hindus are operating on incorrect priors.

Muslim workers choose an optimal effort investment path based on the time horizon. At any given time t and set of history s (which determines prior π_t of Hindu teammates), Muslim workers choose an effort level. Their problem can be written as:

$$V = E^{\psi}(\sum_{t=0}^{T} P_t(a_t, s_t) | \pi_o) \tag{2.12}$$

where ψ denotes the mapping from a set of histories (from 0 to $t-1$) to actions a_t = (e_H, e_L) and P_t denotes expected payoff in each period conditional on the set of history and preferred action choice in that period. State s_t (history of high vs low output events) defines the current belief (π_t) of the Hindu worker regarding their Muslim co-worker.

Note that for any t=k, the problem above can be re-written as

$$V_k^{\psi}(s) = \{ \sum_{s' \in S_{k+1}} P_k(s,a) + \mu_k(s'|s,a) V_{k+1}^{\psi}(s') \}, k = N-1, ..., 0 \tag{2.13}$$

where π denotes mapping from each possible history $h_t = (s_o,\ a_o, \ldots, s_{t-1},\ a_{t-1})$ to actions $a_t = \psi_t(h_t)$. $\mu(.)$ denotes the probability of a future state (belief of the Hindu worker) conditional on actions and current state. The optimal effort path for a Muslim worker is then a mapping from state histories to actions ψ^* such that,

$$\psi_k^*(s) \in argmax_{a \in e_H, e_L}\{\sum_{s' \in S_{k+1}} P_k(s,a) + \mu_k(s'|s,a)V_{k+1}^{\psi*}(s')\} \tag{2.14}$$

Markov Equilibrium

It is clear from equation (2.9) that statically there are two equilibria of this game, one where both workers exert e_H and the other where both exert e_L. Since this game is repeated, it is possible to have strategies that are a function of the history of the game, as well as beliefs of Hindu workers. I am not going to rule out the possibility of some complicated equilibria based on such strategies. Instead, I will be looking at an equilibrium where individuals condition behaviour on commonly known beliefs of Hindus.

***Definition**: $\bar{P}$ is the probability that a Hindu believes a Muslim is exerting e_H beyond which it is statically payoff maximizing for a Hindu to also contribute e_H.*

$\bar{P}$ is the threshold value such that if π_t is greater than this value, or in other words if Muslims are believed likely enough to be contributing high effort, Hindus will exert high effort in response. Note that $\bar{P}$ is exogenous and is obtained by comparing net expected payoff to a Hindu worker from exerting high effort versus exerting low effort, given π_t.

For a given value of π_t, the payoff from exerting e_H and e_L are as follows:

$$P(e_H,\pi_t) = \pi_t\{p_H O_H + (1-p_H)O_L\} + (1-\pi_t)\{p_{HL}O_H + (1-p_{HL})O_L\} - c(e_H) \tag{2.15}$$

$$P(e_L,\pi_t) = p_t\{p_{HL}O_H + (1-p_{HL})O_L\} + (1-\pi_t)\{p_L O_H + (1-p_L)O_L\} \tag{2.16}$$

Comparing (2.15) and (2.16) high effort yields high greater payoff iff (2.15) > (2.16), which is the case when

$$\pi_t > \frac{c(e_H) - (p_{HL} - p_L)(O_H - O_L)}{(p_H + p_L)(O_H - O_L)} = \bar{P} \tag{2.17}$$

This gives us $\bar{P}$. From equation (2.9) it can be seen that the numerator in the RHS is positive. Given this, I now proceed to characterizing the equilibrium.

Proposition: *There exists an equilibrium in which, for a given remaining interaction length T, there is a $\pi_T \leq \bar{P}$, such that for $\pi \geq \pi_T$, Muslims will exert e_H. If $\pi < \pi_T$ Muslims exert e_L. Along the equilibrium path, Hindus exert e_L when $\pi < \bar{P}$ and e_H otherwise.*

Proof: A formal is provided at the end of the section, I provide the intuition for the proof here. If π is lower than $\bar{P}$ (at a certain period t), or in other words if the belief of the Hindu worker is not high enough to exert e_H, it is a static best response for the Muslim worker to also exert e_L. However, the Muslim worker can "invest" in shifting priors of the Hindu worker, if it leads to higher payoff in expectation by transitioning to a high output static equilibrium. In order for that to be worthwhile, there must be enough periods in expectation with $\pi > \bar{P}$, such that the cost of exerting e_H (while the Hindu worker exerts e_L) is compensated for and net payoff to the Muslim worker is greater than exerting e_L (in expectation). At time t, given a remaining interaction length T, π_T is the minimum (threshold) value (of π) for e_H to be worth for the Muslim worker.[50] If the belief of the Hindu worker at t (π_t) is below π_T , then not enough interactions are left (in expectation) for the Muslim worker's investment in shifting the beliefs of the Hindu worker to be worth it.[51] e_L is then the best response in that period.

The Hindu worker's belief at time t (π_t) is essentially the state-variable in this Markov equilibrium. The Hindu worker's action along the equilibrium path is thus e_L if in that period $\pi_t < \bar{P}$ and e_H otherwise.

Additional Implications

I now note a few additional implications of this model which are empirically testable.

1. *On average (during the intervention), Hindus would blame low output on Muslims, Muslims would not blame Hindus. Muslims would invest in shifting Hindu beliefs.*

   I collected data on actual interactions between workers (accusations for contributing low effort, blame etc.) which I use to test this prediction.

[50] π_T is determined by the remaining length of interaction in the game and does not depend on the number of periods that have already elapsed.

[51] A couple of things are worth mentioning here. First, even though exerting e_H might not be worth it at time t, a lucky sequence of high output events can change that, such that in some future period it might be worth it. Alternatively, even though e_H might be worth it in some period, a series of bad outcomes can lead to beliefs of the Hindu worker drifting downwards whereby the Muslim worker may not find it worthwhile to exert high effort anymore in the future.

2. *If Hindus have had past contact with Muslims, we are more likely to see high output in those mixed teams initially.*

   The idea here is that if Hindus have had enough experience of working with Muslims in the past, such that their initial belief π is greater than $\bar{P}$, Hindus and Muslims will coordinate immediately on a high output static equilibrium. I use information on pre-randomization teams of individuals as a proxy for past contact with Muslims in order to test this in the data.

3. *Closer the beliefs of Hindus to $\bar{P}$, the faster (in expectation) is convergence to high output.*

   This is related to the point above. The closer initial beliefs of Hindu workers are to $\bar{P}$, fewer are the number of periods with e_H required from Muslim workers, before the transition to high output static equilibrium is made. This means mixed teams in which Hindu workers have lower priors at baseline, might not see output differences between mixed and non-mixed HD teams complete dissipate during the intervention period.

2.10.4 Proof of proposition

Proposition: *There exists an equilibrium in which, for a given remaining interaction length T, there is a $\pi_T \leq \bar{P}$, such that for $\pi \geq \pi_T$, Muslims will exert e_H. If $\pi < \pi_T$ Muslims exert e_L. Along the equilibrium path, Hindus exert e_L when $\pi < \bar{P}$ and e_H otherwise.*

Proof: Suppose a Hindu and a Muslim worker are working together in a team for periods t = $1, ..., T$, where T is finite but can be arbitrarily large. We assume $\pi < \bar{P}$, whereby the Hindu worker exerts low effort e_L initially. At time period 0, the Muslim worker maximizes expected future payoff. High effort is optimal in the beginning for the Muslim worker iff

$$V_k^{e_H} = \{ \sum_{s' \in S_{k+1}} P_k(s, e_H) + \mu_k(s'|s, a) V_{k+1}^{e_H}(s') \} \geq T P^{e_L} \tag{2.18}$$

for $k = T-1, ..., 0$. Suppose at time 0, (2.18) is true. Then, any other investment path rather than e_H at the beginning (specifically one where the Muslim worker initially exerts e_L and then e_H) is sub-optimal. To see this, suppose that this were not the case by contradiction. Then there exists some t_1 and t_2 such that,

$$t_1 P^{e_L}_{a(H)=e_L} + t_2 P^{e_H}_{a(H)=e_L} + (T - t_1 - t_2) P^{e_H}_{a(H)=e_H} \geq \{\sum_{s' \in S_{k+1}} P_k(s, e_H) + \mu_k(s'|s,a) V^{e_H}_{k+1}(s')\}_{k=T-1,\dots,0} \geq TP^{e_L} \tag{2.19}$$

where t_1 and t_2 respectively denote time periods during which the Muslim worker expects to put low effort and high effort respectively (while the Hindu worker still has not updated their prior above $\bar{P}$). The notation $P^{j}_{a(H)=k}$ denotes that expected payoff to the Muslim worker in a period he exerts effort j and the Hindu worker's action is k. We can similarly split the payoff from exerting high effort and write the inequality as

$$t_1 P^{e_L}_{a(H)=e_L} + t_2 P^{e_H}_{a(H)=e_L} + (T - t_1 - t_2) P^{e_H}_{a(H)=e_H} \geq \tilde{t}_1 P^{e_H}_{a(H)=e_L} + (T - \tilde{t}_1) P^{e_H}_{a(H)=e_H} \geq TP^{e_L} \tag{2.20}$$

where $\tilde{t}_1$ denotes the number of periods until which the the Muslim worker expects to exert e_H while the Hindu worker's $\pi < \bar{P}$. Re-writing the above we have (and ignoring the last inequality),

$$t_1(P^{e_L}_{a(H)=e_L} - P^{e_H}_{a(H)=e_H}) + t_2(P^{e_H}_{a(H)=e_L} - P^{e_H}_{a(H)=e_H}) + TP^{e_H}_{a(H)=e_H} \geq \tilde{t}_1(P^{e_H}_{a(H)=e_L} - P^{e_H}_{a(H)=e_H}) + TP^{e_H}_{a(H)=e_H} \tag{2.21}$$

Notice that t_2 in expectation (at t= 0) must be larger than $\tilde{t}_1$. This is because if the Muslim worker starts off with e_L, in expectation it will take longer to shift the prior of the Hindu worker above $\bar{P}$.

Now I show that if the belief of the Hindu worker is not too low, then the Muslim worker will exert e_H. I split the payoff of the Muslim worker into two parts: before and after (expected) period j, such that for all t $\leq j$, $\pi_t \leq \bar{P}$ and $\pi_t > \bar{P}$ for $t > j$. The Muslim worker will exert e_H iff

$$t_j P^{e_H}_{a(H)=e_L} + (T - t_j) P^{e_H}_{a(H)=e_H} \geq TP^{e_L} \tag{2.22}$$

Suppose this is true. Consider the following extreme scenarios and the consequent prior of Hindu workers: (1) in each period before j, high output is produced and (2) in each period before j, low output is produced. The priors in cases (1) and (2) respectively are:

$$(1): \frac{\pi_0 p^{j-1}_{HL}}{\pi_0 p^{j-1}_{HL} + (1-\pi_0) p^{j-1}_{L}} \tag{2.23}$$

$$(2): \frac{\pi_0 (1 - p_{HL})^{j-1}}{\pi_0 (1-p_{HL})^{j-1} + (1-\pi_0)(1-p_L)^{j-1}} \tag{2.24}$$

(1) and (2) give the lower and upper bound on the beliefs of the Hindu worker about the type of the Muslim worker at time j. The expected prior at time period j can therefore be written as a linear combination of the expressions above. Re-writing, we therefore have,

$$\frac{\pi_0 p_{HL}^{j-1}}{\pi_0 p_{HL}^{j-1} + (1-\pi_0) p_L^{j-1}} + A.\frac{\pi_0 (1-p_{HL})^{j-1}}{\pi_0 (1-p_{HL})^{j-1} + (1-\pi_0)(1-p_L)^{j-1}} \tag{2.25}$$

$$= \frac{\pi_0}{\pi_0 + (1-\pi_0)(\frac{p_L}{p_{HL}})^{j-1}} + A\frac{\pi_0}{\pi + (1-\pi_0)(\frac{1-p_L}{1-p_{HL}})^{j-1}} \tag{2.26}$$

where A is a negative constant. In period j we therefore must have

$$\frac{\pi_0}{\pi_0 + (1-\pi_0)(\frac{p_L}{p_{HL}})^{j-1}} + A\frac{\pi_0}{\pi + (1-\pi_0)(\frac{1-p_L}{1-p_{HL}})^{j-1}} > \bar{P} \tag{2.27}$$

Since the L.H.S. is increasing in j, while the R.H.S. is fixed, we clearly have a value of j such that inequality is satisfied. However, it cannot be so large that equation (2.22) is not satisfied.

In this case π_0 is the starting belief of the Hindu worker. Note that the L.H.S of equation (2.27) is decreasing in π_0, which suggests if π_0 is too small a larger j is required. Given T fixed, the smallest value of π that allows equation (2.22) to be satisfied (i.e. when the equation holds with exact equality) is essentially the threshold value π_T such that if $\pi_0 \geq \pi_T$, then the Muslim worker exerts e_H in period 0. Note that this threshold is updated every period based on the number of interactions that remain.

A Hindu worker simply operates on their prior π_t (state variable) in each period in this Markov equilibrium. If $\pi_t > \bar{P}$, then Hindus exert e_H and e_L otherwise. Both actions are best responses given priors.

Chapter 3

Economic Consequences of Kinship: Evidence from U.S. Bans on Cousin Marriage

3.1 Introduction

"Despite their capacity to form capital, kinship societies remain poor. To explore the economics of kinship societies is thus to explore the economics of underdevelopment." (Bates, Greif, and Singh, 2004)

The weakening of ties among extended family has, since Weber (1951), been associated with development. Recent work by Enke (2019) suggests that loose kinship ties became advantageous in the industrial era, and are linked to urbanization and economic growth. Indeed, Henrich (2020) argues that loosened kinship ties were key to the historical development of Europe.[52] Consistent with this, strong family ties have been linked to lower contemporary growth rates through their effect on generalized trust, geographic mobility, and female labor force participation (Alesina and Giuliano, 2014). The causal relationship underlying this link, however, is still unclear. Family ties may react flexibly to changes in incentives rather than being fundamental causes of economic outcomes (Bau, 2021). A prominent theory in anthropology holds that the increasing economic role of women is what drives the loosening of kinship bonds, while sociologists have argued that urbanization dissolves kinship ties.[53]

[52]See also Schulz, Bahrami-Rad, Beauchamp, and Henrich (2019); Greif and Tabellini (2017); Fukuyama (2011); Korotayev (2000). Notably, the decline of tribes in Europe and the rise of the nuclear family in the late medieval period long preceded the industrial revolution (Greif, 2006).

[53]See Murdock (1949); Naroll (1970) on 'main sequence theory' in anthropology, and Wirth (1938); Tönnies (1957); Fischer (1975) on the role of urbanization on kinship ties. Farber (2000) summarizes the latter view using the German proverb "city air makes one free" and briefly reviews both literatures. Empirical analysis of this link is complicated by the correlation of strong kinship ties with other historical characteristics, such as disease burdens and type of agriculture, which may directly affect

This chapter uses an exogenous decline in marriage between first cousins to estimate the effect of weakening kinship ties on income. We do this using US data from the 19th and 20th centuries, where state-level bans on first-cousin marriage allow us to identify the causal effect of cousin marriage.[54] While now rare in the US, we estimate that 5% of marriages were between first cousins in the first half of the 19th century.[55] We show that the decline of cousin marriage had direct and meaningful economic consequences, including higher income and more schooling. Our results suggest this effect is not driven by the genetic consequences of cousin marriage. Instead, consistent with the weakening of kinship ties, we find a large increase in rural-to-urban migration.[56] Our findings rely on two key contributions: a method for calculating cousin marriage borrowed from human biology, and an identification strategy that exploits the timing of state bans on cousin marriage.

Our measure of cousin marriage comes from the excess frequency of marriages where spouses share a surname. The rate of isonymous ("same-surname") marriages has been widely used to estimate rates of cousin marriage in a population (Crow and Mange, 1965).[57] This paper is the first use of marital isonymy in economics, and we apply it to a far larger set of marriages than, to our knowledge, has been done in other fields. While this method comes at the cost of substantial measurement error, our procedure adjusts for both false positives (unrelated spouses who share a surname) and false negatives (cousins who do not share a surname). We apply this method to a dataset of 18 million US marriage records from 1800 to 1940. These publicly available records, digitized and transcribed at scale for use by amateur genealogists, contain the date and place of marriage and the (pre-marital) names of the spouses. This allows us to track cousin marriage rates over time by surname. We then link these surname-level rates of cousin marriage to individual economic outcomes using 1850 to 1940 US Census returns that include full names.

Our causal analysis exploits the timing of state-level bans on cousin marriage. Thirty-two US states

development. See for example Walker and Bailey (2014); Denic and Nicholls (2007).

[54]Fittingly, the prohibition of marriage between cousins is thought to have been central to the dissolution of clans in Europe and the loosening of kin bonds (Goody, 1983; Schulz et al., 2019).

[55]Among some subsets of the US population, high rates of cousin marriage have persisted until quite recently (Brown, 1951; Reid, 1988; Thomas et al., 1987).

[56]The link between family ties and geographic mobility is highlighted in Alesina et al. (2015); Munshi and Rosenzweig (2016); Greif and Tabellini (2017); Schulz (2019).

[57]We use the terms 'cousin marriage,' 'consanguinity' and 'inbreeding' interchangeably. See Colantonio et al. (2003) for a review of this literature. More recently, DNA analysis has been used to validate results from isonymy, as reviewed in Calafell and Larmuseau (2017).

have banned first-cousin marriage, starting with Kansas in 1858. We restrict our analysis to this set of states, and use differences in the timing of bans to identify treatment effects. These bans are unlikely to have affected all residents equally, and we exploit this to identify the effects of cousin marriage. Specifically, surnames with high rates of cousin marriage are assumed to be more exposed to potential state bans. This is reasonable given surnames exhibit strong persistence in cousin marriage rates over time, and hence surnames with initially low rates of cousin marriage are unlikely to have been directly affected by these bans. Indeed, we find that surnames with high initial rates of cousin marriage see disproportionately large drops in these rates in states with early bans.

Our surname-state-level treatment interacts surname-level variation in cousin marriage rates with state-level variation in the duration of bans on such marriages. That is, we use the interaction of (1) the rate of cousin marriage for a surname in the pre-period (1800-1858), and (2) how long a state has had a ban on cousin marriage. Rather than compare state-wide outcomes, this allows us to compare a targeted population of high-cousin-marriage families across states with early versus late bans. It also allows for state-specific fixed effects to control for any confounding state-wide variation. Causal interpretation of our coefficients rests on a key identifying assumption: the timing of state bans on cousin marriage should not be correlated with factors that affect the relative outcomes of surnames with initially high versus low rates of cousin marriage. We address the two main threats to identification in the following ways.

The first concern is that families with high initial rates of cousin marriage in states with early bans may be different than those in states with late bans. For example, these families may have been either more or less rural, and these differences could have persisted until 1940. We address this concern using the 1850 census to estimate baseline outcomes for surname-state cells to control for pre-existing differences. We also report results from placebo regressions that use 1850 measures as outcomes. These regressions directly test for pre-existing differences between high and low cousin marriage surnames in states with early bans on cousin marriage relative to states with late bans.

The second threat to identification is that states with earlier bans may have enacted other policies that differentially affected families with initially high rates of cousin marriage. For example, the timing of compulsory schooling laws is correlated with that of bans on cousin marriage. We address this by including relevant contemporaneous state-level policies interacted with a surname's rate of cousin marriage in the pre-period. If bans on cousin marriage were proxying for these other policies, these additional interactions should attenuate their effect.

Our first result is that state bans on cousin marriage led to large reductions in measured rates of cousin marriage. Specifically, the duration of state bans on cousin marriage interacted with a surname's rate of cousin marriage in the pre-period, which serves as our treatment variable, strongly predicts rates of cousin marriage in the post-period. We find that state bans reduce cousin marriage rates by about half. We confirm this result in an appendix using an entirely different method and a dataset drawn from genealogical records. This alternative dataset allows us to identify 'true' cousin marriages, rather than infer them from frequencies of same-surname marriages. Both methods find similar magnitudes of cousin marriage reductions following a ban, which suggests that our isonymy measures, while noisy, are correct on average.

Using variation in the extent of a surname's exposure to a state ban, we find that reductions in cousin marriage led to higher incomes. That is, we find that state bans on cousin marriage caused disproportionate increases in incomes for men whose surnames had high initial rates of cousin marriage. By 1940, exposure to a ban on cousin marriage led to 4-6% higher income for surnames with high initial rates of cousin marriage. We also find an increase of 0.2 years of schooling. These magnitudes compare surnames with initial rates of cousin marriage that are one log point apart. This would represent, for example, the comparison of surnames with 16% ('high') cousin marriages to those with 6% ('low').

Crucially, we show that this gap between the 1940 incomes of individuals with differential exposure to cousin marriage bans was absent in 1850, prior to the first ban. This suggests our results are not driven by pre-existing differences. We complement this placebo by showing outcomes for census rounds between 1850 and 1940, and show that the income gap between individuals with different levels of exposure to state bans increases over time in a pattern consistent with our mechanism. That is, relative gains in income for high-cousin-marriage surnames only appear a few generations after bans start being enacted.

A potential explanation for these results comes from the relationship between inbreeding and congenital health problems. Our findings suggest this channel does not drive the increase in income and schooling. We do not find any evidence that bans on cousin marriage affected rates of institutionalization due to physical or mental illness.

We then test whether these results are instead caused by weakened kinship ties. The literature on kinship and strong family ties points to two key outcomes we can measure: geographic mobility and female labor supply. We find a large increase in rural-to-urban migration. Exposure to cousin marriage bans leads to a 5% increase in the likelihood of living in an urban area for surnames with high initial rates of cousin marriage. However, we do not find any effect on inter-state migration. Bans also leads to increased labor supply for women, but no change in labor force participation. Unlike in Alesina and Giuliano (2014), however, we do not find effects on the labor supply of young or of elderly men.

The rest of the chapter proceeds as follows: Section 2 presents our data and the method we use to estimate rates of cousin marriage. Section 3 describes our empirical strategy, including a discussion of the state bans on cousin marriage. Results are in Section 4, and Section 5 concludes.

3.2 Data

This section begins with a description of our dataset of marriage records. We then discuss how we use the rate of same-surname marriages to calculate rates of cousin marriage across surnames and over time. Finally we briefly discuss the Census data we use for measuring outcomes. Appendix B.1.1 describes genealogical data which, while not used in our main analysis, validates our use of isonymy (i.e. same-surname unions) to infer cousin marriage rates, as discussed below.

3.2.1 Marriage records

The marriage records in our dataset come from handwritten documents which have been scanned, transcribed and made publicly available online by Family Search (`familysearch.org`). We retrieved this data for all US states between 1800 and 1940. The transcribed marriage records typically include names of both spouses, and the date and location of marriage. Appendix B.1 includes a scanned image of a sample marriage record, details about what other information these records contain, and our data cleaning procedure.

How good is the coverage of our marriage records? A surprisingly stable benchmark for the US is an annual rate of approximately 10 new marriages per 1000 people (Stevenson and Wolfers, 2007). Our records, averaged annually over the period 1800 to 1940, include about 4 marriages per 1000 people. This suggests that while our dataset is not comprehensive, it accounts a substantial share of marriages in a given year.

Table 3.1 provides summary statistics of our marriage records. The first column includes all marriage records, while the second and third columns include only records either before or after the first US state ban on cousin marriage. While the rate of marriages where the spouses share a surname is low, it is noticeably lower in the latter period of our sample. Further, while we provide more nuance below, a rough benchmark is that the rate of cousin marriage in a population is about four times the rate of isonymy, suggesting cousin marriage rates declined from approximately five percent in the pre-period to about three percent in the post period.

3.2.2 Measuring cousin marriage using marriage records

In the absence of direct measures of cousin marriage during our period of interest, we use a method taken from population genetics to estimate these rates from our dataset of marriages.[58] The basic insight behind the method is straightforward: First cousins, who share two grandparents, will sometimes also share a surname. A family where cousins frequently marry will therefore tend to have a higher share of same-surname (isonymous) marriages than one where they do not. This section describes the formal application of isonymy to our dataset of marriages, including corrections to account for isonymous marriages between unrelated individuals, and marriages between cousins that are not isonymous – that is, both false positives and false negatives.

The use of surnames at marriage to estimate rates of cousin marriage was first proposed by Darwin (1875). Crow and Mange (1965) formalized this approach and showed that the rate of inbreeding in a human population can in some cases be derived from marriage records. That seminal paper spurred a large literature applying their technique to various populations. (Lasker 1985 and Colantonio et al. 2003 review this literature. For examples of marital isonymy applied to US populations see Swedlund and Boyce 1983; Jorde 1989; Relethford 2017.) The link between isonymy and inbreeding has more recently been bolstered by studies which combine surnames with DNA results (Sykes and Irven, 2000; Gymrek

[58]The only dataset we know of with direct measures of consanguinity is Familinx (Kaplanis et al., 2018), which is derived from online genealogies. However, as we describe in Appendix B.1.1, it is anonymized and hence cannot be used for our main analysis, which requires surname-level variation. It is useful, however, in allowing us to perform a number of validation exercises.

et al., 2013; Calafell and Larmuseau, 2017).

Some isonymous marriages are between unrelated people who happen to share a surname. That is, not all Smiths are cousins. To deal with this, we make use of Crow and Mange (1965)'s decomposition of total isonymy into its random and nonrandom components. Total or observed isonymy P is simply the fraction of marriages where spouses share a surname. (Throughout this paper, we refer to the pre-marital, or 'maiden', surnames of marriage partners. Once married, almost all couples in our setting share a surname, which is uninformative.) Random isonymy P^r is defined as the share of marriages we would expect to be isonymous in a population if individuals chose their partners at random. This rate is derived solely from the distribution of surnames in a pool of marriage partners. As per Crow and Mange (1965), random isonymy in a set of marriages is defined as

$$P^r = \sum_s n_s^m \times n_s^f,$$

where n_s^m and n_s^f are the shares of males and females, respectively, with surname s. In our case, we use the surname-level rate of random isonymy, which is essentially a measure of how common that surname is in the population. Total, or observed, isonymy can then be decomposed into its random and nonrandom components as follows:

$$P = P^r + P^n - P^r P^n.$$

Nonrandom isonymy P^n is thus the excess share of isonymous marriages – deviation from the rate we would expect if individuals were marrying at random. We use nonrandom isonymy to calculate cousin marriage rates since, in expectation, it nets out marriages between unrelated partners who happen to share a surname. Nonrandom isonymy, then, adjusts for false positives in calculating cousin marriage rates from isonymy.

Likewise, not all cousin marriages are isonymous. An individual's first cousins can be divided into four types, which are labeled as the offspring of either their (1) father's brother, (2) father's sister, (3) mother's brother, or (4) mother's sister. In a patrilineal society, where children take the surname of their father, only marriages between the first type leads to isonymy.[59] This is illustrated in Appendix figure

[59]Second cousins and more distant relations may also, of course, share a surname. One of the contributions of Crow and Mange (1965) is to show that the degree of inbreeding between two marriage partners is proportional to their probability of

B.3. In the second type, for example, the father passes down his surname but his sister's children take their father's name. If all four types are equally likely, one quarter of cousin marriages will be isonymous. Hence, a first approximation of the rate of cousin marriage in a population is four times the rate of isonymy. Multiplying the isonymy rate by the correct factor, then, adjusts for false negatives in calculating cousin marriage rates from isonymy.

The relationship between isonymy and cousin marriage relies on the assumption, alluded to above, that consanguineous relations occur through male and female ancestors in equal proportion. That is, all four types of cousin marriage are equally likely. Globally, this assumption does not always hold, notably in societies that distinguish linguistically between types of first cousins. Many Arab societies, for example, have a preference for marriage between cousins whose fathers are brothers (Korotayev, 2000). However, no such preference seems to have existed in the US at the time, which is consistent with the lack of linguistic distinction in European languages between types of first cousins (Schneider and Homans, 1955; Swedlund and Boyce, 1983). We test this in Appendix B.1.1 using genealogical data and find that the proportion of each type of cousin marriage is roughly one quarter and shows no secular trend.[60] Further, we use a log transformation of cousin marriage rates in our analysis, which means our results are not sensitive to linear transformations.

These are the steps we take to construct our measures of cousin marriage for a surname s. We begin by selecting a location and time period under consideration, for example Tennessee from 1859 to 1940. Within this setting, let N be the total number of individuals (such that the number of marriages is $N/2$) and N_s be the number of individuals with surname s. N_s^m and N_s^f denote the number of males and females with surname s. We drop a surname from our analysis if N_s^m or N_s^f is less than 50, as we deem these samples too small to provide usable estimates of isonymy. Otherwise, perform the following steps:

Step 1: Observed isonymy P_s

1. Define N_{ss} as the number of individuals with surname s whose marriage partner also has surname

isonymy.

[60] Two other relevant assumptions are that naming practices are consistent (a child always receives their father's surname) and that illegitimacy, adoption and surname changes are negligible (Crow and Mange, 1965). Following the literature on isonymy in the US, we take the first for granted (see for example Swedlund and Boyce 1983). Illegitimacy and adoption are important to geneticists, as it creates a mismatch between inherited genes and inherited surnames. In our case, this distinction is unimportant if children bear the surname of the family that raised them, and hence we do not attempt to correct for them. Surname changes were common for Blacks during our period of interest (most did not have inheritable surnames prior to emancipation) and so, partly for this reason, we exclude Blacks from our analysis.

s.

2. Calculate the observed isonymy rate for this surname: $P_s \equiv N_{ss}/N_s$.

Step 2: Random isonymy P_s^r

[resume]Subdivide observations into state-decade marriage pools, where decades are denoted with subscript d. To simplify notation we use lowercase notation for surname ratios, e.g., $n_{sd} \equiv N_{sd}/N_d$ and $n_{sd}^k \equiv N_{sd}^k/N_d^k$ for $k = m, f$. Calculate random isonymy for each decade pool as follows:

$$P_{sd}^r \equiv \frac{n_{sd}^m n_{sd}^f}{(n_{sd}^m + n_{sd}^f)/2} \equiv \frac{n_{sd}^m n_{sd}^f}{n_{sd}}$$

The denominator can be simplified to n_{sd} because by construction there are equal numbers of males and females (since the unit of observation is a marriage record). The numerator above is the product of the share of males and females with surname s. Notice that if either is zero, random isonymy is zero. The denominator simply normalizes this product by the share of individuals with surname s. If a surname is held by an equal numbers of males and females, the formula above simplifies to $P_{sd}^r = n_{sd}^m = n_{sd}^f$. Aggregate these decade marriage pools for each surname, weighting by the number of individuals in each pool:

$$P_s^r = \sum_d P_{sd}^r \times \frac{N_{sd}}{N_s}.$$

Step 3: Nonrandom isonymy P_s^n

[resume]Calculate nonrandom isonymy using the values of P_s and P_s^r defined in the preceding steps:

$$P_s^n \equiv \frac{P_s - P_s^r}{1 - P_s^r},$$

which is taken from the following decomposition of observed isonymy, $P_s = P_s^r + P_s^n - P_s^r P_s^n$.

Step 4: Cousin marriage rate

[resume]Finally, we calculate the cousin marriage rate by assuming (as described above) that one quarter of cousin marriages are isonymous. The cousin marriage rate is then four times the non-

random isonymy rate, bounded below by zero:

$$CousinMarr_s \equiv \max\{0, P^n_s\} * 4.$$

This procedure is slightly different for pre-period isonymy. Since the number of marriage records in 1800-1858 is substantially smaller, we calculate country-wide surname measures of cousin marriage. That is, the 'setting' is now all of the US, from 1800 to 1858, rather than doing the above steps state by state as we do in the post period. The only modification to the steps listed above is in our calculation of random isonymy (Step 2), where we still use a state o and decade d to define a marriage pool. However, we aggregate across state-decade pools in the following way:

$$P^r_s = \sum_{od} P^r_{sod} \frac{N_{sod}}{N_s}.$$

We now turn to the application of this method to our specific context, the US between 1800 to 1940.

3.2.3 US cousin marriage rates

We use our dataset of marriage records to measure cousin marriage at the level of a surname or surname-state. For concreteness, Table 3.2 presents marriage data on Wallaces, Wheelers and Greens from Tennessee, and the calculation of their isonymy and cousin marriage rates in the post period. Of 648 individuals who appear in the marriage records with the surname Wallace in the post-period, 24 married another Wallace. This 3.7% rate of isonymy is much larger than the expected rate of marriages between Wallaces under random mating, 0.04%. Net of this random component, nonrandom isonymy for Wallaces is 3.67%. Since only one of the four (roughly equiprobable) types of cousin marriage lead to isonymy, we multiply this by four to reach a cousin marriage rate of 15%. Table 3.2 provides the same statistics for Wheelers and Greens of Tennessee, for comparison. Note that in some cases the number of isonymous marriages observed may be less than predicted by random mating, in which case nonrandom isonymy will be negative. In such cases, we treat the cousin marriage rate as being equal to zero.

Random isonymy is the rate of isonymy we would expect if a given group were to marry at random. This requires a decision as to what constitutes a marriage pool or marriage market. That is, if individuals were marrying at random, what is the pool of other individuals they could choose to marry. This should obviously be limited by both time and geography. We subdivide the pre- and the post-periods into decade-long segments, and use the decade-state as a marriage pool for the calculation of random isonymy. Within each of these decade-state marriage pools, we calculate the random isonymy rate for each surname, as described above. We then take the average random isonymy rate for each surname across a period, weighted by the number of observations in each of its two subperiods. This rate of random isonymy is always positive, and is larger for more common surnames.

The use of surnames as a core unit of analysis merits some justification. While it would be ideal to link individuals through generations using direct family links, we believe that surnames provide a useful proxy. First, there are many of them. Our marriages dataset contains 30,000 surnames that appear in more than 200 marriages. Further, while some surnames are common (Smiths account for about 1% of the population), most people hold relatively uncommon names. The top 100 surnames account for only 18% of our marriage records.

Second, cousin marriage rates for given surnames are highly persistent over time. We show this in Figure 3.1, where we calculate cousin marriage rates for each surname (aggregating across states) in the pre and the post periods, and plot them against each other. That is, surnames with high cousin marriage rates in 1800-1858 also have high rates in 1859-1940. This is reassuring as it suggests that surname-level rates of isonymy, which we use to estimate cousin marriage, are measuring stable traits. In turn, the stability of cousin marriage rates is consistent with Giuliano and Nunn (2020), who show persistence of such practices over very long periods of time.

Finally, recent work using DNA sequencing finds that surnames are highly predictive of shared ancestry. Starting with Sykes and Irven (2000), a number of studies have found that Y-chromosome sequences are strongly correlated amongst males who share a surname. Indeed, Gymrek et al. (2013) show that the linking of Y-chromosome sequences to surnames is so reliable that anonymous publicly available DNA sequences can often be linked to specific individuals. Calafell and Larmuseau (2017) provide a review of this literature. One consistent finding is that this genetic link is weaker for the most common names (e.g. Smith), which is taken as evidence of these names having multiple independent origins. This would be the case if unrelated males independently took the name Smith, in this case as a marker of their profession. This is in part why we use non-random isonymy, rather than observed isonymy, as our principal measure of cousin marriage prevalence.[61]

3.2.4 Census data

Data on individual outcomes come from the 1850 to 1940 restricted complete count US Census from IPUMS (Ruggles et al., 2015). We use data on income, education, location of residence, migration and labor supply. Appendix B.1.4 provides a complete list of the outcome variables used in our analysis and their definitions.

We link an individual's census outcomes to a surname-state rate of cousin marriage. For most of our analysis, we use their father's state of birth, rather than their current (e.g. 1940) state of residence, to account for potential inter-state migration.[62] Doing so excludes anyone from our sample whose father was born outside the US. This is useful since the link between immigrants and pre-period rates of cousin-marriage is tenuous. We also restrict our sample to Whites, since many Blacks took on formal surnames only after the Civil War (Byers, 1995; Litwack, 1979) and hence we cannot use surnames to link 1940 Blacks to their historical marriage patterns. We now turn to a description of our empirical analysis, which links our measures of cousin marriage from marital records to these census outcomes.

3.3 Analysis

We combine our measure of cousin marriage with exogenous variation in the propensity to marry first cousins to estimate a causal effect on economic outcomes. This exogenous variation comes from a series of state bans on first-cousin marriage, enacted from 1858 onward. We start this section by introducing these state bans, after which we describe our empirical strategy that exploits this policy variation.

[61] Our results are robust to the exclusion of common surnames, as mentioned in section 3.4.

[62] The 1940 census only asked for father's state of birth for a random 5% of the population, which reduces our usable sample considerably. Prior census rounds ask for father's state of birth for the entire population, though these contain fewer useful outcomes (no wage data, and no grades of schooling).

3.3.1 State bans

Thirty two US states have enacted legislation banning first-cousin marriage, starting with Kansas in 1858. Table 3.3 lists each US state and the year in which it enacted a ban on first-cousin marriage, if ever.[63] There is wide variation in the timing of bans. Kansas was joined by eight states in the 1860s, two in the 1870s, 1880s and 1890s, six in the 1900s, five in the 1910s, and six thereafter. Table 3.3 also reports the number of years cousin marriage had been banned as of 1940, the last year of our analysis.

The US is unique with respect to this type of legislation: It is the only member of the OECD to restrict first-cousin marriage, and the only country globally with sub-national bans (Bittles, 2012). Why did the US ban cousin marriage while Europe and its other offshoots did not? And what explains the state-level variation that we use for causal identification? Paul and Spencer (2016) summarize their conclusion on these questions: "In short, it would seem that laws against cousin marriage are explained by the same factors as legislation permitting compulsory sterilization: relatively poor and powerless targets, an increasingly pessimistic view of heredity, a new willingness to regulate on behalf of the public's health, and a decentralized political system easily swayed by highly motivated activists."

Indeed, in the most sustained treatment of the topic, Ottenheimer (1996) argues that US attitudes turned so decisively against cousin marriage in the 19th century due largely to a growing belief in its negative health consequences. Much of this, he argues, was due to sensationalist news articles or studies such as the Bemiss Report (Bemiss, 1858) which exaggerated the health risks of cousin marriage. The UK and much of Europe, however, saw attitudes towards cousin marriage change around the same period. The more pronounced shift in the US, according to Ottenheimer (1996), was due in part to the influence of theories of civilizational progress which saw family structures of Native American tribes as evidence that cousin marriage was a form of backwardness (Morgan, 1877; Ottenheimer, 1990). Further, the association of cousin marriage with royalty in Europe may have dampened legislative zeal in prohibiting it there (Paul and Spencer, 2016).

While these accounts may explain the general change of attitudes, they do little to explain legislative

[63]While our analysis does not differentiate between types of bans, there are some differences in their details across states. For example, Indiana allows first cousins to marry if they are both above 65. Illinois allows them to marry if they are both above 50 or either is sterile. See Paul and Spencer (2016) for more details on these bans, and for references to the specific legal statutes by which they were enacted. See also Bratt (1984) for a discussion of these bans from a legal perspective. We do not know of systematic data on enforcement of these bans. However, Figure B.7 presents historical news articles that illustrate at least some enforcement.

variation within the US. Some complementary accounts, in contrast, do address state-level variation. The first, by Farber (1968), suggests that the greater individualism and heterogeneous ethnic origins of settlers in the Midwest and West led them to more forcefully oppose first-cousin marriages as a means of assimilation. Ottenheimer (1996), in return, argues that a parsimonious theory fits the data better: Widespread national change in attitudes towards first-cousin marriage only took legal shape when new marriage laws were drafted as states joined the Union. Older states, therefore, were less likely to amend their long standing marriage statutes. Finally, Yamin (2009) argues that activists and lawmakers pushed in some places to extend the reach of the state with an aim to reshape families. This movement, which reached its peak in the Progressive Era, likewise led states to introduce compulsory schooling, child labor laws and compulsory sterilization.

We find some support for these theories in predicting which states ban cousin marriage. For example, states that entered the union later are more likely to have enacted a ban: of states that achieved statehood prior to 1850, 53% eventually enacted a ban, compared to 80% of those who entered afterwards. In contrast, conditional on having ever banned cousin marriage, the state characteristics discussed above do not significantly predict the timing of these bans. We conduct simple empirical tests of these proposed explanations in Appendix B.1.5, by testing whether the following characteristics predict the timing of cousin marriage bans: year of statehood, year of compulsory school, and share foreign born in 1850. We also test for the proportion of population which was urban in 1850, the literacy rate, and measures of religious composition. We also test whether the timing of bans is correlated with state-level prevalence of cousin marriage. We find that it does not. We further show that the timing of bans is uncorrelated with state-level trends in cousin marriage rates.

Consistent with the relatively weak predictive power of these theories, some historians have emphasized the haphazard nature of the legislation against cousin marriage. Discussing these bans, Paul and Spencer (2008) highlight “the ease with which a handful of highly motivated activists–or even one individual–can be effective in the decentralized American system, especially when feelings do not run high on the other side of an issue. The recent Texas experience, where a state representative quietly tacked an amendment barring first-cousin marriage onto a child protection bill, is a case in point.” (Paul and Spencer, 2008, p. 2628)

However we cannot rule out that systematic differences exist between states that predict their decision of whether and when to ban first-cousin marriage. To deal with some of this potential confounding variation, our identification strategy focuses exclusively on differences in the timing of these bans That is, we exclude states that never banned cousin marriage from our analysis.[64]

What if bans were enacted earlier or later in states that were more individualistic and ethnically diverse, more recently settled, or more concerned with legislating the health of families? Firstly, all of our regressions include state fixed effects. This deals with the most obvious concerns related to the selection of states into enacting a cousin marriage ban. However, bias may arise if the state characteristics correlated with such bans also have a differential effect on surnames with high rates of cousin marriage. This could be the case, for example, if people with these surnames lived in more remote areas, and states that passed cousin marriage bans were more likely to impose compulsory schooling laws which may have had a disproportionate effect on this population. After presenting our main results, we show that these state characteristics do not drive our results. We now turn to our empirical strategy, which exploits variation in the timing of state bans.

3.3.2 Empirical specification

The goal of this analysis is to study the causal effect of group-level cousin marriage rates on individual-level economic outcomes. To do so, we estimate the effects of cousin marriage bans on rates of cousin marriage and on economic outcomes. Crucially, we identify surnames with high rates of cousin marriage, and treat these as being more exposed to a potential ban. This is justified in part by the high level of persistence of surname-level cousin marriage rates, as shown in Figure 3.1. Our treatment, then, is at the surname-state level, and consists of the interaction of two continuous variables. The first measures how long a ban has been in place in a given state by 1940, and the second is a surname's rate of cousin marriage prior to the first ban. Our baseline specification is the following:

$$Y_{isor} = \pi[YrsBan_o \times \ln(CousinMarr_s^{pre})] + \delta X_i + \rho_s + \rho_o + \rho_r + \psi_{isor}, \tag{3.1}$$

[64] As we discuss in section 3.4, our main results hold when we also include states that never instituted a ban.

where i denotes an individual, s denotes surname, o denotes father's birth ('origin') state, and r denotes current state of residence. Y_{isor} denotes an outcome of interest. X_i are controls for a quadratic function of age, while α_s, α_o and α_r are surname, state of origin, and state of residence fixed effects, respectively. We cluster standard errors at the surname-state (origin) level. Our results are robust to clustering standard errors at the state level instead.

For each individual, $YrsBan_o \times \ln(CousinMarr_s^{pre})$ is the interaction between (a) the fraction of years between 1858 and 1940 that cousin marriage was banned in their father's birth state o as per table 3.3, and (b) pre-period cousin marriage rates for their surname s. For example, consider a "John Bailey" in 1940 whose father was born in Pennsylvania. Pennsylvania banned cousin marriage in 1902, and the Bailey rate of cousin marriage from 1800-1858 is 8.7%. The treatment value for this person then would be $0.46 \times ln(0.087)$, where 0.46 is the fraction of years between 1858 and 1940 where Pennsylvania banned cousin marriage.

The uninteracted main effects $YrsBan_o$ and $\ln(CousinMarr_s^{pre})$ do not appear in (3.1) because of the inclusion of state of origin and surname fixed effects, ρ_o and ρ_s. That is, we control flexibly for any state-level differences, including any that may be correlated with the timing of bans. Similarly, this controls for any differences across surnames that are correlated with their rates of cousin marriage. Our identifying assumption is that the timing of state bans on cousin marriage is not correlated with factors that affect the relative outcomes of surnames with high versus low rates of 1800-1858 cousin marriage. This would be violated if, for example, states which banned cousin marriage early also enacted other policies which differentially benefited families that had high rates of cousin marriage in the pre-period. We directly test for this by including interactions of state characteristics with pre-period cousin marriage rates.

A number of choices in the specification above merit further discussion. We start with our choice of pre and post periods. The pre period starts in 1800, the first year for which we have marriage records, and ends in 1858, the year Kansas enacted the first state ban on first-cousin marriage. The post period ends in 1940, the last year for which we have census outcomes. We aggregate marriage records over such long periods in part because most surname-state cells have few marriage observations per year. Getting reasonably precise estimates of cousin marriage using marital isonymy requires a sample with many marriages. These long time-scales are further justified because of the potentially long-lasting effects of cousin marriage. A child of parents who are cousins may pass on a strong family-orientation to their own children. Attachment to kin is highly persistent within societies and changes slowly, over generations (Giuliano and Nunn, 2020). It is also likely that the bans will not immediately affect economic outcomes. It will initially affect the choice of marriage partners, which will eventually influence decisions of where to live, how much schooling to acquire, and what occupation to pursue.

Another key specification choice is how we measure cousin marriage rates. Following the cross-country (or cross-region) literature on the effects of cousin marriage (Schulz et al., 2019; Schulz, 2019; Henrich, 2020), we use the logarithm of the rate of cousin marriage to account for the presumed non-linearity in its effects.[65] However, the log transformation emphasizes small differences across surnames with rates of cousin marriage near zero. This is problematic because the number of marriages we use to calculate isonymy rates for most surnames does not allow us to statistically distinguish low rates of cousin marriage from zero.[66] Further, surnames with fewer same-surname marriages than expected given random mating have a measured rate of cousin marriage of exactly zero. To avoid dropping these surnames and to attenuate some of the noise, we censor our cousin marriage rates from below. Specifically we use rates of cousin marriage where we take the log of $\max\{\varepsilon, P_s^n\}$, rather than $\max\{0, P_s^n\}$, where P_s^n is the rate of non-random isonymy (excess rate of same-surname marriages) for surname s. In our baseline specification, we set $\varepsilon = 0.015$ and consider alternative values in our robustness checks. This transformation allows us to focus on variation along the range of values for which we can distinguish cousin marriage from noise. We also show our results are robust to simply using the level of cousin marriage rather than the log, and therefore avoid this censoring procedure altogether. We now turn to the application of the empirical strategy described in this section to our data.

[65] Henrich (2020), for example, writes that "the impact, on both the social world and people's psychology, of increasing the prevalence of cousin marriage from zero to 10% is much bigger than the effect created by increasing it by the exact same amount, from 40% to 50%."

[66] We show in appendix B.1.6 that non-random isonymy values below 0.015 are typically statistically indistinguishable from 0.

3.4 Results

This section begins by showing that our treatment variable, which exploits state-level variation in the timing of cousin marriage bans, is a strong predictor of cousin marriage rates from 1859-1940. That is, the bans did reduce rates of cousin marriage, in proportion to how long they were in place during this period. This is followed by our main result: bans on cousin marriage led to higher incomes and schooling in 1940 for individuals whose surnames were exposed to these bans.

We then explore mechanisms behind this finding, starting with the health effects of cousin marriage. While we have limited ability to measure health outcomes, our results suggest these do not drive our main results. We then study whether this increase in income is caused by a loosening of kinship bonds. While we cannot test for this directly, we test for prominent outcomes from the literature on kinship and strong family ties. In a review, Alesina and Giuliano (2014) highlight the following effects of strong family ties: lower labor force participation of women, youth and the elderly; lower rates of geographic mobility and urbanization; less trust; lower political participation; and higher self-reported happiness and health. Farber (2000), reviewing the sociology and anthropology literature on kinship systems, similarly emphasizes the link between kinship ties and urbanization, geographic mobility and the economic role of women in society. Of these, our data allow us to test the effects of declining cousin marriage on labor supply, as well as geographic mobility (including the decision to move to an urban centre). We find strong evidence in support of the hypothesis that reductions in cousin marriage encourage urbanization, no evidence for a change in inter-state migration, and some evidence of a modest increase in female economic activity.

3.4.1 Effect of bans on cousin marriage rates

Our first result is to show that bans on first-cousin marriage did decrease rates of cousin marriage as measured using our marriages dataset. Specifically, we find that surnames with higher rates of isonymy in the pre-period (1800-1858) see a disproportionate fall in isonymy in the post-period (1859-1940) in states with early bans.

Table 3.4 presents these results. Our treatment variable, the state-level years of cousin marriage ban interacted with surname-level pre-period cousin marriage, has a strong negative impact on cousin marriage in the post-period. Column 1 presents results using the censored log transformation of cousin marriage, while column 2 presents the raw rate of cousin marriage. To interpret magnitudes, compare cousin marriage rates across early and late ban states for a relatively 'exposed' surname–that is, one with a high initial rate of cousin marriage. For a surname with a 10% rate of cousin marriage from 1800-1858, banning in 1858 versus 1940 leads to a reduction of cousin marriage of 2.3 percentage points from 1859-1940. For a surname with the mean rate of cousin marriage in the pre-period (approximately 5%), the decline would be just over one percentage point, or half the mean rate of cousin marriage in the post-period.

Our specification models the effect of a cousin marriage ban as being linear in years. This is reasonable if, each year, marriages are simply affected by whether a ban is in place that year or not. However, if for example enforcement increases gradually, the relationship may not be linear. We present nonparametric results in Appendix Table B.6, where we divide the post-period into five sixteen-year bins. We show that the effect of bans in reducing cousin marriage increases monotonically in their duration.

Further, we show that the finding that cousin marriage bans were effective in reducing rates of cousin marriage does not depend on our method for calculating cousin marriage rates or on our empirical strategy. Appendix B.1.1 replicates this result using an entirely distinct dataset derived from genealogical records to directly identify cousin marriages, rather than inferring them from same-surname marriage records.[67] Any potential problems with this alternative dataset should be orthogonal to the potential issues of measuring cousin marriage using isonymy in marriage records.[68] That dataset, since it does not include surnames, does not lend itself to our surname-specific analysis. Instead, we create a state-year panel of cousin marriage rates, and identify the effect of a state ban on cousin marriage after controlling for state and year effects. This analysis suggests that cousin marriage rates fall by about half in state-years with a ban on cousin marriage, which closely matches the magnitude of results from our main empirical strategy.

[67] Data is from Kaplanis et al. (2018), downloadable at familinx.org

[68] Specifically, one potential issue is that bans disproportionately affected marriages between cousins that share a surname, which would lead us to overstate the reduction in cousin marriage. We show in Appendix B.1.3 that this was not the case.

3.4.2 Income and schooling

We use surname-state variation in exposure to cousin marriage bans to study their impact on income and schooling. Our most striking result is that exposure to bans on cousin marriage leads to higher incomes. Table 3.5 presents results for two measures of 1940 income: individual wage earnings, and imputed income based on occupation.[69]

Cousin marriage bans lead to substantial increases in both measures, as shown in columns 1-4 of Table 3.5. Here, we regress income on our treatment variable, as well as surname, state of origin and state of residence fixed effects, and a quadratic control for age. We restrict our analysis to Whites whose fathers were born in the US, since it is only for these individuals that surnames can reasonably be linked to pre-period marriage records. We restrict to males in most specifications because their surname lineage is identifiable both pre and post-marriage. We also restrict our sample to the prime working ages of 18-50.

Columns 1 and 3 present results from our basic specification. We find that state bans on cousin marriage caused disproportionate increases in incomes for men whose surnames had high initial rates of cousin marriage. By 1940, exposure to a ban on cousin marriage led to 4-6% higher income for surnames with high initial rates of cousin marriage. This magnitude, as with the ones below, compares surnames with initial rates of cousin marriage that are one log point apart. This would represent, for example, the comparison of surnames with 16% ('high') cousin marriages to those with 6% ('low').

In Columns 2 and 4 we also include surname-state level pre-treatment controls to account for potential pre-existing differences in outcomes. Specifically, we use two variables that are available in the 1850 census: (1) imputed income (log occupational income) and (2) a measure of urbanization (the log of the population size of an individual's place of residence). While both predict 1940 income, coefficients on the treatment variable are almost unchanged.

Similarly, schooling is measured using two 1940 outcomes: an individual's highest grade completed, and a binary variable for whether an individual is currently attending school. Since the latter is primarily informative for individuals who are of secondary school age, we restrict that sample to 12-18 year-olds.

[69] The 1940 census offers a better measure of income than preceding censuses, since for the first time individuals were asked not only for their occupation but also for their wage income. Non-wage income (for example from a family business or farm output), however, was captured only by asking whether the respondent received more than $50 of non-wage income in the past year. Occupational income is a measure generated by the Census Bureau to allow for measures of income where these are not available. These variables are described in more detail in Appendix B.1.4.

Results from columns 5-8 show that bans on cousin marriage increased schooling. As above, magnitudes can be interpreted as follows: banning cousin marriage led to 0.25 additional years of schooling for surnames with high initial rates of cousin marriage. The probability of being in school (for 12-18 year-olds) in 1940 increases by 1.7 percentage points (column 7), though this effect is only marginally significant.

An important concern with the results above is that the relative outcomes of males with high-cousin-marriage surnames may have been different in states with early bans even in the absence of these bans. For example, cousin marriage may have been a less rural phenomenon in states with early bans, which could bias our findings. To address this concern, we test for pre-existing differences between high and low cousin marriage surnames that are correlated with the timing of state bans.

We do this in Table 3.6 using measures of income, schooling and urbanization that are available in the US census from 1850 to 1940.[70] Since the first ban was passed in 1858, we should expect the effect of the bans to only show up gradually over the census years. Reassuringly, we find the effects to be small (and statistically insignificant) in the 1850 census when no bans had yet been enacted. They gradually get larger in magnitude over the next census rounds, with the effects being statistically significant starting in 1920, at which time three quarters of the potential treatment period (1859-1940) had gone by. These results suggest that our results using 1940 census outcomes are driven by the causal impact of cousin marriage bans, which gradually change individual-level outcomes as the treatment period elapses.

3.4.3 Congenital health effects

One possible explanation for the positive effects of a reduction in cousin marriage on income and schooling is that they are due not to the cultural shift engendered by a change in family structure, but rather due to the direct genetic effects of a reduction in consanguinity.

A substantial literature has studied the effects of inbreeding on child development due to the expression of recessive genes. However, evidence on the magnitude of these effects is mixed. Saggar and Bittles (2008) review this literature and conclude that inbreeding is associated with modestly higher risk of neonatal and post-neonatal mortality. A recent paper by Mete et al. (2020) finds that in Pakistan, children born into consanguineous unions tend to be stunted and have lower cognitive scores. This litera-

[70] We show results for every other decennial census round. The 1890 census records were destroyed in a fire, hence the somewhat irregular pattern of census rounds.

ture, however, almost universally relies on correlations of inbreeding with health outcomes, with limited ability to control for factors which might be correlated with consanguinity. Mobarak et al. (2019) provide a more causal analysis using an instrument derived from the number of opposite-sex cousins available at age of marriage. They find far more modest results and argue that previous observational estimates are upwardly biased.

To the extent that inbreeding does lead to worse health outcomes and cognitive deficits, some of the effects of cousin marriage on economic outcomes could come through this channel. We test this using census data on whether an individual lives in a long-term care facility such as a hospital, mental institution or home for the physically handicapped. The results are presented in table 3.7. For consistency, we provide results for 18 to 50-year-old males in columns 1-2, though we also report results for all Whites in column 3. In columns 4-6 we present similar results using the 1930 census instead of the 1940 census. The rate of institutionalization in the 1930 census is over three times larger, which may provide more power in detecting an effect.[71]

We do not find any evidence that cousin marriage bans reduced the probability of either being hospitalized or of being admitted to a mental institution. While these outcomes are coarse measures of health or cognitive ability, they provide some evidence that our results on the effects of cousin marriage on economic outcomes are not driven by biological channels. Moreover, our results here are consistent with more recent studies that have tried to establish a causal relationship between consanguinity and health and find small or no effects.

We now turn to outcomes which are linked to another channel by which reductions in cousin marriage may have led to higher income: the weakening of kinship ties. Specifically, as described above, we study the effects on labor supply, urbanization and geographic mobility.

3.4.4 Labor supply

An important channel through which the strength of kinship bonds has been linked to economic outcomes is through changes in female labor supply. We study this both along the intensive and extensive

[71] It is not clear why the rate is lower in 1940 than 1930. Deinstitutionalization, the widespread closure of mental asylums, had not yet begun in 1940. The change may derive from differences in how particular institutions were classified across census rounds.

margin: the number of weeks worked in the past year, and an indicator for having worked one or more weeks. We present these results for men as well as women. There is no reason to believe that cousin marriage would affect the labor supply decision of men in general, though Alesina and Giuliano (2014) find that stronger family ties are also correlated with lower labor force participation of the young and elderly. We use labor supply decisions of men aged between 14-18 and those above 51 along with our main sample of those between 18-50 to formally test for this.

Interpreting results for women in our sample requires some clarification. Most of our regressions use outcomes only for men since their surnames can be linked back to their father, paternal grandfather, and so forth. The same is true for unmarried women. Married women, however, do not report their pre-marital surnames in the census, making it impossible to trace their ancestors. Married women are therefore linked to their husband's surname and husband's father's state of birth. For married women, then, the treatment should be interpreted as coming through their husband. A husband's family's rate of cousin marriage could be linked to his wife's labor supply decision either through spousal selection (e.g. cultural homophily) or directly through his influence on her labor supply decisions.

Columns 2-4 of Table 3.8 suggest that male labor supply is unrelated to rates of cousin marriage. This is true for both the extensive and intensive margins as well as for the young and the elderly. The evidence for women is mixed. Column 1 suggests that, on the extensive margin, cousin marriage does not lead to an increased propensity to supply labor outside the household. Bans on cousin marriage did however increase female labor supply on the intensive margin. Relative to a surname with one log point lower cousin marriage in the pre-period, an early ban leads to almost 1 additional week of work, from a mean of 15. These results come with the caveat that our ability to measure effects for women is less direct than for men, as discussed above.

3.4.5 Geographic mobility and urbanization

Another outcome linked in the literature to the strength of kinship bonds is the choice of where to live. As Alesina et al. (2015) emphasize, strong family ties make moving away from home more costly. We study two sets of outcomes related to this decision which the census captures: the decision of whether to live in an urban area, and the decision to migrate across state lines.

Urbanization is measured firstly using a dummy variable for whether an individual in 1940 is coded as living in an urban, rather than a rural, location.[72] Second, we use another dummy variable to denote if an individual reported to be living on a farm. Third, we use a continuous measure of urbanization: the log of the population size of one's location of residence. Columns 1-6 of Table 3.9 show that our treatment variable strongly predicts increases in urbanization across all the three measures described above. That is, early bans on cousin marriage lead to disproportionately high likelihoods of urbanization for surnames with high initial rates of cousin marriage. Exposure to a ban leads to a 4.8 percentage point increase in urbanization, a 4 percentage points decrease in the probability of living on a farm, and a 30 percent increase in the population size of one's locality of residence. These effects are substantial in magnitude and may explain part of the increase in income from a reduction in cousin marriage.

In columns 7-10, we test for increased inter-state migration. The coefficients suggest precise null effects on both 5-year migration as well as lifetime migration, defined as living outside one's state of birth. Overall, the results in this section suggest that bans on cousin marriage led to greater rural-to-urban migration within states, but no increase in inter-state migration.

Why did bans on cousin marriage lead people to be more likely to live in urban areas? One potential explanation is almost mechanical: If a large share of local potential marriage partners are your first cousins, you may be forced to move to a larger population area to get married. We believe this channel is unlikely to be quantitatively important. Assuming a fertility rate of five (more than double the average rate for Whites in 1940), a person seeking a marriage partner would have 40 first cousins. Over 99% of individuals in 1940 lived in counties with more than 5,000 inhabitants. While the effect of the ban on one's marriage pool depends on their age, sex and marital status, it seems unlikely that these bans had a substantial effect on the raw number of potential local marriage partners. Instead, bans on cousin marriage, by weakening kinship bonds, may have reduced attachment to a family home and hence reduced the perceived cost of moving to a nearby town or city to find work, as emphasized in Alesina et al. (2015); Munshi and Rosenzweig (2016); Greif and Tabellini (2017); Schulz (2019).

[72] The Census Bureau considered cities and incorporated places of 2,500 inhabitants or more as urban. It also included other local subdivisions with population of 10,000 and population density above 1000 per square mile. See Appendix B.1.4 for more details on this classification.

3.4.6 Robustness

Our findings are robust to a wide range of placebos and robustness checks. First, while we believe that the log of the cousin marriage rate is the appropriate measure for this analysis, we show that our results are not driven by this transformation. We present our main results using the raw (untransformed) measure of cousin marriage in Table B.7. We confirm that being treated with an early ban on cousin marriage leads to higher income, more schooling, more urbanization and increased female labor supply. In Appendix B.1.6 we also discuss our choice of ε threshold when censoring values of isonymy. We justify our choice of threshold and provide results for alternative values in table B.8.

We show that our results are also robust to dropping the most common and the least common surnames from our sample (table B.9) as well as to the inclusion of all states in our sample, rather than just those that eventually ban cousin marriage (table B.10). In this specification we treat states that never banned cousin marriage akin to those that banned after 1940, i.e., zero years of ban between 1859-1940. Furthermore, we show that our main results are robust to using 1930 census outcomes in table B.11.

Next, we take up the discussion from section 3 on the determinants of state bans on cousin marriage. We test whether any of the following act as confounders in our regression results: the timing of statehood, the degree of ethnic heterogeneity, and the willingness of state legislators to intercede in family decisions. We measure the first using the number of years elapsed since achieving statehood by 1940, to mirror our treatment of the state bans. Ethnic heterogeneity is proxied using the state share of 1850 census respondents born outside the US. State legislation in family decisions is proxied using the number of years elapsed since compulsory high school legislation was enacted by 1940. In each of these specifications we interact these state characteristics with our measures of pre-period surname-level cousin marriage. State fixed effects in all our regressions control for the overall effects of any state characteristics. By adding these interactions to our regression specifications we test whether the differential outcomes of high-cousin marriage surnames in states that banned cousin marriage are partly driven by other state characteristics. We find no evidence for this – in each of the three cases our results do not change substantially in magnitude or significance, as shown in tables B.12 to B.14.

3.5 Conclusion

This paper uses 19th and 20th century U.S. state level bans on cousin marriage to provide causal micro-evidence of the impact of consanguineous marriages on a range of economic outcomes. Borrowing a method from population genetics, we show that excess rates of same-surname marriages can provide credible estimates of cousin marriage rates by surname, by state, and over time. We show that bans on first-cousin marriage led to a reduction in the rate of those marriages, and also led to higher incomes, more schooling, rural-to-urban migration and increased female labor supply. These effects do not seem to be driven by the genetic impacts of cousin marriage. Instead, we argue that the economic gains we document are driven largely by changes in social relationships that stem from weakened kinship ties.

These effects, while striking in magnitude, are consistent with work in anthropology and sociology that studies the characteristics of strong kinship ties. Henrich (2020), for example, summarizes a large body of ethnographic and historical research showing that tight (intensive) kinship is associated with greater cooperation within a kin group, at the cost of geographic and social mobility and participation in anonymous markets and broader impersonal institutions. The results from this paper are consistent with the view that kinship norms evolve as economic conditions change, but that they do so slowly, over the course of many generations.

While clearly of historical significance, we believe these results may also be relevant for contemporary development outcomes, since intensive kinship is still prevalent in many societies. Figure 3.2 shows estimated national contemporary rates of cousin marriage plotted against incomes per capita. These data suggest high rates of cousin marriage in many countries, and a striking cross-country correlation with development and political institutions (Schulz, 2019; Akbari et al., 2019; Woodley and Bell, 2013). The causal estimates in this paper of the impact of kinship are not directly applicable to such societies, where kinship ties may substitute for weak formal institutions. Nevertheless, our results do suggest that as economies undergo structural transformation, leading to the development of better institutions, there could be high returns from family structure transitions that weaken kinship ties.

In future work, we plan to study the effect of cousin marriage on household formation and family structure. A substantial literature has studied the gradual emergence of the "European Marriage Pattern" (Hajnal, 1965) and whether this set of practices (notably late marriage and nuclear neolocal households) was central to the economic success of Europe (e.g. Dennison and Ogilvie, 2014). We plan to test whether the decline of cousin marriage made this family structure more common, and its link to changes in migration and human capital acquisition. Using street addresses on census returns, we also intend to study the relationship between kinship tightness and geographical clustering. Do brothers migrating to cities, for example, tend to live near each other? Finally, does a decline in cousin marriage lead to fewer sons taking on their father's first name? If we do observe fewer "Jr's," that suggests these naming practices can serve as easily-observed proxies of kinship in America (Taylor, 1974).

3.6 Tables and figures

3.6.1 Tables

Table 3.1: Summary statistics: Marriage records

	All marriage records (1800-1940)	Prior to first ban (1800-1858)	Post first ban (1859-1940)
Number of marriages	17.73 million	3.45 million	14.28 million
Isonymous (same-surname)	0.0085	0.0120	0.0075

Table 3.2: Calculating cousin marriage rates from isonymy, examples from Tennessee

Surname	Wallace	Wheeler	Green
1859-1940 marriage records (post-period)			
Individuals with surname	648	4781	1696
Married to same surname spouse	24	64	2
Observed isonymy P_s	0.0370	0.0134	0.0012
Random isonymy P_s^r	0.0004	0.0035	0.0011
Nonrandom isonymy P_s^n	0.0367	0.0099	0.0001
Cousin marriage rate	14.66%	3.97%	0.05%

Cousin marriage rates are calculated using the following formula: $CousinMarr_s = \max\{P_s^n, 0\} \times 4$.

Table 3.3: Year of enactment of state laws banning first-cousin marriage

State	Year	Years as of 1940	State	Year	Years as of 1940
Alabama	Never ban	-	Nebraska	1911	29
Arizona	1901	39	Nevada	1861	79
Arkansas	1875	65	New Hampshire	1869	71
California	Never ban	-	New Jersey	Never ban	-
Colorado	1864	76	New Mexico	Never ban	-
Connecticut	Never ban	-	New York	Never ban	-
Delaware	1921	19	North Carolina	Never ban	-
Florida	Never ban	-	North Dakota	1862	78
Georgia	Never ban	-	Ohio	1869	71
Idaho	1921	19	Oklahoma	1890	50
Illinois	1887	53	Oregon	1893	47
Indiana	1877	63	Pennsylvania	1902	38
Iowa	1909	31	Rhode Island	Never ban	-
Kansas	1858	82	South Carolina	Never ban	-
Kentucky	1946	0	South Dakota	1862	78
Louisiana	1900	40	Tennessee	Never ban	-
Maine	1985	0	Texas	2005	0
Maryland	Never ban	-	Utah	1907	33
Massachusetts	Never ban	-	Vermont	Never ban	-
Michigan	1903	37	Virginia	Never ban	-
Minnesota	1911	29	Washington	1866	74
Mississippi	1923	17	West Virginia	1917	23
Missouri	1889	51	Wisconsin	1914	26
Montana	1919	21	Wyoming	1869	71

Source: Paul and Spencer (2016). 'Years as of 1940' refers to the number of years a state has had a ban in place as of 1940. Alaska and Hawaii are omitted since they achieved statehood post-1940. Neither has ever banned first-cousin marriage, nor has Washington, D.C.

Table 3.4: Impact of bans on cousin marriage rates

	(1)	(2)
	Log (CousinMarr^{post})	**CousinMarr^{post}**
Years Ban (Fraction) × Log (CousinMarr^{pre})	-0.307***	
	(0.0709)	
Years Ban (Fraction) × CousinMarr^{pre}		-0.258***
		(0.0732)
Mean Dep. Var.	-2.75	0.0239
N	1,207,523	1,207,523
Adj. R^2	0.450	0.458

* p<0.10, ** p<0.05, *** p<0.010. Standard errors clustered at the surname-origin state (father's birth state) level in parentheses. All regressions include surname, state-of-origin and state-of-residence fixed effects. Cousin marriage rates are calculated (as described in section 3.2.2) at the surname-level in the pre-period and surname-state level in the post period and linked at the surname-origin state (of father's birth) level with the 1940 Census records. Years Ban (Fraction) refers to the number of years each state had a ban on cousin marriage in place by the year 1940, divided by 82, which is the number of years that had passed since the first ban in Kansas in 1858. Sample includes White males aged 18 to 50 in 1940 unless otherwise specified.

Table 3.5: Impact of cousin marriage bans on income and schooling

	Log Wage Income (1940)		Log Occupational Income (1940)		Highest Grade Completed (1940)		In School (1940)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years Ban (Fraction) × Log (CousinMarr^{pre})	0.0605**	0.0542**	0.0392***	0.0337**	0.246***	0.200**	0.0170*	0.0176
	(0.0245)	(0.0273)	(0.0140)	(0.0155)	(0.0794)	(0.0922)	(0.00915)	(0.0111)
Surname-state pre-treatment controls								
Log Occupational Income (1850)		-0.00483		0.0330***		0.188***		0.0134***
		(0.0170)		(0.0122)		(0.0540)		(0.00511)
Log Urbanization (1850)		0.0423***		0.0344***		0.108***		0.00999***
		(0.00461)		(0.00340)		(0.0147)		(0.00147)
Surname F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State of Origin F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State of Residence F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	White Males (18-50)		White Males (18-50)		White Males (18-50)		White Males (12-18)	
Mean Dep. Var.	6.13	6.14	2.87	2.88	10.17	10.10	0.790	0.783
N	682,847	617,084	976,414	885,324	1,187,725	1,071,133	1,129,743	991,424
Adj. R^2	0.247	0.244	0.122	0.122	0.120	0.119	0.225	0.229

* p<0.10, ** p<0.05, *** p<0.010. Standard errors clustered at the surname-origin state (father's birth state) level in parentheses. Cousin marriage rates are calculated (as described in section 3.2.2) at the surname-level in the pre-period and surname-state level in the post period and linked at the surname-origin state (of father's birth) level with the 1940 Census records. Years Ban (Fraction) refers to the number of years each state had a ban on cousin marriage in place by the year 1940, divided by 82, which is the number of years that had passed since the first ban in Kansas in 1858. Sample includes White males aged 18 to 50 in 1940 unless otherwise specified.

Table 3.6: Income, schooling and urbanization: Placebo regressions from 1850-1940 Censuses

Census	*1850*	*1870*	*1900*	*1920*	*1940*
Fraction of treatment period (1859-1940)	0% (Placebo)	15%	51%	76%	100%
			Log Occupational Income		
Panel A: Income					
Years Ban (Fraction) × Log (CousinMarr^{pre})	-0.0023	0.004	0.005	0.0185**	0.0392***
	(0.0121)	(0.0088)	(0.0076)	(0.0082)	(0.0139)
Mean Dep. Var.	2.90	2.82	2.86	2.97	2.87
N	643,809	1,129,136	1,983,051	2,649,824	976,414
			In School (12-18 year olds)		
Panel B: Schooling					
Years Ban (Fraction) × Log (CousinMarr^{pre})	-0.00558	-0.0102	0.0284	0.00682	0.0165*
	(0.0152)	(0.0130)	(0.0446)	(0.00764)	(0.00914)
Mean Dep. Var.	0.625	0.630	0.601	0.689	0.789
N	337,692	602,610	41,869	1,119,576	1,130,714
			Log population size of locality of residence		
Panel C: Urbanization					
Years Ban (Fraction) × Log (CousinMarr^{pre})	0.0312	0.0489	0.0327	0.0814**	0.303***
	(0.03402)	(0.0320)	(0.0379)	(0.0409)	(0.0733)
Mean Dep. Var.	6.59	6.91	7.57	8.28	8.27
N	659,536	1,304,994	2,370,685	3,455,349	1,207,523
			Living on Farm		
Panel D: Farm					
Years Ban (Fraction) × Log (CousinMarr^{pre})	0.00330	-0.00776	-0.0157**	-0.0230***	-0.0388***
	(0.0127)	(0.0100)	(0.00782)	(0.00697)	(0.0113)
Mean Dep. Var.	0.580	0.432	0.430	0.370	0.295
N	729,480	1,304,979	2,370,685	3,455,349	1,209,997
Surname F.E.	Yes	Yes	Yes	Yes	Yes
State of Origin F.E.	Yes	Yes	Yes	Yes	Yes
State of Residence F.E.	Yes	Yes	Yes	Yes	Yes

* p<0.10, ** p<0.05, *** p<0.010. Standard errors clustered at the surname-origin state level in parentheses. Cousin marriage rates are calculated (as described in section 3.2.2) at the state level in the pre-period and surname-state level in the post period. These are linked at the surname-state (of own birth) level for census records before 1900 and at the surname-state (of father's birth) level for records from 1900 onwards. Years Ban (Fraction) refers to the number of years each state had a ban on cousin marriage in place by the year 1940, divided by 82, which is the number of years that had passed since the first ban, in 1858. Sample includes White males aged 18 to 50 unless otherwise specified.

Table 3.7: Impact cousin marriage bans on genetic outcomes

	Living in Hospital, Mental Institution or Home for Physically Handicapped					
	(1)	(2)	(3)	(4)	(5)	(6)
Years Ban (Fraction)	0.0004	0.0009	0.0001	0.0006	0.0010	0.0004
× Log (CousinMarr^{pre})	(0.000601)	(0.000674)	(0.000296)	(0.000598)	(0.000636)	(0.000347)
Surname F.E.	Yes	Yes	Yes	Yes	Yes	Yes
State of Origin F.E.	Yes	Yes	Yes	Yes	Yes	Yes
State of Residence F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Pre-treatment controls (surname-state)	No	Yes	Yes	No	Yes	Yes
Sample	White Males (18-50)		Whites (All)	White Males (18-50)		Whites (All)
Year	1940	1940	1940	1930	1930	1930
Mean Dep. Var. (% of population)	0.1%	0.1%	0.08%	0.37%	0.36%	0.26%
N	1,207,523	1,087,102	3,686,601	4,021,199	3,790,663	7,945,497
Adj. R^2	0.001	0.001	0.002	0.002	0.002	0.002

* p<0.10, ** p<0.05, *** p<0.010. Standard errors clustered at the surname-origin state (father's birth state) level in parentheses. Cousin marriage rates are calculated (as described in section 3.2.2) at the surname-level in the pre-period and surname-state level in the post period and linked at the surname-origin state (of father's birth) level with the 1940 Census records. Years Ban (Fraction) refers to the number of years each state had a ban on cousin marriage in place by the year 1940, divided by 82, which is the number of years that had passed since the first ban in Kansas in 1858. Surname-state pre-treatment controls include occupational income (logged) and population of locality of residence (logged) measured at the surname-state level from the 1850 census records. Sample includes White males aged 18 to 50 in 1940 unless otherwise specified.

Table 3.8: Impact of cousin marriage bans on labor supply

	(1)	(2)	(3)	(4)
Sample	Female (18-50)	Male (18-50)	Male (14-18)	Male (51+)
Panel A: Labor Force Participation				
		Dep Variable: Weeks worked > 0		
Years Ban (Fraction) × Log (CousinMarr^{pre})	0.0144	-0.00246	-0.00724	0.00879
	(0.0113)	(0.0102)	(0.0112)	(0.0258)
Mean Dep. Var.	0.394	0.724	0.197	0.700
N	853,731	1,087,102	701,232	90,637
Adj. R^2	0.039	0.106	0.127	0.206
Panel B: Labor Supply				
	Dep Variable: Number of weeks worked in the past year			
Years Ban (Fraction) × Log (CousinMarr^{pre})	0.948**	0.280	-0.207	-0.434
	(0.490)	(0.492)	(0.390)	(1.412)
Mean Dep. Var.	14.95	28.56	5.83	31.12
N	853,731	1,087,102	701,232	90,637
Adj. R^2	0.055	0.154	0.108	0.166
Surname F.E.	Yes	Yes	Yes	Yes
State of origin F.E.	Yes	Yes	Yes	Yes
State of Residence	Yes	Yes	Yes	Yes
Surname-state pre-treatment controls	Yes	Yes	Yes	Yes

* p<0.10, ** p<0.05, *** p<0.010. Standard errors clustered at the surname-origin state (father's birth state) level in parentheses. Cousin marriage rates are calculated (as described in section 3.2.2) at the surname-level in the pre-period and surname-state level in the post period and linked at the surname-origin state (of father's birth) level with the 1940 Census records. Unmarried women are linked to their father's state of birth while married women to their husband's father's state of birth. Years Ban (Fraction) refers to the number of years each state had a ban on cousin marriage in place by the year 1940, divided by 82, which is the number of years that had passed since the first ban in Kansas in 1858. Surname-state pre-treatment controls include occupational income (logged) and population of locality of residence (logged) measured at the surname-state level from the 1850 census records.

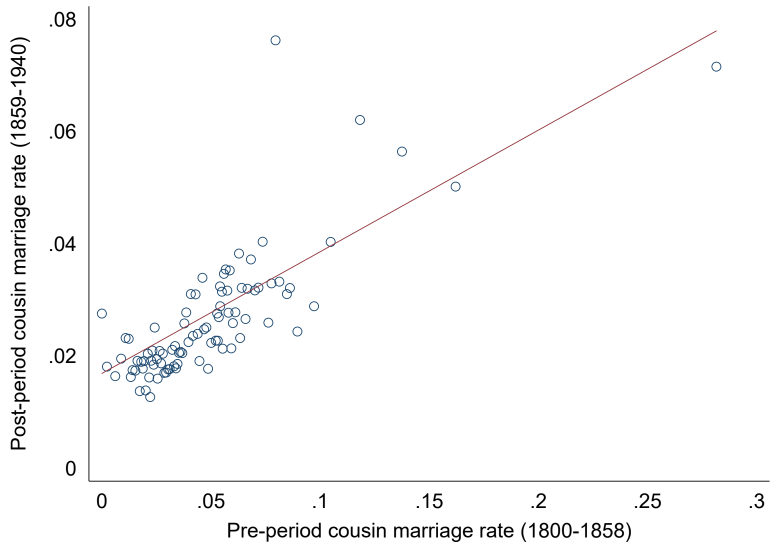
Table 3.9: Impact of cousin marriage bans on urbanization

	Living in Urban		**Living in Farm**		**Log Urbanization**		**Inter-state migration** 5 years		**Inter-state migration** Lifetime	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Years Ban (Fraction) × Log (CousinMarr^{pre})	0.0475***	0.0329**	-0.0397***	-0.0396***	0.303***	0.239***	-0.00203	-0.00540	0.00505	0.0127
	(0.0127)	(0.0139)	(0.0113)	(0.0126)	(0.0733)	(0.0782)	(0.00419)	(0.00435)	(0.00717)	(0.00774)
Surname F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State of Origin F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State of Residence F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Surname-state pre-treatment controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean Dep. Var.	0.446	0.448	0.325	0.324	8.28	8.29	0.047	0.045	0.205	0.196
N	1,207,523	1,087,102	1,207,523	1,087,102	1,207,523	1,087,102	1,156,740	1,041,639	1,203,588	1,083,999
Adj. R^2	0.097	0.101	0.117	0.121	0.132	0.134	0.058	0.059	0.246	0.257

* p<0.10, ** p<0.05, *** p<0.010. Standard errors clustered at the surname-origin state (father's birth state) level in parentheses. Cousin marriage rates are calculated (as described in section 3.2.2) at the surname-level in the pre-period and surname-state level in the post period and linked at the surname-origin state (of father's birth) level with the 1940 Census records. Years Ban (Fraction) refers to the number of years each state had a ban on cousin marriage in place by the year 1940, divided by 82, which is the number of years that had passed since the first ban in Kansas in 1858. Surname-state pre-treatment controls include occupational income (logged) and population of locality of residence (logged) measured at the surname-state level from the 1850 census records. Sample includes White males aged 18 to 50 in 1940 unless otherwise specified.

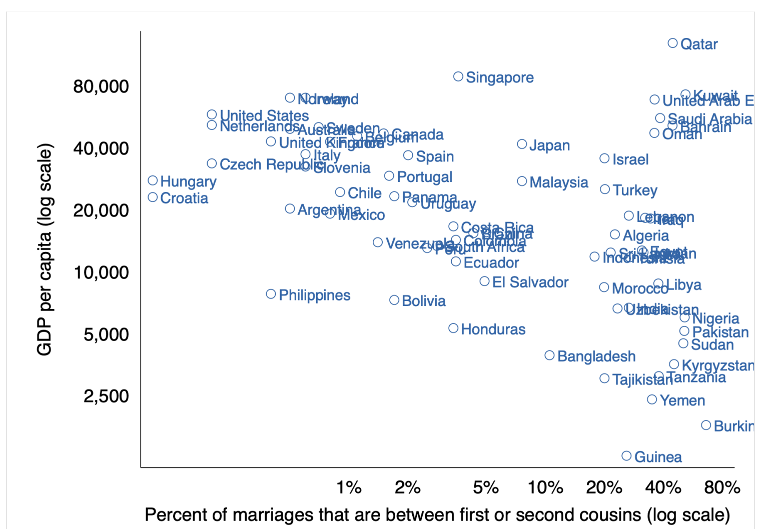
3.6.2 Figures

Figure 3.1: Persistence in cousin marriage rates by surname



Note: This figure is a binscatter of surname-level rates of isonymy in the pre- and post-period.

Figure 3.2: Consanguinity and income (Cross-country correlation)



Sources: Data on consanguineous marriages is from `consang.net` (Bittles 2015). GDP per capita values are from the Penn World Table 2000. *Note*: Consanguineous marriages here are defined in this series as marriages between first or second cousins. GDP per capita values are adjusted for purchasing power parity. Both series are displayed on a log scale.

Chapter 4

Elections, Accidental Deaths and Insurgency: Recipe for India's Conflict Minerals

4.1 Introduction

Natural resource extraction is often associated with conflict, severe corruption and high rent seeking in many countries across the world. Unlocking the true potential of these sectors is crucial to propel developing nations into higher growth paths. However, illegal and unsustainable practices due to the absence and/or weak enforcement of regulations, coupled with disregard for large sections of the poor directly affected by such actions, can often have negative consequences. This could, for example be in the form of environmental degradation and physical displacement of sections of the vulnerable population. A growing strand of literature is now studying these topics (see Burgess et al., 2011; Asher and Novosad, 2018).

A separate literature is motivated by the fact that incumbent politicians may manipulate policies over the electoral cycle for political gains. While the focus has traditionally been on macroeconomic policies in mature democracies, it is now shifting to nascent democracies, particularly focusing on electorally sensitive factors typical to countries in early stages of development. For example, Cole (2009) shows the existence of political cycles in agricultural credit offered by public banks in India – he finds that the amount of credit offered gradually increases in years leading up to state elections. Bhattacharjee (2014) and Baskaran et al. (2015) also focus on India, and find similar cycles in public health provision and rural electrification respectively.

I construct a new state/district-level dataset from India and bring these two strands of literature together in this chapter. I first document the existence of opportunistic political cycles in mineral licensing, mining output and accidents at the state-level. In contrast to much of the existing literature, the cycles exhibit an inverted U-shaped pattern between state assembly elections, with the level of activity minimized in election years. Using district level data, I then examine how these patterns are affected by electoral competition and local conflict, namely India's Naxalite insurgency, which is believed to be fuelled significantly by extortion of mining revenues in mineral rich areas. I find that electoral competition intensifies mining cycles, with politically competitive districts exhibiting larger reductions in fatalities, but only in election years. A similar pattern is observed in the conflict prone districts. I show that this is driven by politicians' objective of minimizing the rebels' resource base during polls, since the latter systematically threaten national and state assembly elections with violence. Finally, while average levels of violence across mining and non-mining districts within states are not significantly different, I find that a reduction in mining intensity leads to reduced electoral violence in mining districts relative to non-mining ones. This suggests strategic behavior on the part of politicians. Politicians do not directly influence small businesses owners, contractors or even poor villagers who the rebels “tax" (Chakravarti, 2014), but with their large scale involvement in the mineral industry in India (Asher and Novosad, 2018), they are able to manipulate activity in a way that suits their electoral agenda.

This chapter makes three important contributions. It points to the existence of political cycles in an industry with both private and public organizations, suggesting collusive actions between politicians and private firms in generating them. Majority of the previous work on political cycles considers policies directly under the purview of either federal or local governments. Secondly, while much of the literature on Naxalite violence only considers static predictors of conflict such as Scheduled Caste/Scheduled Tribe (SC/ST) population proportions,[73] land inequality, presence of minerals, etc., this paper sheds light on the dynamic nature of violence, in a context where realistically, the rebel groups can never achieve their declared objective of overthrowing the Government of India through armed struggle. Such analyses elucidate behavioral aspects of rebellion intensity which are more relevant for conflict resolution strategies. Finally, this paper also contributes to the literature on myopic voting in a developing country context.

[73] Scheduled Castes (SC) and Scheduled Tribes (ST) are officially designated groups of historical disadvantage as a result of the caste system in India.

The fact that politicians respond to electoral competition by attempting to reduce mining fatalities only in election years, shows that the poor, who in general are likely to carry greater grievances, nevertheless evaluate the performance of politicians only over a short horizon.

The findings of this chapter have important implications for policy aimed at diffusing opportunistic cycles in mining fatalities. In addition to my main results, I find mining cycles to be larger in districts with greater Scheduled Tribe (ST) population and smaller in districts with higher literacy rates. These are not unexpected results. A large proportion of mine workers belong to the Scheduled Tribes (Srivastava, 2005), making mining accidents more electorally sensitive in districts with a large voter base belonging to the same caste. Higher literacy rates are usually associated with the demand for greater political accountability and less myopic voting behavior, leading to larger political costs of accidents and deaths in general. In summary, this paper provides important new findings on electorally driven political behavior in an industry marred with severe corruption in India, analyzes the effect of both fixed and time varying factors on such behavior, and also sheds light on the consequences of such behavior.

The rest of the chapter is organized in the following manner. Section 4.2 briefly describes the contextual setting, focusing on the mineral industry in India, the politics around it and also discusses the development of the Naxalite movement since it's emergence in 1967. Section 4.3 presents a conceptual framework discussing the different channels that could generate political business cycles in mining. Section 4.4 describes the sources and construction of the dataset and presents summary statistics of key variables. Section 4.5 lays out estimation procedures and presents results from the state-level analysis, which shows the presence of election cycles in different aspects of mining activity over several decades. Section 4.6 is dedicated to the district-level analysis. It studies the effect of electoral competition and Naxalite conflict on the size of mining cycles. It also provides a comparative analysis of electoral cycles in Naxal violence in mining versus non-mining districts. Robustness checks and caveats follow in section 4.7. Finally, section 4.8 concludes.

4.2 Context

4.2.1 Politics in India

India is a federal democracy with its states having considerable power over their own government and policies. State legislative bodies are elected every 5 years to carry out administration in the 29 states of India. Every state had its first election in 1952, however for reasons such as coalition breaks, imposition of president's rule due to loss of confidence in the dispensation, midterm (unscheduled) elections have been held in many states leading to de-synchronization of the election cycles. From the point of view of analyzing the impact of political cycles, this is ideal since it allows for both cross-sectional variation across states/districts and temporal variation in the outcomes of interest that can be exploited to address my question. Furthermore, midterm elections are often unexpected events, which provides the basis for a natural placebo test; studying whether effects differ across scheduled and unscheduled elections (see Appendix Figure C.2 for an illustration). State elections are held at the constituency level, whereas the lowest administrative division at which I observe outcome measures are districts. The number of districts vary significantly across states, however each district on average is comprised of 9 constituencies. Therefore, in order to analyze heterogenous effects (of election cycles) generated by electoral competition, some form of aggregation of constituency level election results up to districts is necessary. This is discussed in greater detail in section 4.6.

4.2.2 Mining in India

Mining is a significant economic activity in India, contributing approximately 3% to its GDP (Indian Bureau of Mines, 2016) from over 80 different types of minerals. While this might appear to be a small contribution as a whole, mining is the major economic activity in mineral rich areas accounting for large shares of local incomes and generates significant employment. Nevertheless, large scale criminal activity exists in the sector, primarily through collusion between firms and state legislators. Such activities include: mineral prospections without permits, kickbacks for state legislators in exchange for mining permits, illegal extraction and violation of environmental and safety regulations. These activities are facilitated to an extent by the structure of the legal regime associated with mining and the role that state and central governments play in the extraction process. Indian states own all minerals within their geographical boundaries and Members of State Legislative Assemblies (MLAs) have significant control over majority of permits required before extraction can begin (Mohanty, 2017). While federal clearances are required for mining of some selected minerals, states have hold up power over these too. The royalties and taxes paid by mining companies go directly to the state and central governments. Furthermore, 7% of all mining leases in 2014 covered 75% of the total mining area, while 67% of leases covered only 8% of the area – and majority of these smaller areas are leased out to private enterprises (Mohanty, 2017). This is likely to ease collusion between firms and politicians. Indian mines also have an abysmal record when it comes to workplace safety compared to other developing countries such as Brazil and South Africa, making it one of the most dangerous professions in the country. A large proportion of accidents and deaths are caused by roof/side wall falls, events that are easily avoidable with better safety practices (Mandal and Sengupta, 2000).

4.2.3 Naxal Violence: Origin, development and characteristics

In this subsection, I offer a brief history of India's Naxalite conflict, and put forward arguments as to why violence intensity could be tied to the electoral cycle. I also provide anecdotal evidence on the importance of mining revenues as a tax base for the rebels.

Origin and development

India's Naxalite (Maoist) movement originated in a small village in rural West Bengal, "Naxalbari" in 1967, triggered by an attack on a tribal villager by local landlords. Thereafter, it gained momentum with support from key members of the Communist Party of India (Marxist). The period since has been marked by high levels of conflict between separate Naxalite groups, who spread across various states. In 2003, two major factions which promoted the idea of "Allegiance to armed struggle and non-participation in elections", merged to form the Communist Party of India (Maoist), which significantly intensified levels of violence. The popularity and strength of the movement is perceived to stem from underdevelopment in the affected communities. The Naxals are also believed to be banking on the grievances of the tribal population against mining activity, which has resulted in large-scale displacements in Maoist strongholds (Kujur, 2009). The declared objective of the CPI (Maoist) is to overthrow the government of India through protracted armed struggle and establish a liberated zone in the centre of India (Kujur, 2008). In 2006, the Naxalite problem was termed by the Indian Prime Minister Manmohan Singh, as the “the single biggest internal security challenge ever faced by our country".

Indian states have managed to deal with the problem to varying degrees. An outlier is Andhra Pradesh, which has undertaken a combination of effective development policies together with the use of specialised police force to deal with the issue. At the other extreme, some states have also undertaken semi-legal measures to try to curb this problem. A prime example of this is “Salwa Judum", a militia mobilized and deployed as part of anti-insurgency operations in Chhattisgarh with the government eager to flush the area of Naxalites in order to allow smooth operation of the mining companies. “Salwa Judum" herded villagers and tribals into makeshift camps which were rife in human rights violation.[74] In July 2011, the Supreme Court of India declared the militia to be illegal and unconstitutional and ordered it to be disbanded.

The Naxal affected areas qualify for multiple sources of central (federal) government funding for development purposes, such as the National Rural Employment Guarantee Scheme (NREGA) (which is a national scheme), as well as others that are specific only to the affected districts. This shows that the Indian government recognizes that underdevelopment constitutes an important aspect of this conflict. The extent to which state governments are able to use such funds however depends on the political capability and stability of the governments itself (Kujur, 2009). While these measures have reduced the potency of conflict in recent times, it still remains a significant challenge. Data released by the Home Ministry in April 2018 reports 90 districts in 11 states to be still affected by the conflict, of which 35 were identified as the worst affected.[75]

Relevant characteristics

Election cycles in mining intensity cannot be studied in isolation from the Naxalite conflict for multiple different reasons. I discuss important characteristics of the nature of conflict in this section to elucidate this.

[74] “Salwa Judum victims assured of relief". The Hindu. Chennai, India. 16 December 2008.
[75] Indian Express. New Delhi, India. 17 April 2018.

First, mineral resources are an important component of the rebels' tax base in many mineral rich districts. This has been documented in previous works as well (see Vanden Eynde, 2016). The Maharashtra State Home Minister R.R. Patil in 2010 accused the mining industry of funding the Maoist movement. There are numerous newspaper reports showing anecdotal evidence of this:

> "Where there is mining, there is Maoism, because where there is mining, there is more revenue, and where there is more revenue, there is more extortion," he added."Some of the best-known names in Indian industry are running businesses in the Maoist areas by paying off the Maoists. I don't want to name names, but these are the biggest names in Indian industry."(Interview of a local land rights activist in Hejda, Jharkhand, National Geographic Magazine, April 2015).
>
> "Around 10 armed members of Jharkhand Prastuti Committee (JPC, Naxal Wing) torched three dumpers, damaged an SUV and beat up a guard at a mine of Central Coalfields Limited (CCL) in the small hours today at Charhi, some 32km from the district headquarters, the second such attack in the area within a week. The JPC had been demanding levy from the company for some time and the incident seemed to have been orchestrated to mount pressure on authorities."(The Telegraph, February 2016).

The rebels tax the mining sector in mineral rich districts through violence and mining companies are believed to be susceptible to their demands. Therefore, to the extent that Naxals actively boycott elections through both violent means and civilian collaboration, minimizing the resource base of the rebels might be an important consideration for politicians during elections.

This brings me to the next important characteristic of the Naxalite violence, boycott of national and state assembly elections:

> "As usual, we have appealed to people to boycott the elections because they are a farce. Elections only renew five-year tenures of loot and torture by the elected representative in the present system. Like always, this time too, the government has deployed a huge number of security forces in the name of conducting free and fair elections, which are already exploiting and torturing people. Attacks on villages in the name of search operations, arrests, beating up people, fake encounters are consistently on. Therefore, I can only say that when the government tries to defuse our poll boycott movement through crackdown on the people, there will be certainly a counter to it". (CPI Maoist Special Zonal Committee Secretary to The Times of India, October 2013.)

Lastly, the rebel groups at the same time levy harsh punishments on villagers for collaborating with the police :

> “Maoists killed a middle-aged couple suspecting them of being police informers and dumped their bodies on a road in Madhubanthana area of Giridih district in the small hours of Friday. When villagers woke up in the morning, they found the bodies lying about a hundred metres from each other with their throats slit. The deceased have been identified as residents of Banpura village in Madhubanthana area". (The Telegraph, May 2018)

In summary, mining revenues form an important component of the rebels’ tax base in mineral rich areas, and they systematically appeal for the boycott of elections to civilians. The government on the other hand makes efforts to counter this through collaboration with civilians, some of who become police informants. Reducing the tax base of the rebels during elections (when they are likely to value resources more) therefore might help the government in their objective as well.

4.3 Conceptual framework

In this section, I develop a brief conceptual framework to argue how the incentives faced by politicians/state legislators could lead to political business cycles in mining. As described in the previous section, state legislators have significant involvement in the mining sector, both due to the regulatory structure of the industry as well as through collusion with mining firms. Mining is India’s most dangerous profession, with a death every third day in coal mines (Sasi, 2014). Anecdotal evidence suggests that mining accidents affect the popularity of incumbent politicians with opposition parties often capitalizing on such events (Ray, 2022; Jai, 2019). Incumbents may thus have the incentive to manipulate activity in a way that maintains their rent-seeking ability but at the same time does not hurt their electoral prospects (by shifting activity away to earlier years in their term but not lowering overall activity). Politicians facing greater re-election threat would have higher incentives to undertake such manipulation. Therefore, we are likely to observe more pronounced cycles in areas that are electorally competitive, as well as where the voter base is likely to be more sensitive to these events for e.g. in districts with a higher Scheduled Tribe (ST) population, since a large share of mine workers tend to be from the same community (Srivastava, 2005). Since mining is a capital-intensive industry (with large capital-labor ratios), the dis-employment effects of reduced activity is likely to be small too, giving politicians greater leverage over the degree of manipulation.

Apart from minimizing accidents and fatalities, a second reason why incumbent politicians (in mineral rich areas) could benefit from lowering mining intensity during elections is because it might help to reduce the intensity of Naxalite conflict. As described in section 4.2.3, the rebel groups target elections with violence, and as a result are likely to put greater value on resources closer to elections to fund their activities. Extortion from mining companies constitutes an important part of their tax base. From the perspective of an incumbent politician, being able to control such violence signals positively about their ability to maintain law and order. I thus hypothesize that higher collaboration with civilians is increasingly important for both the government and rebel groups during election years – leading to tensions between the two parties with respect to controlling the population. The government can reduce the rebels' resource base more effectively in mining districts compared to non-mining ones. This is because in non-mining districts their resource base tends to be more disaggregated (for e.g. small business owners, contractors and farmers). In equilibrium, non-mining districts are therefore likely to experience more violence before elections (relative to mining districts), since the tax base of rebel groups remain less affected in these areas. I test these predictions empirically in the district-level analysis.

4.4 Data and descriptive statistics

I combine seven different sources of data, which are described in this section. Data on state-level variables are generally available over a longer time series, whereas district-level measures for a shorter time horizon. Nevertheless, this is mitigated by the fact that I observe elections in over a 100 districts, across 15 states that are not synchronized with each other. An additional advantage of having a large number of districts is the inclusion of district fixed effects, which only a few past papers include.

State-level mining output data from 16 major Indian states (between 1960-2000) are obtained from EOPP Indian States Data (Besley and Burgess, 2004), which I supplemented with additional data from the Department of Statistics, Planning Commission of India. I digitized mining lease data from 23 states (including the 16 for which I have output data) between 1995-2004 from the website of the Indian Bureau of Mines. Finally, I obtained state-level data (for 21 of the 23 states with mining lease data) on fatal mining accidents between 1998-2015 from the Ministry of Labour and Employment, which I also digitized. While the state-level analysis is important in understanding whether the impact of the electoral cycle is significant enough for the effects to show up on aggregate, analysis only at such a high political jurisdiction level severely limits understanding of heterogeneous effects of crucial factors such as electoral competition and conflict propensity. Therefore, I focus on the district-level data for this analysis.

The Directorate General of Mines Safety (DGMS), Ministry of Labour and Employment, records injuries and deaths as a result of both serious and fatal accidents at the district-level for both coal and non-coal mines in India. These are made available online in a yearly issue, "Statistics of Mines", that I have digitized for the period 2010-2015.[76] Figure 4.1 shows all districts included in the empirical analysis. State elections data at the constituency level come from the Election Commission of India's (ECI) website. These data include information on election results and candidate characteristics such as names, caste, wealth and criminal history. Finally, I obtained geocoded data on Naxalite violence intensity at the district level between 2005-2017 from the Uppsala Conflict Data Program (UCDP) and South Asia Terrorism Portal (SATP). These include information on civilian, security forces and rebel deaths resulting from clashes between rebel groups and security forces. I consider the construction of this data set, particularly of the data on mining fatalities and leases, an important contribution of this work.

Sub-figure (a) includes all districts included in the analysis of election cycles in mining fatalities. Amongst these, districts seriously affected by Maoist violence (as classified by the Union Ministry of Home Affairs in 2008)[77] have been marked in red and the rest in green. In the empirical analysis, I study whether mining cycles are larger in Maoist affected districts, since the cost of more mining closer to elections is likely to be higher in these areas, owing to the greater threat of violence. I only use pre-sample classification to avoid confounding by the fact that greater mining activity over the sample period could in turn intensify violence in these areas. Note that the government does not update the list of affected districts in a consistent manner. Updates are made only when there are substantial changes in the number of affected districts and therefore the classification in sub-figure (a) is the latest update I was able to find before 2010. Sub-figure (b) marks all districts with at least one reported death resulting from clashes between security forces and rebels between 2005-2017. Amongst these, mining districts are marked in orange. Data on each event (clash) at the district level were obtained from UCDP, and then aggregated to the district-year level. This sample is used to study election cycles in violence and how they might different across mining and non-mining districts. Table 4.1 provides summary statistics of

[76] Versions of the "Statistics of Mines" are not available in a consistent manner before this period.
[77] LWE Violence/Under Influence Districts, 2008 (http://www.satp.org.)

key variables for both, the state- and district-level analyses.

4.5 State-level analysis

This section first presents estimating equations for the state-level analysis and then empirical results on the impact of the election cycle on three different measures of mining activity: mineral licensing, mining output and fatalities. In order to estimate the impact of elections on mining intensity the following simple model can be run,

$$Y_{st} = \phi E_{st} + \alpha_s + \gamma_t + \delta_s t + u_{st} \tag{4.1}$$

where Y_{st} is the outcome of interest in state s at time t. E_{st} is a dummy taking the value of 1 if there is an election in the state s in year t, and 0 otherwise. α_s is a state fixed effect, γ_t is a year effect and $\delta_s t$ is a state-time trend. The impact of elections is given by ϕ. This model however is too simple and unlikely to be very informative for a couple of reasons. Firstly, elections in Indian states are not strictly scheduled for a specific time of the year every cycle. Majority of the outcomes of interest, such as mining fatalities and output could rise immediately in the post-election period within the election year, which could potentially be confounding. Secondly, and more importantly, this model does not allow estimation of the impact of the entire political cycle, which is important. A simple extension to equation (1) can achieve this. The specification is,

$$Y_{st} = \sum_{k=-4}^{-1} \phi_k E_{st}^k + \alpha_s + \gamma_t + \delta_s t + u_{st} \tag{4.2}$$

where I include dummies for each year of the four years before a state election, omitting the year of election to avoid multicollinearity. All results should therefore be interpreted as relative to the election year. While this model is able to deal with the two caveats mentioned earlier, a potential endogeneity concern still remains. Chief Ministers of Indian states have the authority to move elections early within a year from the scheduled month. Such decisions are non-random and could be correlated with the state of the economy. This is generally a problem for questions that deal with politicians manipulating policies to stimulate the economy during elections, because if Chief Ministers move elections dates according to the state of the economy, it could lead to a spurious relationship between elections and the outcomes of interest. This is much less of a concern in my case, since decisions regarding mineral licensing and production happen at local administrative levels, are unlikely to be coordinated amongst politicians across constituencies/districts and therefore not affect major political decisions such as election dates. Moreover, my results, in contrast to previous works, find the shape of the cycle to be exactly opposite to those existing in the literature alleviating concerns of a spurious relationship between elections and factors positively correlated with a booming economy. One could nevertheless think of an alternative story where moving elections to a later date might be beneficial in the aftermath of a mining accident. Elections however cannot be postponed and therefore such situations are much less of a concern. Nevertheless, I discuss three possible ways to deal with this endogeneity issue, all using instruments. For each state, a placebo election cycle can be used to instrument the actual observed cycle. Call t_s the year in which an election was held in state s before the state appears in the sample. The placebo cycle would assign elections to $t_s + 5$, $t_s + 10$ and so on. The placebo cycle can then be used to instrument the actual cycle to obtain causal estimates.

The alternate strategy, suggested by Khemani (2004), similar in vein to the previous strategy is to “update" the placebo cycle suggested above, every time there is an off-cycle election (call this the scheduled cycle). For example, if state s had elections in years 0, 5, 9 and 11, the placebo cycle will take the value 1 in years 0, 5, 10, 14 and 16. This provides considerably more first stage power and therefore has more “relevance" as an instrument.

Alternatively, the scheduled cycle itself can be used as the right hand side variable. It is equivalent to using the scheduled cycle as an instrument for the actual cycle. It is worth mentioning at this point that majority of my results are insensitive to any of these choices (Scheduled or IV). Furthermore, the OLS and IV results are similar for most of my specifications. The most relevant outcome for which endogeneity of elections might be a concern is mineral licensing, since a booming economy could lead to a spurious relationship between licensing and election timing. I therefore present, OLS, IV and scheduled cycle results for election cycles in licensing, but do not find large differences in this case either. In other cases, I report results using the scheduled cycle as my right hand-side variable, since in general it is more efficient than the IV estimator.[78]

[78] In the few instances where results from the IV approach are different from using the scheduled cycle itself as the main regressor, the

The results are reported in Table 2. Political cycles in mineral licensing are clearly observable. Note that the outcome variables have been logged and therefore the coefficients should be interpreted as percentage change over election levels. Across majority of the specifications, there is a statistically and economically significant negative effect of the penultimate year of the election cycle on the total number of active mining leases (logged) in a state. The results are robust to the addition of state specific time trends. Since the state-level analysis comprises of a small number of clusters (23), I also present wild cluster bootstrap p-values along with state-clustered standard errors. The statistical significance of the results at the 1%, 5% or 10% level are determined by the bootstrap p-values. The most important takeaway from this analysis is the fact that the extensive margin (mining leases) is important in generating mining cycles.

Table 4.3 reports the effect of political cycles on mining output at the state level. Mining output data from 16 major Indian states over the period 1960-2000 are used for this analysis. I report results for log of total output and output per capita. The trends are robust across all the specifications. Results in Table 4.3 show mining output to increase sharply after an election (4 or 3 years from the next election), and this effect attenuates leading up to the next election. The inclusion of state-specific time trends reduces the magnitude of the coefficients; but the effects remain economically significant.

Finally, to conclude the state-level analysis, I present evidence of election cycles in mining fatalities in Table 4. The sample used for this comprises of 21 of the 23 states featuring in the analysis of mining leases (including all 16 used for the output tables), but the Ministry of Labor and Employment have only made data on state-level mining fatalities available from 1998 and therefore the sample includes observations from 1998-2015. I use Poisson regressions since these are count outcomes. A similar pattern is observed. Results show significantly fewer accidents in the penultimate year of the term, which are robust to bootstrapping the standard errors (Kline and Santos, 2012 score bootstrap).

While the effects 4 years away from a scheduled election were positive and significant for some outcomes with respect to state-clustered standard errors, bootstrap p-values make them statistically insignificant. Nevertheless, the marginal effects are economically significant.

Marginal effects imply an additional 4 fatal accidents on average in the year after an election and

estimates are more conservative in the latter case. This is not an issue in the district-level analysis since there was no off-cycle election over the time period considered.

a reduction of approximately 5 fatal accidents in the year immediately before an election, relative to the election year. Given an average death toll of 28 (see Table 4.1, Panel A) conditional on at least one accident in a state in a year, this is an fairly large effect. The results for mining fatalities are robust to a OLS model with the outcome variables log transformed (Appendix Table C.1). Overall, the findings from this section point to the presence of political cycles in different aspects of mining activity between Indian state elections, laying a strong foundation for the district-level analysis aimed at understanding the mechanisms generating these patterns.

4.6 District-level analysis

In this section, I explore different factors that may explain election cycles in mining in India. The unit of analysis is districts, which is equivalent to a U.S. county in terms of geographic coverage. There are two main advantages of performing this analysis at the district level. Firstly, greater electoral competition on aggregate at the state-level might not be a significant determinant of the size of mining cycles, owing to the decentralized structure of India's mining sector, as described in section 4.2.2. Therefore, focusing on the mining districts is necessary. Secondly, there is significant variation within states with respect to the proportion of districts that are affected by the Naxalite conflict, whereby only a state-level study of the impact of conflict on mining cycles will lead to loss of important within state variation. Digitization and cleaning of mining accidents and ancillary electoral outcomes data is still a work in progress. I restrict the analysis of mining cycles at the district level to the years 2010-2015 in order to ensure the use of a consistent, balanced panel. None of the 15 states included in the district level analysis had a off-cycle election between 2010-2015 whereby scheduled cycles and actual cycles are equivalent in this case. I first report results on the size of mining cycles at the district level and then proceed to the analysis of factors driving these cycles.

The specification used for this analysis is,

$$Y_{dst} = \alpha_d + \sum_{k=-4}^{-1} \phi_k E_{st}^k + \lambda_{Rt} + u_{dst} \tag{4.3}$$

where Y_{dst} is the outcome variable of interest in district d of state s at time t. α_d is a district fixed effect and $\sum_{k=-4}^{-1} E_{st}^{k}$ are dummies denoting years to the next scheduled election in state s at time t. Once again, the election year is omitted. I include region-year fixed effects (λ_{Rt}) in the district level analysis.[79] The states of India are divided into six regions by the Reserve Bank of India. Region-year fixed effects (λ_{Rt}) control for macroeconomic fluctuations that could affect mineral production differently across regions. Results from a total of 104 districts are reported in Table 4.5. The standard errors are clustered at the state-year level.[80] In Column (1), the outcome variable "Any Bad Event" is a dummy coded 1 if at least one person was either seriously injured or killed in a mining field in a district in a given year. Coefficients imply up to 24 percentage points increase in the probability of such an event in the three years immediately after an election relative to the election year. The effect in the penultimate year of the term is much smaller and statistically insignificant. Columns (2)-(5) consider the intensive margin. The outcome variables are total accidents and total casualties (deaths and serious injuries). The results are robust to considering both a OLS and Poisson framework as shown. In the OLS specifications, outcome variables are logged and coefficients should therefore also be interpreted as percentage point changes over election levels.[81] Marginal effects from the Poisson model imply 3 additional accidents in the year immediately after an election which is similar to the OLS estimate.[82] Overall, I find a large impact of the political cycle on mining fatalities at the district level, that aggregates to states.

4.6.1 Electoral competition and mining cycles

When studying factors that determine the size of mining cycles, it is important to consider district characteristics that are both fixed and time varying. I first examine whether electoral competition is a determining factor and thereafter consider fixed factors that are likely to be important. Given that elections are the basis for mining cycles, electoral competition should be a strong predictor of the size of cycles. As mentioned earlier, the lowest administrative level at which outcomes are observed are districts, whereas state-level elections are held at the constituency level and therefore, some form of aggregation is necessary. MLAs (Members of State Legislative Assemblies) have significant control over granting of mineral

[79] State-year effects are of course collinear with political cycle dummies.

[80] State-year is the treatment level since it determines the year of the political cycle a district is at.

[81] Since the outcome variables are logged, the coefficients can be interpreted approximately as percentage changes.

[82] In order to ensure that I do not drop observations with 0 values, I add 1 to each observation before the log-transformation. Adding smaller numbers does not affect the results.

licenses and are also likely to be able to manipulate mining activity, irrespective of whether they belong to the state ruling party or not (Mohanty, 2017; Asher and Novosad, 2018). I thus create a variable "Close Election Proportion" (C_{dst}), which denotes the proportion of constituencies within a district where the margin of victory in the previous election was no more than a certain threshold level. I then modify Equation (3) in the following manner to first examine whether political competition affects mining accidents in general (i.e. on average; call this the "level effect"),

$$Y_{dst} = \alpha_d + \sum_{k=-4}^{-1} \phi_k E_{st}^k + \pi C_{dst} + \lambda_{Rt} + u_{dst} \tag{4.4}$$

where Y_{dst}, α_d , E_{st}^k and λ_{Rt} are defined exactly as before. C_{dst} is the proportion of close elections in a district in the previous state assembly election. The variable C_{dst} is constructed in the following manner. I first denote an election to be close if the victory margin of the winning candidate in a constituency was no more than 10% of total polled votes[83](all results are robust to considering this threshold to be any value between 5%-10%) and then calculate the proportion of constituencies in each district that satisfy this criteria. The value of C_{dst} for a district is therefore,

$$C_{dst} = \frac{1}{N_{cd}} \sum_{k=1}^{N_{cd}} \mathbb{1}(Margin_{kt} < 10\%) \tag{4.5}$$

where, N_{cd} is the number of constituencies in district d, and $Margin_{kt}$ denotes the margin of victory in constituency k in the previous election. I choose the 10% threshold because the distribution of close elections that this generates is not too skewed in any direction since that could lead to the results being driven by extreme values. As mentioned, the results are not sensitive to this choice. The average win margin in constituencies that are considered to have had close elections by this criteria is about 4%. The district-level distribution of close elections is shown in Figure 4.2.

The specification in equation 4.4 is restrictive in the sense that it will not detect concentration of mining fatalities in particular years of the political term. To the extent that voters are myopic, making mining fatalities more costly during elections, politically competitive districts could simply concentrate greater activity to earlier years of the electoral term as opposed to a reduction in intensity uniformly over

[83]The average margin of victory at the constituency level across all districts in my sample is 13%.

the term (“shape effect"). In order to elucidate whether this is the case I also run the following model,

$$Y_{dst} = \alpha_d + \sum_{k=-4}^{-1} \phi_k E_{st}^k + \tilde{\pi} C_{dst} + \sum_{k=-4}^{-1} \gamma_k (E_{st}^k \times C_{dst}) + \lambda_{Rt} + u_{dst} \tag{4.6}$$

where I interact C_{dst} with each of the political cycle dummies. This allows for a different relationship between political strength and mining accidents for every year of the electoral term.

There are 7 districts for which I could not find election results for all constituencies, consistently over multiple elections. I do not include them in this analysis. The results are presented in Table 4.6. Once again, standard errors are clustered at the state-year level. The outcome variable in each case is the total number of events in a district-year (logged). I choose to present my results in this form for easier interpretation of the coefficients. As shown in Table 4.5, marginal effects from a Poisson model are very similar. In each case, I also present baseline results for comparison. The baseline coefficients are however not significantly different from those in Table 4.5, as can be observed in Columns (1) and (4) of Table 4.6. In columns (2) and (4), the coefficient on C_{dst}, π (from specification 4.4), is estimated to be 0, ruling out the possibility that electoral competition affects average levels of fatalities over the political cycle. This is important since it implies that electorally competitive districts are not inherently different in terms of characteristics that affect mining intensity in general. I discuss this in more detail at the end of the section.

Columns (3) and (6) present results based on specification 4.6. These estimates however clearly show that political cycles in fatalities are significantly larger in districts that experienced greater competition in the previous elections. Note that the overall effect in any year of the term is given by the sum ($\phi_k + \tilde{\pi} + \gamma_k$). The coefficients ($\gamma_k$) on the interaction between C_{dst} and the political cycle dummies are positive, statistically significant and larger than the coefficient on C_{dst} ($\tilde{\pi}$) alone in the earlier years of the cycle, implying larger positive changes in fatalities from the election year in competitive districts. The coefficients on the interactions are however much smaller than $\tilde{\pi}$ and statistically insignificant in the last two years of the term whereby competitive districts also exhibit a sharper drop in fatalities as the next elections approach. The coefficient $\tilde{\pi}$ alone in this specification estimates the impact in the next scheduled election year, which is negative and significant for both outcomes. The largest difference generated by political competition is observed in the year immediately after an election and in the next scheduled election year, pointing to tactical manipulation by authorities responsible for mining operations. This is strong evidence of a setting where voters are myopic. Authorities take advantage of this and manipulate mining activity in a way that is likely to minimize electoral loss.

The results in Table 4.6 are best presented through event study graphs that provide a visual representation. In Figures 4.3 and 4.4, I present log predicted values of accidents and casualties in mines over the electoral cycle; these are based on the model in equation 4.6. The dotted lines represent 95% confidence intervals. I present a comparative analysis in each case between a notional district with 0% close elections and one with 50% close elections[84] in the previous state assembly polls. Figures 3 and 4, based on results from Columns (3) and (6) of Table 4.6 respectively, show a significantly larger increase in fatalities in the post-election period in the district with 50% close elections compared to the district that had no close elections in the past cycle; relative to their respective election levels. In fact, log predicted number of accidents is also larger in the former. However, in the years leading up to the next election, the drop in fatalities in the district with 50% close elections is much sharper and overall accidents are fewer as well. In sum, figures 4.3 and 4.4 clearly show that electoral competition magnifies the size of mining cycles.

A couple of observations from these graphs merit additional discussion. To the extent that districts with high electoral competition are characteristically different in unobservables from those with low competition (time varying factors that are not accounted for through fixed effects), particularly with respect to factors such as "ability" to conduct mineral prospections, licensing with lesser opposition (political) pressure, one could expect a "level effect" of electoral competition on mining intensity over the electoral term. I do not find evidence of this. However, larger mining cycles in competitive districts ("shape effect"), without a significant difference in the overall level of intensity cannot be attributed to such "ability" factors. The dynamics are therefore important in suggesting that there is indeed a behavioral effect generated by electoral competition; one where competitive districts concentrate greater activity to earlier years of the political cycle and undertake additional measures to avoid fatalities only during elections.

[84] Please refer to Appendix Figures C.8 and C.9 for a comparison with a notional district with 90% close elections.

4.6.2 Mining cycles in the Red Corridor

The "Red Corridor" of India is comprised of districts that are affected by the Maoist insurgency. They are also some of the most mineral rich districts in India, and as motivated in section 4.2.3, access to key mineral resources is an important source of funding for the Maoists. With the declared agenda of overthrowing the Indian government through armed struggle, boycotting elections form a core component of the Maoist agenda. Therefore, whether mining districts are part of the Red Corridor or not seems to be an important characteristic to consider when studying the size of mining cycles. I address this question in this section.

Why should one ex-ante believe that mining cycles may be different in the Red Corridor compared to other districts? The mining industry in India is particularly vulnerable to manipulation by local politicians and to the extent that rebel groups value resources more during elections, governments could optimally try to shut this channel of funding down to reduce violence. In comparison to other sources of extortion such as small industrialists, contractors, and even poor villagers,[85] which politicians do not directly control, mineral extraction can be reduced more effectively. Therefore, controlling for the mean level of mining intensity across districts, those in the Red Corridor should on average concentrate greater mining activity to years away from elections. Table 4.7 formally tests this. I classify districts to be part of the Red Corridor (a dummy variable) if they were declared in 2008 to be amongst the Maoist affected districts by the Union Ministry of Home Affairs. I use pre-sample classification to avoid the possibility that mining activity over the sample period could intensify violence, which in turn could lead to the inclusion of certain districts into this group, creating an endogeneity issue. Figure 4.1(a) marks mining districts that are part of the Red Corridor in red. The effect of being in the Red Corridor on the size of mining cycles is estimated in exactly the same way as that of electoral competition; except that it is a fixed district characteristic and I therefore only include interactions with election cycle dummies when district fixed effects are included (the main effect is of course absorbed in the district effects) but add a separate dummy for Red Corridor with state fixed effects. The specification (with district fixed effects) is,

$$Y_{dst} = \alpha_d + \sum_{k=-4}^{-1} \phi_k E_{st}^k + \sum_{k=-4}^{-1} \gamma_k (E_{st}^k \times RC_d) + \lambda_{Rt} + \epsilon_{dst} \tag{4.7}$$

[85] Srivastava (2009) "Extortnomics : Maoists raise Rs 2000 crore every year".

where RC_d takes the value 1 if the district is classified as part of the Red Corridor and 0 otherwise. All other variables are defined exactly as before. Table 4.7 presents results with total accidents and total casualties (logged) as the outcome variables respectively.

The first important result in Table 4.7 is that mining cycles exist in districts that are not classified as part of the Red Corridor. Hence, the objective of minimizing fatalities around elections is significant on its own in generating these cycles. Second, the hypothesis that districts in the Red Corridor are likely to observe greater concentration of mining activity in years away from elections (larger cycles) is also observed to be true.

In columns (2) and (5) (specifications with district fixed effects), interactions between the Red Corridor dummy and the political cycle dummies are positive and statistically significant in earlier years of the cycle, implying a larger cycle on average in districts of the Red Corridor. Though positive, the coefficient on the interaction term $One\ Year \times Red\ Corridor$ is much smaller and statistically insignificant. Note that this is not mechanically driven by the fact that districts in the Red Corridor have higher mining intensity in general. The dependent variable (casualties or accidents) is in logs and therefore coefficients represent percentage point changes over election years. Figure 4.5, based on column (3) of Table 4.7 depicts this. In each case log predicted accidents in the election year has been normalized to 0 and y-axis values in each year of the political cycle represent changes from election levels. Districts in the Red Corridor experience a significantly larger increase in the number of accidents in the years immediately after an election compared to other districts. In the penultimate year of the term however, the difference in the relative changes over the election year is insignificant across the two sets of districts.[86]

A similar pattern is observed in the specifications with state-fixed effects (Columns (3) and (6)). Since there is variation within states with respect whether districts are part of the Red Corridor or not (refer to Appendix Figure C.3), I include a dummy for Red Corridor (in order to estimate its overall effect), in addition to it's interaction with election cycle dummies. Note that the interaction terms are positive, though statistically insignificant and reduce in magnitude over the electoral cycle, whereby the coefficient on the Red Corridor dummy alone (which is negative and significant), dominates (when they are summed) as elections approach. This implies larger cycles in districts of the Red Corridor i.e., they experience

[86] The equality of the coefficients $One\ Year$ and ($One\ Year + One\ Year \times Red\ Corridor$) cannot be rejected.

a larger drop in fatalities in the pre-election period and a sharper increase in the post-election period (please refer to Appendix figures C.4 and C.5 for a visual representation).

Overall, I find strong evidence of the size of the mining cycle to be significantly larger in districts in the Red Corridor, which is consistent with the hypothesis that constraints on mining activity closer to elections are on average greater in Naxal affected areas.

4.6.3 Election cycles and Naxalite conflict intensity

Why are mining cycles larger in the Red Corridor? To answer this question it is first necessary to understand the dynamics of Naxalite conflict intensity over state electoral cycles. As motivated in section 4.2.3, the rebels traditionally follow a election boycott strategy, often through the use of violence and it is therefore likely that levels of violence increase as elections approach. To the extent that governments are more sensitive about their reputation in the election season, increased levels of violence during this period is likely to benefit the Maoist agenda. Marginal value of resources is thus likely to be higher for rebel groups during polls, leading to greater conflict in pursuit of acquiring more resources.[87] In this sub-section, I first show empirical results supporting this hypothesis and then study whether the intensity of electoral violence is different across mining districts and non-mining districts.

Table 8 presents results from poisson regressions of different measures of conflict intensity on the political cycle dummies. In columns (2),(5) and (8), I include all districts (mining and non-mining) with at least one reported casualty as a result of clashes between rebels and security forces between 2005-2017. In Column (2) the outcome variable total deaths is the sum of security forces, rebel and civilian deaths from all clashes in a district in a particular year.[88] A large and statistically significant increase in total deaths can be observed in the penultimate year of the term. As mentioned previously, elections in Indian states could be held in any month of the year, whereby regressions of conflict outcomes on a election dummy alone might not be the best measure of the effect of elections. This is especially important in this case, since unlike other studies that consider government policies which are fixed over the year (fiscal policy, agricultural credit), there could be large variations within a year in the intensity of conflict.

[87] A large proportion of weapons that Maoists acquire are through attacks on security forces (Prakash, 2014).

[88] Civilians are not the intended targets of rebel groups in these particular clashes but casualties are caused if they are caught in the crossfire.

The one year before effect is thus likely to be “cleaner".[89] Empirically, a regression of conflict outcomes on the election dummy alone shows no significant effect of the election year, however a dummy for the penultimate year alone is positive and highly significant across all measures of conflict between security forces and rebels. Columns (5) and (8) show effects on the number of rebel deaths and security forces deaths respectively. Similar to Column (2), large positive effects can be observed in the penultimate year. Note that there is no statistically significant difference in average intensity of violence across mining and non-mining districts within states. In Columns (1), (4) and (7) I regress different measures of conflict intensity on a “Mineral Dummy" coded 1 if the district has mineral deposits and 0 otherwise. I include state-fixed effects thereby comparing districts within each state. None of the coefficients are precisely estimated. This result is not to be confused with previous results in the literature that find mineral presence to be a determinant of conflict onset and intensity. Previous works such as Hoelscher et al. (2012) consider the extensive margin of conflict and show that in general, districts that have mineral deposits are more prone to the presence of Naxals compared to districts that do not. My results are closer to Ghatak and Eynde (2017) who show that within affected states, mineral presence does not significantly affect conflict intensity. In fact, the analysis here goes a step further. I find no significant difference in the levels of violence, comparing only affected districts within states, but with and without mineral presence.

In columns (3), (6) and (9) I interact each of the political cycle dummies with the “Mineral Dummy". The coefficients on the interaction terms $One\ Year \times Mineral\ Dummy$ are negative and significant for total deaths and security forces deaths. The null hypothesis of a zero net effect for none of the conflict outcome measures can be rejected in the penultimate year of the political cycle (relative to the election year) in mining districts ($H_0 : One\ Year + One\ Year \times Mineral\ Dummy = 0$; p-values 0.715, 0.728 and 0.178 respectively). However, the coefficients on the $One\ Year$ dummy alone remain strongly significant and also increase in magnitude across all the different measures of conflict intensity. Overall, the results show that as elections approach the increase in levels of violence is significantly stronger in non-mining districts. In other words, mining districts exhibit a smaller cycle of violence.

[89] Violence levels are likely to increase until elections are over and not continue in the months after. Therefore, the effect in the penultimate year is likely to be uncontaminated. Furthermore, to the extent that increased clashes between rebel groups and security forces are over the capture of resources required to boycott elections (weapons, explosives etc.), the effect in the penultimate year is likely to be larger. It would of course be different for clashes between civilians and rebels with the latter trying to physically prevent voters from voting.

Figure 4.6 plots log predicted total deaths (security forces, rebels and civilians) over the electoral cycle in mining and non-mining districts (Please refer to figures C.6 and C.7 in the Appendix for plots with rebel and security forces casualties as separate outcomes). It clearly shows that conflict intensity is significantly higher in the year before elections but only in non-mining districts. In mining districts, log predicted total deaths is not significantly in the election year from any other year of the political term.

A pattern consistent with the idea that governments undertake active measures to minimize the rebels' resource base during elections, emerges from the analyses in Tables 4.7 and 4.8. As described earlier, politicians are likely to be able to control mining activity more than other sources of extortion of rebels, and therefore optimally reduce mining intensity during elections to minimize disruptive activities by rebel groups. Not only do the rebels levy monetary taxes on mineral companies, anecdotal evidence of them looting explosives from mining fields also exist,

> "Naxal guerillas raided a mining facility of Steel Authority of India Limited (SAIL) in Chhattisgarh's Durg district late Thursday evening and looted about two tonnes of explosives." (Business Standard, 2013)

Government officials, aware of the rebels' targets have considered removing mining explosives from public sector mining fields.[90] In Table 4.9, using data on kilograms of explosives used in coal mines over the period 2010-2015 across 38 major coal mining districts in India (obtained from the Directorate General of Mines Safety, Ministry of Labour and Employment), I find significant reductions in the use of coal explosives in the election year as well the one before that, but only in districts of the Red Corridor. Note once again that the results are not mechanically driven by greater use of explosives in the Red Corridor in general, since point estimates denote percentage changes over other years. 38 districts are unfortunately not enough to estimate the impact of the entire political cycle. I only add year effects instead of region-year fixed effects in this case, since regional variation in the location of major coal mines in India is not large.[91] The coefficients on the interactions ($Election_{-1} \times Red\ Corridor$) are negative and larger than those on the $Election_{-1}$ dummies across all outcomes. They are significant for log detonators and non-permitted explosives,[92] implying a reduction in the use of coal mine explosives in Maoist affected

[90] "Mining Explosives Attract Maoists" (https://www.downtoearth.org.in/news/mining-3510).

[91] Results are nevertheless robust to the inclusion of region-year effects.

[92] "Permitted explosives are especially designed to produce a flame of low volume, short duration, and low temperature"(U.S. Bureau of Mines).

districts before elections.[93] Though weak, positive coefficients on the $Election_{-1}$ dummy alone could be driven by the fact that explosives are moved to districts without Maoist influence.

In summary, the empirical findings from this section are consistent with existing anecdotal evidence and provide strong support for my hypothesis that the objective of minimizing Maoist activity during elections is an important determinant of the size of mining cycles. The resource base of rebel groups is different across mining and non-mining districts and the government is better equipped to affect their funding base in regions with key mineral presence. A reduction in mining intensity around elections therefore results in less increase in violence in mining districts.

4.7 Extensions

4.7.1 Robustness

In this section, I perform robustness checks to ensure that the results are not sensitive to the choice of econometric models and also provide additional evidence that strengthens my main findings.

Since specifications with mining fatalities and conflict casualties as the left hand-side variable involve count outcomes, it is important to ensure that the results are not sensitive to either the choice of a Poisson or OLS framework. In Table 4.5, I show district level cycles in mining fatalities are robust to considering both a Poisson and OLS model. In Table C.1 of the Appendix, I show that results from Table 4.4 (state-level cycles in fatal accidents) are also robust to a OLS framework with the outcome variables log transformed. While Table 4.8 estimates political cycles in Naxalite violence using a Poisson framework, Figures 4.6, C.6 and C.7 plot results from OLS specifications with logged outcome variables. The patterns remain robust. Finally, unlike previous works, it is not obvious that the election year should be the reference category in this study, as discussed earlier. In Table C.2 (Appendix) I run specifications with dummies for the election year and the years immediately before and after, in order to estimate the impact of the political cycle. The overall results and marginal effects are largely similar.

A large proportion of the population affected by mining activity, both in terms of mining induced physical displacement and mining casualties by virtue of being mine workers, belong to the Scheduled Tribes (ST) (Srivastava, 2005). Therefore, districts with a high proportion of ST population can be ex-

[93] A zero net effect however cannot be rejected for Non-permitted explosives (p-value 0.32).

pected to have larger mining cycles, since mining fatalities will be more electorally sensitive in constituencies with a large base of ST voters. Note however that the proportion of ST population in districts has been documented to be a strong predictor of Maoist violence (Gomes, 2015; Hoelscher et al., 2012) whereby a larger cycle in these areas could simply driven by ST population shares acting as a proxy for Maoist violence and/or an interaction of both of these effects.[94]

I cannot differentiate between these channels or comment on the relative importance of each; but at the least, larger cycles in districts with high ST population proportions would act as a robustness check of the results in Table 4.7 and potentially also imply that the composition of voter base matters as a determining factor. Table C.3 in the appendix tests this. I present results for both accidents and casualties. In each case, I first show the impact of elections on outcomes alone (election dummy) and then include an interaction between the election dummy and the proportion of ST population (obtained from the 2011 Census of India) in each district. While this is after my sample for fatalities begin, it is not a major concern since caste compositions are fairly stable across districts in India and are unlikely to have changed significantly over such a short period of time. The coefficients should be interpreted as percentage changes over non-election years.[95] For both accidents and casualties, the interaction terms are negative and economically significant (though statistically insignificant),[96] implying a sharper drop in fatalities in districts with greater ST population. This is important since it adds credibility to both, the effect of electoral competition and conflict on mining cycles. It acts as a robustness check for larger mining cycles in the Red Corridor, and potentially shows that voter demographics are also important.

Another factor to naturally consider as important in determining the size of mining cycles is literacy rates. The mining districts in India are amongst the worst performing in terms of socio-economic outcomes, which can facilitate larger cycles if politicians are not adequately held accountable for accidents. Literate voters might be less myopic in their voting behavior whereby accidents in general are likely to be penalized irrespective of when they occur. Districts with high literacy rates should therefore experience smaller cycles. This is exactly what is observed in Table C.3. In columns (3) and (6), I include interactions between the election dummy and the proportion of population in each district that is literate. These

[94] The mean of ST population percentages across districts in the Red Corridor is 23%, compared to only 14% in the other districts in my sample.

[95] The coefficient on the interaction term must be multiplied by the proportion of Scheduled Tribes in a district to obtain the overall marginal effect.

[96] For other outcomes such as deaths and fatal accidents it is significant (results not reported).

data were also obtained from the 2011 Census of India. The coefficients on the election dummies alone are large, negative and statistically significant, whereas those on the interaction terms are positive and significant implying smaller cycles in districts with higher literacy rates.

In summary, the results from Table C.3 provide strong support for my main findings and in addition have implications for policy aimed at diffusing opportunistic cycles in mining fatalities, primarily through means that increase political accountability.

Given the claim that the rebels bank on local grievances against mining activity in qualitative studies (Kujur, 2009), it is possible that mine accidents, especially deaths, aggravate grievances that rebels utilize to propagate their agenda. An important way of doing this would be through increased recruitment to the group. A larger recruitment base is thus likely to help facilitate this process. In Table C.4 of the appendix, I first show mining fatalities in districts to be positively correlated with conflict onset and intensity. Thereafter in columns (2), (5) and (6), I interact mining fatalities with the proportion of ST population in districts, documented to be a strong predictor of Maoist violence primarily though increased opportunity of recruitment to groups (Gomes, 2015; Hoelscher et al., 2012). I find this correlation to be stronger in districts that have a higher share of ST population. The coefficients on the triple interaction term between fatalities, ST proportion and the election dummies are negative and significant implying a weakening of this correlation in the election year, which concurs with my main results. These coefficients should not be interpreted as causal estimates. However, they do provide correlational evidence of the fact that political costs of mining accidents are on average greater in districts that have higher conflict propensity due to their demographic composition and characteristics.

Overall through the range of robustness checks performed, I find strong evidence in support of the two primary channels I hypothesize are at play in generating mining cycles. Additionally, the analysis points to possible interaction of these two mechanisms, which would be interesting to address in future work.

4.7.2 Discussion

I note two potential limitations to the analysis in this paper. In the district level analysis, I only use mining fatalities data as proxy for mining intensity, primarily because it is the objective of reducing fatalities, as I show, that is important for politicians during elections. Fatalities data are obtained from government sources and one can potentially argue that there is deliberate underreporting of accidents closer to elections, which is driving my results. However, this is unlikely for two reasons. The data for each year are not released until after the end of the year. Thus, to the extent that elections are the basis of mining cycles, there is no incentive to simply underreport fatalities, unless they actually reduce. It is difficult to suppress the spread of information on deaths and accidents locally at the constituency level, especially since opposition party candidates are likely to use them as campaigning tools. Therefore, underreporting is unlikely to be the first order consideration for incumbent politicians. Note that the results on the effect of electoral competition on the size of mining cycles (Table 4.6) in such a situation would imply greater underreporting in a consistent manner in electorally competitive districts relative to non-competitive ones, but only as elections approach. This is also unlikely. Furthermore, the effects are observed for output and mineral licensing as well, when aggregated at the state-level. The government might actually be tempted to overreport activity during elections, at least when it comes to these two measures. The fact that the same pattern is observed across all three measures provides strong evidence against a situation where the results would be generated simply by systematic underreporting during elections.

The effect of electoral competition on the size of mining cycles is analyzed using election outcomes over at least two elections for every district in the sample. This could potentially lead to an endogeneity problem, if there is a spurious relationship between mine accidents and election results. One way around this problem would be to restrict the analysis to only one election cycle per district. However, with a little over 100 districts, this would mean significant loss of power leading to less precise estimates. Furthermore, since elections are not synchronized, it would also mean using different panel sizes for districts in each state, which is not ideal. It is possible to use the same panel size and at the same time consider only one election outcome per district, using variation from districts within states (using state fixed effects). I have tried this. However, this leads to a sample that is not representative, since it entails using a time period over which a large proportion of states in the sample did not have elections. Nevertheless, it might be possible to use past election results to observe behavior in the years immediately after the election, and predicted margin of victory in the second half of the political term. This is dependent on the availability of consistent and reliable data on opinion polls.

4.8 Conclusion

Focusing on India, this chapter documents the presence of opportunistic political cycles in several aspects of mining activity, different from the ones traditionally observed in the developed or the developing world. While political cycles in economic activity generally exhibit a U-shape pattern over electoral cycles, with greater activity during elections, I show that mining in India follows a counter cyclical pattern with mineral extraction minimized in election years.

The analysis in this chapter makes two factors driving these patterns apparent. Mining fatalities are politically unfavourable and electorally competitive districts undertake additional measures to minimize accidents compared to non-competitive ones during elections. This could stem from reduced output and/or better safety practices, especially because studies have found fatality records in Indian mines to fare poorly even when scaled for output compared to records in the U.S.A or developing countries such as South Africa (Mandal and Sengupta, 2000). I do not address the mechanisms directly in this paper, but it remains an important area of future work. However, the fact that accidents are reduced significantly in competitive districts (relative to non-competitive ones) only in election years, is evidence that mining fatalities are electorally sensitive, but also that voters are myopic and politicians therefore address this only during polls. Additionally, the fact that cycles are smaller in districts with higher literacy and larger in districts with greater Scheduled Tribe (ST) population showcase strategic behavior on the part of authorities responsible for mining operations.

The second factor driving mining cycles, also electorally tied, is the propensity of Maoist conflict. Based on anecdotal evidence and previous works that show extortion of mining revenues constitute an important revenue base for rebels, I hypothesize that minimizing this source of funding is important for politicians during elections, since rebel groups target the electoral process to spread their agenda. Consistent with this hypothesis, I find mining cycles to be larger in the Red Corridor, with greater reductions in activity in conflict prone districts during elections. This in turn results in smaller cycles of violence in Maoist affected mining districts relative to non-mining ones.

My results have important implications for addressing political cycles in mining fatalities and also for conflict resolution strategies. Fatalities in mining fields need to be addressed systematically, and a larger share of the population with basic education, by being less myopic in their voting behavior can help reduce the magnitude of these cycles. Furthermore, given that a large proportion of mine workers belong to the Scheduled Tribes (ST), political reservation for ST candidates in mineral rich areas could result in greater attention to the issue. Related to this, if rebels do bank on grievances generated as a result of mining deaths, reservation could also lead to reduced violence intensity in these areas. This is another important area I plan to focus on in future work.

Election cycles in violence, as documented in Table 8 clearly point to the time varying nature of the value of resources to rebel groups. This has implications for deployment of security personnel to protect legitimate mining activity. This is also true with respect to rebels' tax base is non-mining districts; however further research is required into first identifying and thereafter understanding policy that would be effective in achieving this. Furthermore, while policies such as subsidised rainfall insurance and guaranteed employment schemes have been successful in reducing rebel recruitment in the past, the findings of this paper suggest that their effectiveness is likely to vary over the electoral cycle. This is not to suggest intensifying such schemes during elections, but solely from the objective of minimizing violence, greater government responsiveness to income shocks at times when rebels value recruitment more, is important. Future work could focus on building a theoretical framework to formalize the different channels generating mining cycles. More importantly, a model will allow counterfactual policy analysis which is key. Will greater political reservation for backward classes reduce violence and conflict in mineral rich areas ? Will it affect the size of mining cycles? Can changes to the regulatory structure of the mining sector reduce political cycles in accidents? These are questions of vital importance, answers to which can help fix large-scale market failures in India's mining industry – unlocking its true potential could add USD 250 billion to India's GDP creating 13-15 million jobs through both direct and indirect contribution by 2025.[97]

[97] Strategy Plan for Ministry of Mines 2015 (https://mines.gov.in/writereaddata/UploadFile/Strategy.pdf)

4.9 Tables and figures

4.9.1 Tables

Table 4.1: Summary statistics of key variables

Variable	Mean	Standard Deviation	N
Mineral Output/Accidents/Leases			
Log Mining Output (1960-2000)	8.05	2.31	592
Mining's Share of State Output	.032	.028	592
Fatal Mining Accidents (1998-2015)	16.71	33.89	350
Deaths from Fatal Accidents	16.67	33.79	350
Deaths from Fatal Accidents (Accidents >0)	28.50	40.29	350
Log Number of Mining Leases (1995-2004)	5.01	1.78	221
Political Variables (for Output)			
Election Year	0.22	0.42	592
Scheduled Election Year	0.23	0.42	592
Four Years from Scheduled Election	0.22	0.42	592
Three Years from Scheduled Election	0.20	0.41	592
Two Years from Scheduled Election	0.16	0.34	592
One Year from Scheduled Election	0.19	0.38	592
Political Variables (for Accidents)			
Election Year	0.22	0.41	350
Scheduled Election Year	0.20	0.40	350
Four Years from Scheduled Election	0.19	0.40	350
Three Years from Scheduled Election	0.20	0.40	350
Two Years from Scheduled Election	0.18	0.38	350
One Year from Scheduled Election	0.23	0.42	350
Political Variables (for Leases)			
Election Year	0.21	0.41	221
Scheduled Election Year	0.20	0.39	221
Four Years from Scheduled Election	0.20	0.40	221
Three Years from Scheduled Election	0.20	0.41	221
Two Years from Scheduled Election	0.19	0.39	221
One Year from Scheduled Election	0.20	0.40	221

(a) Panel A: State-year level

Variable	Mean	Standard Deviation	N
Mining Fatalities (2010 - 2015)			
Total Accidents	6.10	16.96	605
Fatal Accidents	1.08	1.84	605
Serious Accidents	5.02	16.08	605
Total Serious Injuries	5.15	16.67	605
Total Deaths	1.36	2.36	605
Naxalite Conflict (2005 - 2017)			
Security Forces Deaths from Naxalite Conflict	1.10	5.94	1417
Rebel Deaths from Naxalite Conflict	1.01	3.63	1417
Political Variables (for Fatalities)			
Scheduled Election Year/ Election Year	0.17	0.37	605
Four Years from Scheduled Election	0.22	0.42	605
Three Years from Scheduled Election	0.23	0.42	605
Two Years from Scheduled Election	0.18	0.38	605
One Year from Scheduled Election	0.19	0.39	605

(b) Panel B: District-year level

Notes: The unit of observation in Panel A is a state-year. The output values are normalized with respect to 1973 prices. Mining fatalities and lease data are obtained from 23 states (including the 16 for output) between periods 1998-2015 and 1995-2004 respectively. Political variables are dummies for each year of the political cycle. The unit of observation in Panel B is a district-year. All values reported are averages across districts but over different time periods as mentioned in the table.

Table 4.2: Mining lease distribution and years to scheduled election

	OLS		IV		Scheduled Cycle	
	(1)	(2)	(3)	(4)	(5)	(6)
Four Years	0.0439	0.022	-0.038	-0.130	0.000	-0.047
	(0.0447)	(0.0370)	(0.101)	(0.0967)	(0.0424)	(0.035)
	[0.359]	[0.484]	[0.716]	[0.151]	[0.924]	[0.124]
Three Years	-0.0696	0.007	0.000	-0.065	0.005	-0.039
	(0.0476)	(0.0287)	(0.0935)	(0.0420)	(0.0563)	(0.029)
	[0.157]	[0.790]	[0.997]	[0.189]	[0.925]	[0.137]
Two Years	-0.0631	0.0186	-0.029	-0.093	-0.014	-0.060
	(0.0375)	(0.0338)	(0.113)	(0.0814)	(0.0648)	(0.0553)
	[0.167]	[0.507]	[0.815]	[0.309]	[0.802]	[0.215]
One Year	-0.117***	-0.057**	-0.070	-0.178***	-0.042	-0.125**
	(0.041)	(0.028)	(0.118)	(0.091)	(0.055)	(0.043)
	[0.004]	[0.023]	[0.594]	[0.007]	[0.564]	[0.012]
Outcome Mean	5.01	5.01	5.01	5.01	5.01	5.01
	(1.78)	(1.78)	(1.78)	(1.78)	(1.78)	(1.78)
State Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes
State Time Trends	No	Yes	No	Yes	No	Yes
N	221	221	221	221	221	221
R^2	0.986	0.992	0.986	0.991	0.986	0.992

Notes : * p<0.10, ** p<0.05, *** p<0.010 with respect to wild cluster bootstrap p-values. State-Year Level observations between 1995-2004. Dependent variable is log number of active leases in a given state in a year. Standard Errors clustered at the state level in parenthesis. P-values from Wild cluster Bootstrap (Davidson and MacKinnon, 2010 for IV specifications) with 1000 replications presented in square [] brackets. Scheduled cycles used as IV for actual cycles.

Table 4.3: Mining output and years to scheduled election

	Log Per Capita		**Log Output**	
	(1)	(2)	(3)	(4)
Four Years	0.114**	0.0613*	0.112**	0.0569
	(0.0414)	(0.0343)	(0.0446)	(0.0363)
	[0.0210]	[0.0870]	[0.0400]	[0.1210]
Three years	0.170**	0.100**	0.170**	0.0952*
	(0.0609)	(0.0455)	(0.0647)	(0.0482)
	[0.0260]	[0.0450]	[0.0280]	[0.0680]
Two Years	0.150*	0.0776	0.146	0.0663
	(0.0736)	(0.0560)	(0.0786)	(0.0629)
	[0.0520]	[0.2050]	[0.1200]	[0.3250]
One Year	0.114	0.0427	0.116	0.0393
	(0.0658)	(0.0400)	(0.0578)	(0.0414)
	[0.1510]	[0.294]	[0.1510]	[0.3625]
Outcome Mean	-2.31	-2.31	8.05	8.05
	(1.87)	(1.87)	(2.30)	(2.30)
State Effects	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes
State Time Trends	No	Yes	No	Yes
N	590	590	592	592
R^2	0.893	0.936	0.921	0.954

Notes : * p<0.10, ** p<0.05, *** p<0.010 with respect to wild cluster bootstrap p-values. State-Year level observations covering 16 major Indian states between 1960-2000. Standard errors clustered at the state-level reported in parenthesis. Wild cluster bootstrap p-values with 1000 replications in square [] brackets. “Log Per capita” refers to log of per capita mining output. “Log Output" refers to log of total mining output. Output values normalized with respect to 1973 prices.

Table 4.4: Fatal mining accidents and years to scheduled election (Poisson)

	Accidents		**Injuries**		**Deaths**	
	(1)	(2)	(3)	(4)	(5)	(6)
Four Years	0.213	0.227	0.646	0.928	0.205	0.225
	(0.191)	(0.222)	(0.419)	(0.512)	(0.179)	(0.198)
	[0.296]	[0.391]	[0.240]	[0.178]	[0.325]	[0.338]
Three Years	0.285	0.275	0.469	0.704	0.277	0.270
	(0.388)	(0.420)	(0.460)	(0.551)	(0.374)	(0.400)
	[0.805]	[0.905]	[0.340]	[0.252]	[0.676]	[0.757]
Two Years	-0.0510	-0.154	0.0586	0.0822	-0.193	-0.297
	(0.190)	(0.176)	(0.259)	(0.336)	(0.246)	(0.234)
	[0.823]	[0.412]	[0.818]	[0.825]	[0.491]	[0.260]
One Year	-0.233*	-0.340**	-0.952**	-0.949*	-0.197*	-0.304**
	(0.122)	(0.129)	(0.374)	(0.393)	(0.120)	(0.131)
	[0.087]	[0.027]	[0.046]	[0.085]	[0.100]	[0.042]
Outcome Mean	16.71	16.71	3.02	3.02	16.67	16.67
	(33.89)	(33.89)	(7.05)	(7.05)	(33.79)	(33.79)
State Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes
State Time Trends	No	Yes	No	Yes	No	Yes
N	350	350	350	350	350	350
Log-Likelihood	-1186	-1018	-385	-306	-1382	-1206

Notes: * p<0.10, ** p<0.05, *** p<0.010 with respect to Kline and Santos (2012) score bootstrap p-values. State-Year Level observations between 1998-2015. Only fatal accidents (accidents with at least one death) feature in this sample. Dependent variables are the number of events in a year in a given state. Standard errors clustered at the state level reported in parenthesis. P-values from Wild cluster Bootstrap with 1000 replications presented in square [] brackets. The sample used for this analysis includes data on 21 Indian states.

Table 4.5: Mining fatalities and years to scheduled election (district-level)

	Any Bad Event	**Accidents**		**Casualties**	
	(1)	(2)	(3)	(4)	(5)
Four Years	0.236***	0.214***	0.475***	0.234***	0.506***
	(0.0758)	(0.0511)	(0.138)	(0.0600)	(0.153)
Three Years	0.179***	0.207***	0.402***	0.203***	0.384**
	(0.0656)	(0.0547)	(0.131)	(0.0693)	(0.157)
Two Years	0.222***	0.199***	0.508***	0.211***	0.532***
	(0.0616)	(0.0530)	(0.141)	(0.0691)	(0.167)
One Year	0.0930	0.0385	0.179	0.0478	0.196
	(0.0608)	(0.0651)	(0.160)	(0.0717)	(0.171)
Outcome Mean	0.64	6.10	6.10	6.50	6.50
	(0.49)	(16.96)	(16.96)	(17.61)	(17.61)
District Effects	Yes	Yes	Yes	Yes	Yes
Region-Year Effects	Yes	Yes	Yes	Yes	Yes
N	605	605	605	605	605
$R^2/Log-Likelihood$	0.535	0.880	-1063	0.853	-1124
Estimation	OLS	OLS	Poisson	OLS	Poisson

Notes : * p<0.10, ** p<0.05, *** p<0.010. District-Year level observations covering 104 districts across 15 states between 2010-2015. Standard errors clustered at the state-year level reported in parenthesis. Standard deviations for outcome means are reported in parenthesis. "Any Bad Event" is a dummy variable coded 1 if at least one person was either seriously injured or died in a mining field. "Accidents" refer to the sum of both fatal and serious accident and "Casualties" refer to the total number of peopled killed and injured in mining fields in a district in a given year. Outcomes Casualties and Accidents for OLS regressions are subject to a log $(x+1)$ transformation.

Table 4.6: Electoral cycle, political competition and mining fatalities

	Log Accidents			Log Casualties		
	(1)	(2)	(3)	(4)	(5)	(6)
Four Years	0.237***	0.237***	-0.0224	0.268***	0.269***	-0.065
	(0.0594)	(0.0603)	(0.115)	(0.0680)	(0.0690)	(0.124)
Three Years	0.339***	0.338***	0.178*	0.357***	0.356***	0.147
	(0.0750)	(0.0764)	(0.105)	(0.0891)	(0.0895)	(0.121)
Two Years	0.244***	0.243***	0.169*	0.271***	0.270***	0.165
	(0.0687)	(0.0701)	(0.102)	(0.0860)	(0.0877)	(0.129)
One Year	0.0477	0.0467	-0.0652	0.0667	0.0656	-0.0434
	(0.0749)	(0.0762)	(0.0926)	(0.0832)	(0.0841)	(0.104)
Close Election Proportion		0.020	-0.286**		0.021	-0.366**
		(0.0959)	(0.139)		(0.104)	(0.148)
Four Years × Close Election Proportion			0.556***			0.717***
			(0.213)			(0.214)
Three Years × Close Election Proportion			0.349**			0.451**
			(0.178)			(0.189)
Two Years × Close Election Proportion			0.165			0.240
			(0.187)			(0.205)
One Year × Close Election Proportion			0.266			0.270
			(0.172)			(0.182)
Outcome Mean	6.36	6.36	6.36	6.89	6.89	6.89
	(17.59)	(17.59)	(17.59)	(18.27)	(18.27)	(18.27)
District Effects	Yes	Yes	Yes	Yes	Yes	Yes
Region-Year Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	558	558	558	558	558	558
R^2	0.849	0.849	0.885	0.817	0.817	0.861

* p<0.10, ** p<0.05, *** p<0.010. District-Year Level observations between 2010-2015 covering 97 districts. Standard errors clustered at state-year level in parenthesis. The outcome means reported are the actual means and not log values. Standard deviations for outcome means are reported in parenthesis. Log Accidents refer to the natural log of the total number of accidents (both serious and fatal) in a district in a particular year. Log Casualties refer to the natural logarithm of total deaths and injuries in mining fields in a given district in a year. Outcome variables are subject to a log $(x+1)$ transformation.

Table 4.7: Electoral cycle, Red Corridor and mining fatalities

	Log Accidents			Log Casualties		
	(1)	(2)	(3)	(4)	(5)	(6)
Four Years	0.214***	0.188***	0.183*	0.234***	0.223***	0.218**
	(0.0521)	(0.0556)	(0.0954)	(0.0610)	(0.0637)	(0.101)
Three Years	0.207***	0.151**	0.143	0.203***	0.147**	0.139
	(0.0549)	(0.0608)	(0.0939)	(0.0695)	(0.0718)	(0.103)
Two Years	0.199***	0.122**	0.106	0.211***	0.135*	0.122
	(0.0530)	(0.0594)	(0.0929)	(0.0691)	(0.0742)	(0.106)
One Year	0.0385	-0.0101	-0.00388	0.0478	-0.0120	-0.00218
	(0.0653)	(0.0677)	(0.102)	(0.0720)	(0.0725)	(0.106)
Red Corridor			-0.383*			-0.389*
			(0.231)			(0.235)
Four Years × Red Corridor		0.0972	0.116		0.0401	0.0590
		(0.0969)	(0.249)		(0.109)	(0.252)
Three Years × Red Corridor		0.191**	0.242		0.185**	0.237
		(0.0884)	(0.243)		(0.0911)	(0.246)
Two Years × Red Corridor		0.243***	0.252		0.230**	0.238
		(0.0887)	(0.255)		(0.0975)	(0.264)
One Year × Red Corridor		0.171	0.142		0.189	0.149
		(0.122)	(0.296)		(0.129)	(0.306)
Outcome Mean	6.10	6.10	6.10	6.50	6.50	6.50
	(16.96)	(16.96)	(16.96)	(17.61)	(17.61)	(17.61)
State Effects	No	Yes	No	No	Yes	No
District Effects	Yes	No	Yes	Yes	No	Yes
Region-Year Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	605	605	605	605	605	605
R^2	0.844	0.196	0.844	0.811	0.192	0.811

* p<0.10, ** p<0.05, *** p<0.010. District-Year Level observations covering 104 districts between 2010-2015. Standard errors clustered at state-year level in parenthesis. The outcome variables are the number of accidents and the sum of deaths and injuries in mining fields in a district-year (logged) respectively. The actual mean values (not logged) are reported for the outcome variables. “Red Corridor" is a dummy variable coded 1 if a district had been classified by the Union Ministry of Home Affairs to be part of the areas affected by the Naxalite conflict in 2008, which is prior to the beginning of the fatalities sample. Figure 4.1(a) marks all such districts in red and those that are not part of the “Red Corridor" in green. Outcome variables are subject to a log $(x+1)$ transformation.

Table 4.8: Election cycles and Naxalite conflict (Poisson regressions)

	Total Deaths			Security Forces Deaths			Rebel Deaths		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mineral Dummy	-0.0462			0.0335			-0.212		
	(0.508)			(0.594)			(0.416)		
Four Years		-0.322	-0.244		-1.039***	-0.941**		0.220	0.203
		(0.238)	(0.329)		(0.293)	(0.431)		(0.282)	(0.305)
Three years		0.174	0.0290		0.143	-0.129		0.311	0.290
		(0.224)	(0.320)		(0.318)	(0.427)		(0.339)	(0.360)
Two Years		0.234	0.0248		0.318	0.0335		0.123	-0.0873
		(0.370)	(0.485)		(0.423)	(0.584)		(0.507)	(0.579)
One Year		0.812**	1.106***		0.673*	1.002**		1.012**	1.193***
		(0.320)	(0.363)		(0.368)	(0.445)		(0.434)	(0.423)
Four Years × Mineral Dummy			-0.128			-0.194			0.0663
			(0.468)			(0.646)			(0.446)
Three Years × Mineral Dummy			0.122			0.269			-0.00811
			(0.492)			(0.576)			(0.432)
Two Years × Mineral Dummy			0.208			0.330			0.282
			(0.431)			(0.554)			(0.418)
One Year × Mineral Dummy			-0.969**			-1.183*			-0.590
			(0.496)			(0.707)			(0.389)
p-value ($One\ Year$ + $One\ Year \times Mineral\ Dummy = 0$)			0.715			0.728			0.178
Outcome Mean	2.16	2.16	2.16	1.10	1.10	1.10	1.01	1.01	1.01
	(9.09)	(9.09)	(9.09)	(5.94)	(5.94)	(5.94)	(3.63)	(3.63)	(3.63)
State Effects	Yes			Yes			Yes		
District Effects	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Region-Year Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1417	1417	1417	1417	1417	1417	1417	1417	1417
Log-Likelihood	-4851	-2457	-2388	-2901	-1443	-1381	-2480	-1370	-1358

* $p<0.10$, ** $p<0.05$, *** $p<0.010$. District-Year Level observations covering 2005-2017. Standard errors clustered at state-year level in parenthesis except for regressions in Columns (1), (4) and (7) where I cluster standard errors at the district level. Standard deviations for outcome means are reported in parenthesis. The outcome variable in columns (1)-(3) is the total number of deaths from clashes between security forces and naxal rebels. These also include civilians who died in these clashes. In columns (4)-(6) and (7)-(9), the outcome variables are number of rebel and security forces deaths respectively. There are 109 districts in these regressions.

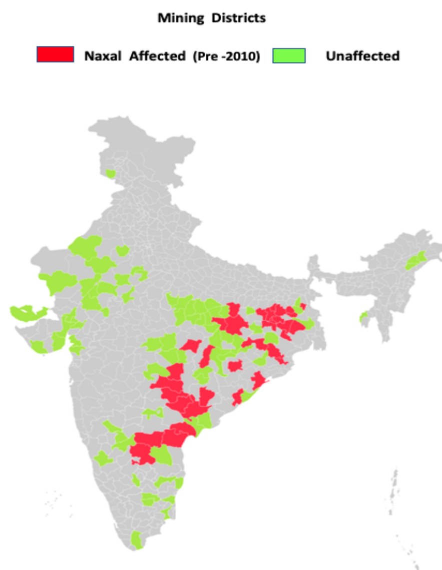
Table 4.9: Elections and coal mine explosives

	Log Detonators		**Log Total Explosives**		**Log Non-Permitted**	
	(1)	(2)	(3)	(4)	(5)	(6)
$Election$	-0.0148	0.0170	0.0661	0.156	0.0867	0.155
	(0.126)	(0.166)	(0.0807)	(0.121)	(0.0961)	(0.127)
$Election_{-1}$	-0.0226	0.176	0.0625	0.140	0.139	0.259**
	(0.136)	(0.126)	(0.0923)	(0.114)	(0.0899)	(0.0989)
Election × Red Corridor		-0.263*		-0.273**		-0.287**
		(0.147)		(0.128)		(0.129)
$Election_{-1}$ × Red Corridor		-0.583***		-0.158		-0.310**
		(0.161)		(0.117)		(0.121)
Outcome Mean	12.85	12.85	15.27	15.27	15.21	15.21
	(2.51)	(2.51)	(1.78)	(1.78)	(1.85)	(1.85)
District Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	220	220	221	221	207	207
Adj. R^2	0.928	0.929	0.923	0.924	0.919	0.920

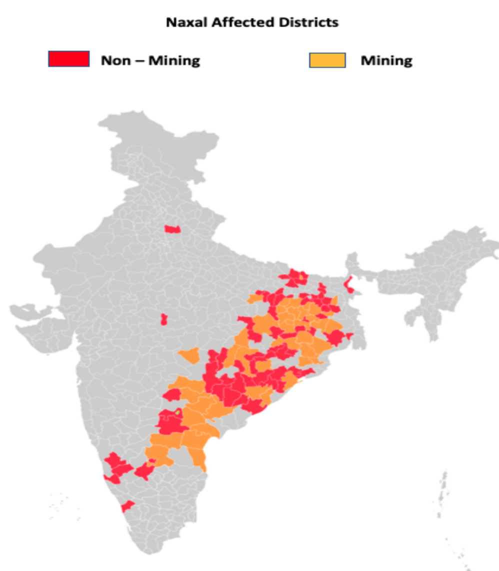
* p<0.10, ** p<0.05, *** p<0.010. District-Year level observations for covering 38 major coal mining districts in India between 2010-2015. Standard errors clustered at the state-year level reported in parenthesis. Standard deviations for outcome means are reported in parenthesis. Each district reports total kilograms of explosives and detonators used in a year. The mean of the logged values have been reported for each outcome variable. "Election_{-1}" is a dummy variable denoting whether the state which a district belongs to, is scheduled to have a election in the following year. "Red Corridor" is a dummy variable taking the value 1 if a district had been classified by the Union Ministry of Home Affairs as affected by Naxal violence in 2008. Non-permitted explosives are less environmentally friendly and "should not be used in underground coal mines where there is any possible risk of igniting combustible gases or coal dust" (U.S. Bureau of Mines).

4.9.2 Figures

Figure 4.1: Sample for district-level study



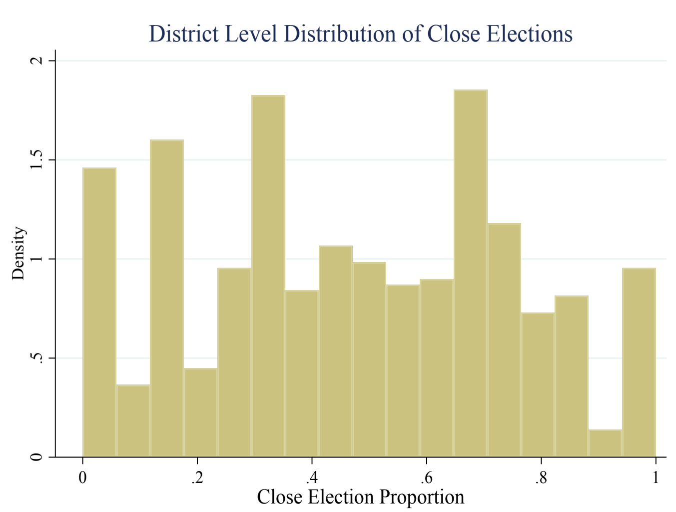
(a) Mining Districts



(b) Naxal Affected Districts (2005-2017)

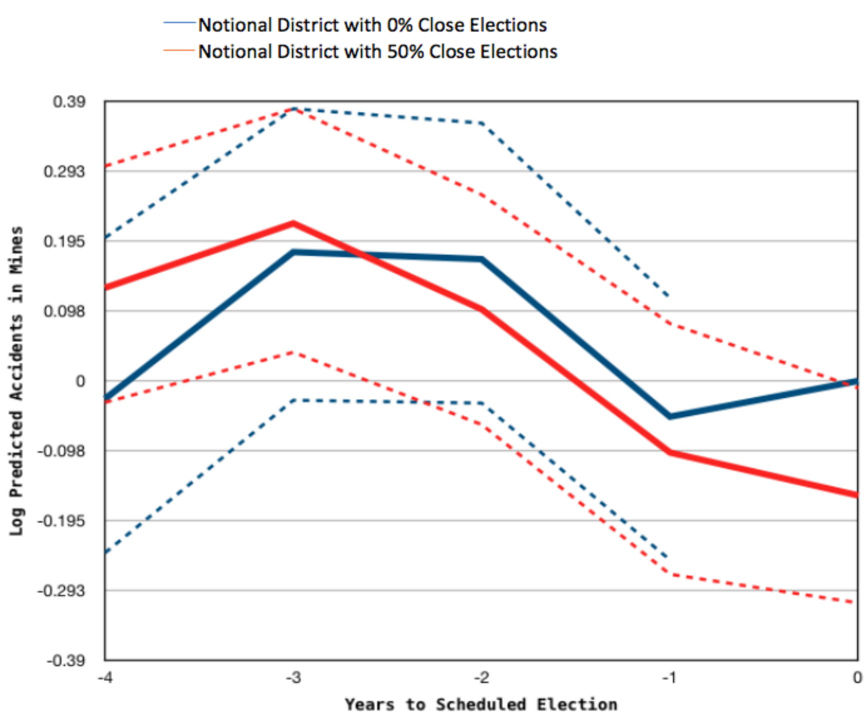
Note: This figure marks all the districts that form the entire sample for the district level study. Sub-figure (a) includes all the mining districts in my sample. The Naxal affected districts in 2008 are marked in red in sub-figure (a). Sub-figure (b) includes all districts where at least one death has been reported between 2005-2017 as a result of clashes between Naxalite rebels and government security forces. The mining districts are marked in orange.

Figure 4.2: Distribution of close elections



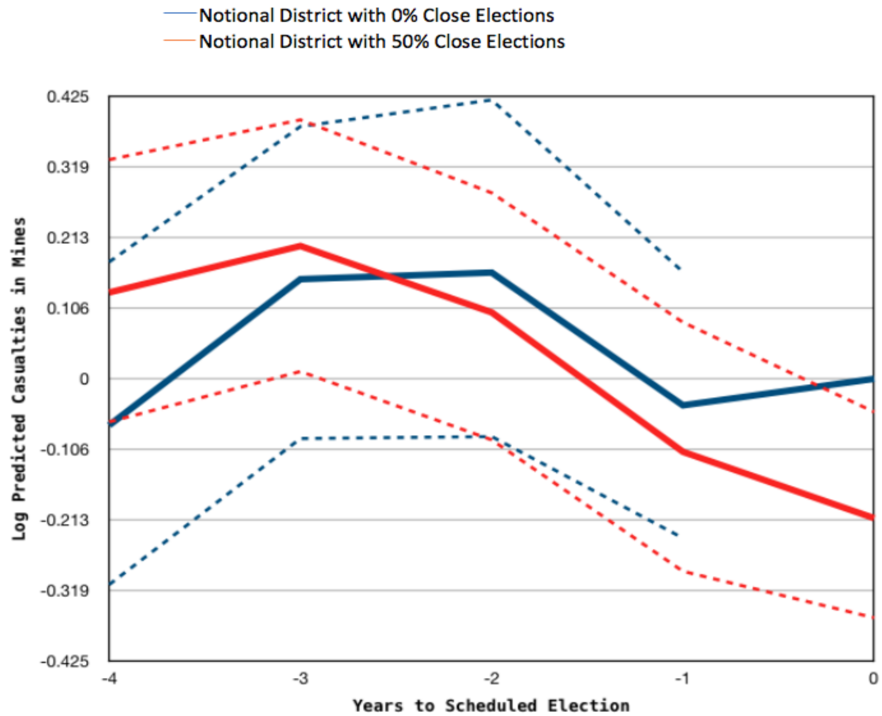
Note: This figure shows the distribution of close elections across all the districts in my sample. An election is considered to be close if the win margin in a constituency within a district was less than 10 % of total polled votes.

Figure 4.3: Accident cycles and electoral competition



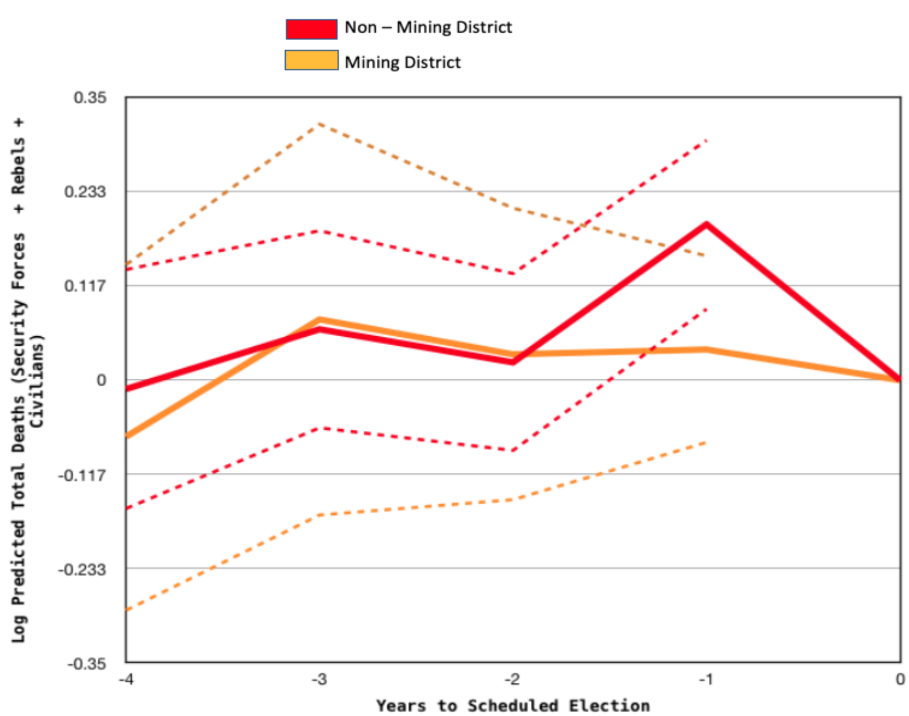
Note: This diagram plots log predicted accidents in mines over the next electoral cycle for notional districts with 0% and 50% close elections in the previous elections respectively. It is based on Column (3) of Table 4.5. Note that the value in election year is 0 by construction for a district with no close elections. Dotted lines represent 95% confidence intervals.

Figure 4.4: Fatality cycles and electoral competition



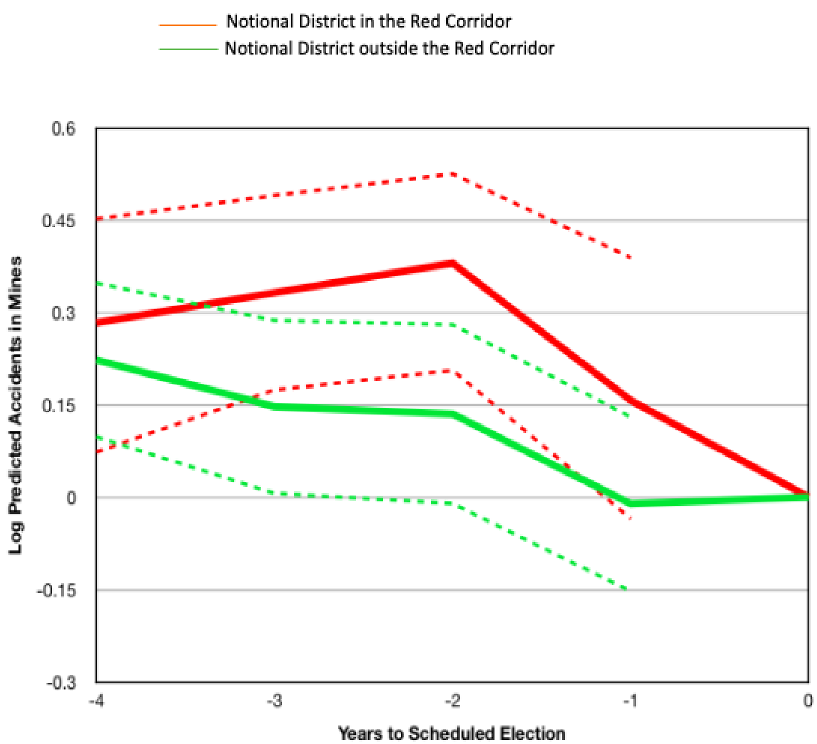
Note: This diagram plots log predicted casualties in mines over the next electoral cycle for notional districts with 0% and 50% close elections in the previous elections respectively. It is based on Column (6) of Table 4.5. Note that the value in election year is 0 by construction for a district with no close elections. Dotted lines represent 95% confidence intervals.

Figure 4.6: Total conflict deaths over the electoral cycle



Note: This figure plots log predicted total deaths from clashes between naxal rebels and security forces over the electoral cycle in mining and non-mining districts. The value in the election year is normalized to 0 in each case. It is based on specifications of the form in Table 8 Columns (3), (6) and (9) but from a OLS model with logged outcome variable. Mining districts are those that feature in my sample of mining fatalities. Dotted lines show 95% confidence intervals.

Figure 4.5: Accident cycles and Red Corridor



Note: This figure plots log predicted accidents in mining fields over the electoral cycle for a district in the red corridor (red line) and a district that is not in the red corridor (blue line) based on equation (6) with district fixed effects. The value in the election year is normalized to 0 in each case. I use pre-sample classification of districts into the Red Corridor (fixed district characteristic). Hence, equation (6) only includes interactions between RC and the electoral cycle dummies. This figure is based on results from Column (3) of Table 4.7. Dotted lines show 95% confidence intervals.

Chapter 5

Conclusion

This thesis consists of three distinct chapters at the frontier of research in development economics. Chapter 2 shows that the effects of religious diversity on team productivity and worker attitudes in a firm depends on the type of production technology in use. I partnered with a large processed food manufacturing plant in West Bengal, India and randomly allocated its Hindu and Muslim workers to be in either religiously mixed (Hindus and Muslims) or homogeneous (Hindus only) teams. There are multiple production tasks at the factory. I classify these tasks into high- (HD) and low-Dependency (LD) types with time-use data on the degree of continuous coordination required amongst teammates to ensure uninterrupted production and the dependence on teammates for breaks.

I find that while religious mixing leads to a loss in output in high-dependency tasks, it has no effect in low-dependency work. Consistent with this, worker surveys reveal more frictions amongst workers from mixing in high-dependency work: there are greater accusations, more inter-group blame and lower team cohesion in general relative to homogeneous teams, as well as relative to mixed low-dependency teams. However, the negative effects on output in high-dependency tasks attenuate over time and dissipate completely in four months. The improvements in production are accompanied by improved inter-group relations as well. In low-dependency tasks, while there is no negative output shock from religious mixing, there are little or no positive effects on inter-group relations. Overall, this pattern of results suggests that technology that incentivizes individuals to learn to work together is important in overcoming existing intergroup differences – and leads to improved relations and team performance. However, more speculatively, the tension between the goal of maximizing short-run productivity and that of improving intergroup relations might explain why (in equilibrium) we could see a lot of integration (at work) without intergroup relations improving — the integration might only occur in contexts where it is socially ineffective.

I hope to now conduct a large-scale survey with firm owners and workers across India to understand how firms with different production technologies deal with (the effects of) diversity. Do high-dependency firms segregate workers based on caste and religion? Do low-dependency firms integrate workers but without inter-group relations improving? Which exact aspects of high-dependency tasks make diversity costly? Can some of these interactions be substituted with low-dependency work? A deeper analysis of the role that technology plays in determining the effects of diversity can help us understand how the *costs* of diversity might change as economies undergo structural transformation. My next steps will be aimed at understanding how the interaction between production technology and ethnic diversity matters for broader economic change.

Chapter 3 uses 19th and 20th century state-level bans on cousin marriage in the US to provide causal micro-evidence of the impact of consanguineous marriages on a range of socio-economic outcomes. We borrow a method from population genetics and show that excess rates of same-surname marriages provide credible estimates of cousin marriage rates by surname, by state as well as over time. Our main results show that bans on first-cousin marriage led to a reduction in the rate of such marriages, led to higher incomes, more schooling, greater female labor force supply and rural-urban migration. We argue that these effects are driven by the weakening of family ties rather than a genetic channel.

Our findings are consistent with recent work in anthropology and sociology that studies the characteristics of kin-based societies. A large body of ethnographic and historical research shows that intensive kinship is associated with greater cooperation within a group, but it comes at the cost of geographic and social mobility, as well as participation in mainstream institutions (Henrich, 2020). Our results are consistent with the view that kinship norms evolve simultaneously with economic conditions, but they do so slowly over multiple generations. Finally, while clearly of historical significance, we feel our results might be relevant for contemporary development outcomes, since tight kinship is still prevalent in many societies. While the causal estimates of the of effect of kinship are not directly applicable to such societies, our results do suggest that as economies undergo structural transformation, leading to the development of better institutions, the returns to family structure transitions that weaken kinship ties could potentially be quite high.

While the first two chapters, broadly speaking, study the effects of social relations/integration on economic outcomes chapter, 4 focuses on political business cycles in the mining industry in India. I study several aspects of mining intensity (output, licensing and accidents) and find that in contrast to majority of other economic activities that track the electoral cycle, mining exhibits an inverted U-shaped pattern between state assembly elections, with the level of activity minimized in election years. I present evidence that the magnitude of these cycles are determined primarily by two factors: electoral competition and the intensity of Naxalite conflict, an ongoing left-wing insurgency against the Indian government. While mining accidents are costly during elections in general, I show that cycles in conflict prone areas are exacerbated in order to minimize the tax base of rebel groups, who thrive on extortion of mining revenues and target elections with violence. Overall, these results suggest that we might observe a decoupling of the usual relationship between economic activity and election timing for high-risk industrial activities or for those that increase the propensity of civil conflict.

Bibliography

A.Churchill, S. and M. Danquah (2020). Ethnic diversity and informal work in Ghana. Technical report, WIDER Working Paper.

Afridi, F., A. Dhillon, S. X. Li, and S. Sharma (2020). Using social connections and financial incentives to solve coordination failure: A quasi-field experiment in India's manufacturing sector. *Journal of Development Economics 144*, 102445.

Afridi, F., A. Dhillon, and S. Sharma (2020). The ties that bind us: Social networks and productivity in the factory. *CEPR Discussion Paper No. DP14687*.

Akbari, M., D. Bahrami-Rad, and E. O. Kimbrough (2019). Kinship, fractionalization and corruption. *Journal of Economic Behavior & Organization 166*, 493–528.

Alam, M. S. (2010). Social exclusion of Muslims in India and deficient debates about affirmative action: Suggestions for a new approach. *South Asia Research 30*(1), 43–65.

Alesina, A., Y. Algan, P. Cahuc, and P. Giuliano (2015). Family values and the regulation of labor. *Journal of the European Economic Association 13*(4), 599–630.

Alesina, A. and P. Giuliano (2010). The power of the family. *Journal of Economic growth 15*(2), 93–125.

Alesina, A. and P. Giuliano (2014). Family ties. In *Handbook of economic growth*, Volume 2, pp. 177–215. Elsevier.

Alesina, A. and E. Spolaore (1997). On the number and size of nations. *The Quarterly Journal of Economics 112*(4), 1027–1056.

Allport, G. W., K. Clark, and T. Pettigrew (1954). The nature of prejudice. *Addison-wesley Reading, MA*.

Altonji, J. G. and C. R. Pierret (1998). Employer learning and the signalling value of education. *Internal Labour markets, Incentives and Employment*, 159–195.

Altonji, J. G. and C. R. Pierret (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics 116*(1), 313–350.

Angelucci, M., G. De Giorgi, M. A. Rangel, and I. Rasul (2010). Family networks and school enrolment: Evidence from a randomized social experiment. *Journal of public Economics 94*(3-4), 197–221.

Artiles, M. (2020). Within-group heterogeneity in a multi-ethnic society. *Unpublished Working Paper*.

Asher, S. and P. Novosad (2018, April). Rent Seeking and Criminal Politicians : Evidence from Mining Booms. *Working Paper*.

Asher, S., P. Novosad, and C. Rafkin (2018). Intergenerational mobility in India: Estimates from new methods and administrative data. *World Bank Working Paper*.

Ashraf, N. and O. Bandiera (2018). Social incentives in organizations. *Annual Review of Economics 10*, 439–463.

Bandiera, O., I. Barankay, and I. Rasul (2010). Social incentives in the workplace. *The Review of Economic Studies 77*(2), 417–458.

Bandiera, O., I. Barankay, and I. Rasul (2013). Team incentives: Evidence from a firm level experiment. *Journal of the European Economic Association 11*(5), 1079–1114.

Banerjee, S. (1991). 'Hindutva': Ideology and social psychology. *Economic and Political Weekly*, 97–101.

Basant, R. (2007). Social, economic and educational conditions of Indian Muslims. *Economic and Political Weekly*, 828–832.

Baskaran, T., B. Min, and Y. Uppal (2015, 04). Election Cycles and Electricity Provision: Evidence from a Quasi-Experiment with Indian Special Elections. *Journal of Public Economics 126*.

Bates, R. H., A. Greif, and S. Singh (2004). The political economy of kinship societies. *Politics from Anarchy to Democracy: Rational Choice in Political Science*, 66.

Bau, N. (2021). Can policy change culture? government pension plans and traditional kinship practices. *American Economic Review Forthcoming*.

Bazzi, S., A. Gaduh, A. D. Rothenberg, M. Wong, et al. (2017). Unity in diversity?: Ethnicity, migration, and nation building in Indonesia. *Centre for Economic Policy Research*.

Becker, G. (1957). The economics of discrimination. *Chicago: University of Chicago Press*.

Bemiss, S. M. (1858). *Report on influence of marriages of consanguinity upon offspring*. Collins, printer.

Bennett, R. L., A. G. Motulsky, A. Bittles, L. Hudgins, S. Uhrich, D. L. Doyle, K. Silvey, C. R. Scott, E. Cheng, B. McGillivray, et al. (2002). Genetic counseling and screening of consanguineous couples and their offspring: recommendations of the national society of genetic counselors. *Journal of genetic counseling 11*(2), 97–119.

Berger, H. and U. Woitek (1997). Searching for Political Business Cycles in Germany. *Public Choice 91*(2), 179–97.

Besley, T. and R. Burgess (2004). Can Labor Regulation Hinder Economic Performance? Evidence from India. *The Quarterly Journal of Economics 119*(1), 91–134.

Bhalotra, S. R., I. Clots-Figueras, L. Iyer, and J. Vecci (2018). Leader identity and coordination. *IZA Discussion Paper*.

Bhattacharjee, S. (2014, Novermber). Timing of Elections and Infant Mortality : Evidence from India. *UBC Working Paper*.

Bhaumik, S. K. and M. Chakrabarty (2009). Is education the panacea for economic deprivation of Muslims?: Evidence from wage earners in India, 1987–2005. *Journal of Asian Economics 20*(2), 137–149.

Bittles, A. H. (2001). Consanguinity and its relevance to clinical genetics. *Clinical genetics 60*(2), 89–98.

Bittles, A. H. (2012). *Consanguinity in context*, Volume 63. Cambridge University Press.

Boisjoly, J., G. J. Duncan, M. Kremer, D. M. Levy, and J. Eccles (2006). Empathy or antipathy? The impact of diversity. *American Economic Review 96*(5), 1890–1905.

Bratt, C. S. (1984). Incest statutes and the fundamental right of marriage: is oedipus free to marry? *Family Law Quarterly*, 257–309.

Brown, J. S. (1951). Social class, intermarriage, and church membership in a kentucky community. *American Journal of Sociology 57*(3), 232–242.

Buonanno, P. and P. Vanin (2017). Social closure, surnames and crime. *Journal of Economic Behavior & Organization 137*, 160–175.

Burgess et al., R. (2011, September). The Political Economy of Deforestation in the Tropics. Working Paper 17417, National Bureau of Economic Research.

Byers, P. K. (1995). *African American genealogical sourcebook.* Gale/Cengage Learning.

Calafell, F. and M. H. Larmuseau (2017). The y chromosome as the most popular marker in genetic genealogy benefits interdisciplinary research. *Human genetics 136*(5), 559–573.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics 90*(3), 414–427.

Carpenter, J. and E. Seki (2011). Do social preferences increase productivity? Field experimental evidence from fishermen in Toyama Bay. *Economic Inquiry 49*(2), 612–630.

Cattaneo, M. D., R. K. Crump, M. H. Farrell, and Y. Feng (2019). On binscatter. *arXiv preprint arXiv:1902.09608*.

Census, . (2011). Census of India 2011 provisional population totals. *New Delhi: Office of the Registrar General and Census Commissioner*.

Chakravarti, S. (2014). Maoists, money and business. *The Mint*.

Chapparban, S. N. (2020). Religious identity and politics of citizenship in South Asia: A reflection on refugees and migrants in India. *Development 63*(1), 52–59.

Churchill, S. A., M. R. Valenzuela, and W. Sablah (2017). Ethnic diversity and firm performance: Evidence from China's materials and industrial sectors. *Empirical Economics 53*(4), 1711–1731.

Colantonio, S. E., G. W. Lasker, B. A. Kaplan, and V. Fuster (2003). Use of surname models in human population biology: a review of recent developments. *Human Biology*, 785–807.

Cole, S. (2009, January). Fixing Market Failures or Fixing Elections? Agricultural Credit in India. *American Economic Journal: Applied Economics 1*(1), 219–50.

Crow, J. F. and A. P. Mange (1965). Measurement of inbreeding from the frequency of marriages between persons of the same surname. *Eugenics Quarterly 12*(4), 199–203.

Cruz, C., J. Labonne, and P. Querubin (2017). Politician family networks and electoral outcomes: Evidence from the philippines. *American Economic Review 107*(10), 3006–37.

Darwin, G. H. (1875). Marriages between first cousins in england and their effects. *Journal of the Statistical Society of London 38*(2), 153–184.

Davidson, R. and J. G. MacKinnon (2010, March). Wild Bootstrap Tests for IV Regression. Working Papers 1135, Queen's University, Department of Economics.

Denic, S. and M. G. Nicholls (2007). Genetic benefits of consanguinity through selection of genotypes protective against malaria. *Human Biology 79*(2), 145–158.

Dennison, T. and S. Ogilvie (2014). Does the european marriage pattern explain economic growth? *The journal of economic history*, 651–693.

DeSante, C. D. (2013). Working twice as hard to get half as far: Race, work ethic, and America's deserving poor. *American Journal of Political Science 57*(2), 342–356.

Directorate General of Mines Safety, Ministry of Labour and Employment (2018). *Statistics of Mines (2010-2015)*.

Do, Q.-T., S. Iyer, and S. Joshi (2013). The economics of consanguineous marriages. *Review of Economics and Statistics 95*(3), 904–918.

Drazen, A. and M. Eslava (2010). Electoral Manipulation via Voter-Friendly Spending: Theory and Evidence. *Journal of Development Economics 92*(1), 39–52.

Easterly, W. and R. Levine (1997). Africa's growth tragedy: Policies and ethnic divisions. *The Quarterly Journal of Economics 112*(4), 1203–1250.

Edlund, L. (2018). Cousin marriage is not choice: Muslim marriage and underdevelopment. *AEA Papers and Proceedings 108*, 353–57.

Enke, B. (2019). Kinship, cooperation, and the evolution of moral systems. *The Quarterly Journal of Economics 134*(2), 953–1019.

Ermisch, J. and D. Gambetta (2010). Do strong family ties inhibit trust? *Journal of Economic Behavior & Organization 75*(3), 365–376.

Fafchamps, M. and J. Labonne (2017). Do politicians' relatives get better jobs? evidence from municipal elections. *The Journal of Law, Economics, and Organization 33*(2), 268–300.

Farber, B. (1968). *Comparative Kinship Systems: A Method of Analysis.* John Wiley and Sons, Inc.

Farber, B. (2000). Kinship systems and family types. In E. F. Borgatta (Ed.), *Encyclopedia of sociology*. New York: Macmillan Reference USA.

Farber, H. S. and R. Gibbons (1996, 11). Learning and Wage Dynamics. *The Quarterly Journal of Economics 111*(4), 1007–1047.

Fershtman, C. and U. Gneezy (2001). Discrimination in a segmented society: An experimental approach. *The Quarterly Journal of Economics 116*(1), 351–377.

Fischer, C. S. (1975). Toward a subcultural theory of urbanism. *American journal of Sociology 80*(6), 1319–1341.

Fukuyama, F. (2011). *The origins of political order: From prehuman times to the French Revolution.* Farrar, Straus and Giroux.

Ghatak, M. and O. V. Eynde (2017). Economic determinants of the maoist conflict in india. *Economic & Political Weekly 52*(39), 69–76.

Giuliano, P. and N. Nunn (2020). Understanding cultural persistence and change. *Review of Economic Studies*, Forthcoming.

Gomes, J. F. (2015). The Political Economy of the Maoist Conflict in India: An Empirical Analysis. *World Development 68*(C), 96–123.

Gonzalez, M. A. (2002, December). Do Changes in Democracy Affect the Political Budget Cycle ? Evidence from Mexico. *Review of Development Economics* (6(2)), 204–224.

Goody, J. (1983). *The development of the family and marriage in Europe.* Cambridge University Press.

Greif, A. (2006). Family structure, institutions, and growth: the origins and implications of western corporations. *American Economic Review 96*(2), 308–312.

Greif, A. and G. Tabellini (2017). The clan and the corporation: Sustaining cooperation in china and europe. *Journal of Comparative Economics 45*(1), 1–35.

Gymrek, M., A. L. McGuire, D. Golan, E. Halperin, and Y. Erlich (2013). Identifying personal genomes by surname inference. *Science 339*(6117), 321–324.

Hajnal, J. (1965). European marriage patterns in perspective. In D. V. Glass and D. E. C. Eversley (Eds.), *Population in History: Essays in Historical Demography*. E. Arnold London.

Hamilton, B. H., J. A. Nickerson, and H. Owan (2012). Diversity and productivity in production teams. In *Advances in the Economic Analysis of participatory and Labor-managed Firms*. Emerald Group Publishing Limited.

Henrich, J. (2020). *The WEIRDest People in the World: How the West Became Psychologically Peculiar and Particularly Prosperous.* Farrar, Straus and Giroux.

Hjort, J. (2014). Ethnic divisions and production in firms. *The Quarterly Journal of Economics 129*(4), 1899–1946.

Hoelscher, K., J. Miklian, and K. C. Vadlamannati (2012). Hearts and mines: A district-level analysis of the maoist conflict in india. *International Area Studies Review 15*(2), 141–160.

Hotte, R. and K. Marazyan (2020). Demand for insurance and within-kin-group marriages: Evidence from a west-african country. *Journal of Development Economics*, 102489.

Hsieh, C.-T. and B. A. Olken (2014). The “missing middle". *Journal of Economic Perspectives 28*(3), 89–108.

Indian Bureau of Mines (2018). *Statistics of Mineral Information, 2018*.

Jai, S. (2019). Political unrest, mine mishap bring Talcher coalfields to a halt. *The Business Standard*.

Jakiela, P., E. Miguel, and V. te Velde (2011). Combining field and lab experiments to estimate the impact of human capital on social preferences. *Mimeo, UC Berkeley*.

Jha, S. (2013). Trade, institutions, and ethnic tolerance: Evidence from South Asia. *American Political Science Review*, 806–832.

Jorde, L. B. (1989). Inbreeding in the utah mormons: an evaluation of estimates based on pedigrees, isonymy, and migration matrices. *Annals of human genetics 53*(4), 339–355.

Kalpagam, U. et al. (2010). Are Muslims discriminated against in the labour market in India? *Indian Journal of Labour Economics 53*(1).

Kaplanis, J., A. Gordon, T. Shor, O. Weissbrod, D. Geiger, M. Wahl, M. Gershovits, B. Markus, M. Sheikh, M. Gymrek, et al. (2018). Quantitative analysis of population-scale family trees with millions of relatives. *Science 360*(6385), 171–175.

Khan, J. I. (2019). Muslims in Indian labour market: Access and opportunities. *Sage Publications Pvt. Limited*.

Khemani, S. (2004). Political Cycles in a Developing Economy: Effect of Elections in the Indian States. *Journal of Development Economics 73*(1), 125–154.

Klein, M. (1996, 02). Timing Is All: Elections and the Duration of United States Business Cycles. *28*, 84–101.

Kline, P. and A. Santos (2012). A Score Based Approach to Wild Boostrap Inference. *Journal of Econometric Methods 1(1)*, 23–41.

Korotayev, A. (2000). Parallel-cousin (fbd) marriage, islamization, and arabization. *Ethnology*, 395–407.

Kremer, M. (1993). The O-ring theory of economic development. *The Quarterly Journal of Economics 108*(3), 551–575.

Kujur, R. (2008, September). Naxal Movement in India : An Profile. *IPCS Research Papers* (15).

Kujur, R. (2009, February). Naxal Conflict in 2008 : An Assessment. *IPCS Issue Brief* (93).

Lange, F. (2007). The speed of employer learning. *Journal of Labor Economics 25*(1), 1–35.

Lasker, G. W. (1985). *Surnames and genetic structure*, Volume 1. Cambridge University Press.

Lazear, E. P. (1998). Personnel economics for managers. *Wiley New York*.

Lindbeck, A. (1976). Stabilization Policy in Open Economies with Endogenous Politicians. *The American Economic Review 66*(2), 1–19.

Litwack, L. (1979). *Been so long in the storm: The aftermath of slavery*. Alfred Knopf, New York.

Lowe, M. (2021). Types of contact: A field experiment on collaborative and adversarial caste integration. *American Economic Review* (6), 1807–44.

Lowes, S. (2020). Matrilineal kinship and spousal cooperation: Evidence from the matrilineal belt. Technical report, University of California, San Diego.

Mandal, A. and D. Sengupta (2000). The Analysis of Fatal Accidents in Indian Coal Mines. *Calcutta Statistical Association Bulletin 50*(1-2), 95–120.

Marx, B., V. Pons, and T. Suri (2021). Diversity and team performance in a Kenyan organization. *Journal of Public Economics 197*, 104332.

Mas, A. and E. Moretti (2009). Peers at work. *American Economic Review 99*(1), 112–45.

McCallum, B. T. (1978). The Political Business Cycle: An Empirical Test. *Southern Economic Journal 44*(3), 504–515.

Mete, C., L. Bossavie, J. Giles, and H. Alderman (2020). Is consanguinity an impediment to child development? *Population Studies 74*(2), 139–159.

Miguel, E. (2004). Tribe or nation? Nation building and public goods in Kenya versus Tanzania. *World Politics 56*(3), 327–362.

Mitra, A. and D. Ray (2014). Implications of an economic theory of Conflict: Hindu-Muslim violence in India. *Journal of Political Economy 122*(4), 719–765.

Mobarak, A. M., T. Chaudhry, J. Brown, T. Zelenska, M. N. Khan, S. Chaudry, R. A. Wajid, A. H. Bittles, and S. Li (2019). Estimating the health and socioeconomic effects of cousin marriage in south asia. *Journal of biosocial science 51*(3), 418–435.

Mobarak, A. M., R. Kuhn, and C. Peters (2013). Consanguinity and other marriage market effects of a wealth shock in bangladesh. *Demography 50*(5), 1845–1871.

Mohanty, N. (2017). *Political Economy of Mining in India.* Har-Anand Publications.

Montalvo, J. G. and M. Reynal-Querol (2017). Ethnic diversity and growth: Revisiting the evidence. *Review of Economics and Statistics*, 1–43.

Morgan, L. H. (1877). *Ancient Society; Or, Researches in the Lines of Human Progress from Savagery, Through Barbarism to Civilization.* H. Holt.

Moscona, J., N. Nunn, and J. Robinson (2020). Segmentary lineage organization and conflict in sub-saharan africa. *Econometrica 88*, 1999–2036.

Mousa, S. (2018). Overcoming the trust deficit: Inter-group contact and associational life in post-ISIS Iraq. *Science. Vol. 369, Issue 6505, pp. 866-870.*.

Munshi, K. and M. Rosenzweig (2016). Networks and misallocation: Insurance, migration, and the rural-urban wage gap. *American Economic Review 106*(1), 46–98.

Murdock, G. P. (1949). *Social structure.* Macmillan.

Naroll, R. (1970). What have we learned from cross-cultural surveys? 1. *American Anthropologist 72*(6), 1227–1288.

Nath, S. and S. R. Chowdhury (2019). Mapping polarisation: Four ethnographic cases from West Bengal. *Journal of Indian Anthropological Society 54*, 51–64.

Nordhaus, W. (1975). The Political Business Cycle. *Review of Economic Studies 42*(2), 169–190.

Ottenheimer, M. (1990). Lewis henry morgan and the prohibition of cousin marriage in the united states. *Journal of Family History 15*(1), 325–334.

Ottenheimer, M. (1996). *Forbidden relatives: The American myth of cousin marriage*. University of Illinois Press.

Paluck, E. L., S. A. Green, and D. P. Green (2019). The contact hypothesis re-evaluated. *Behavioural Public Policy 3*(2), 129–158.

Parrotta, P., D. Pozzoli, and M. Pytlikova (2012). Does labor diversity affect firm productivity? *IZA Discussion Paper*.

Paul, D. B. and H. G. Spencer (2008). "it's ok, we're not cousins by blood": the cousin marriage controversy in historical perspective. *PLoS Biol 6*(12), e320.

Paul, D. B. and H. G. Spencer (2016). Eugenics without eugenists? anglo-american critiques of cousin marriage in the nineteenth and early twentieth centuries. *Heredity Explored: Between Public Domain and Experimental Science, 1850–1930*, 49.

Persson, T. and G. E. Tabellini (1990). *Macroeconomic Policy, Credibility and Politics*, Volume 38. New York, N.Y.: Harwod Academic Publishers.

Pettigrew, T. F., L. R. Tropp, U. Wagner, and O. Christ (2011). Recent advances in intergroup contact theory. *International Journal of Intercultural Relations 35*(3), 271–280.

Pillalamarri, A. (2019). The origins of Hindu-Muslim conflict in South Asia. *Washington, DC: The Diplomat*.

Prakash, O. (2014). The Use of Technology by the Maoists in the Conflict Against the State In India. *Proceedings of the Indian History Congress 75*, 1277–1284.

Rao, G. (March, 2019). Familiarity does not breed contempt: Generosity, discrimination and diversity in Delhi schools. *American Economic Review Vol. 109, No. 3*.

Ray, U. K. (2022). In Jharkhand, death of 5 labourers is testament to coal mafia's chokehold on poor. *The Wire*.

Reid, R. M. (1988). Church membership, consanguineous marriage, and migration in a scotch-irish frontier population. *Journal of Family History 13*(4), 397–414.

Relethford, J. H. (2017). Comparison of observed and expected levels of genetic diversity based on surname frequencies: An example from historical massachusetts. *American journal of physical anthropology 163*(1), 200–204.

Rogoff, K. and A. Sibert (1988). Elections and Macroeconomic Policy Cycles. *The Review of Economic Studies 55*(1), 1–16.

Ruggles, S., K. Genadek, R. Goeken, J. Grover, and M. Sobek (2015). Integrated public use microdata series: Version 6.0 [dataset]. Minneapolis: University of Minnesota.

Saez, L. and A. Sinha (2010). Political Cycles, Political Institutions and Public Expenditure in India, 1980-2000. *British Journal of Political Science 40*(1), 91–113.

Saggar, A. K. and A. H. Bittles (2008). Consanguinity and child health. *Paediatrics and Child Health 18*(5), 244–249.

Sasi, A. (2014). One death every third day in india's most dangerous job. *The Indian Express*.

Schneider, D. M. and G. C. Homans (1955). Kinship terminology and the american kinship system. *American anthropologist 57*(6), 1194–1208.

Schulz, J. F. (2019). Kin networks and institutional development. Technical report, SSRN Working Paper.

Schulz, J. F., D. Bahrami-Rad, J. P. Beauchamp, and J. Henrich (2019). The church, intensive kinship, and global psychological variation. *Science 366*(6466).

Srivastava, D. (2009). Terrorism and Armed Violence In India. *Institute of Peace and Conflict Studies, IPCS Special Report, no. 71*.

Srivastava, R. (2005, April). Bonded Labor in India : Its Incidence and Pattern. *Cornell University ILR School*.

Stevenson, B. and J. Wolfers (2007). Marriage and divorce: Changes and their driving forces. *Journal of Economic perspectives 21*(2), 27–52.

Swedlund, A. C. and A. Boyce (1983). Mating structure in historical populations: estimation by analysis of surnames. *Human biology*, 251–262.

Sykes, B. and C. Irven (2000). Surnames and the y chromosome. *The American Journal of Human Genetics 66*(4), 1417–1419.

Talbot, I. and G. Singh (2009). *The partition of India*. Cambridge University Press.

Taylor, R. (1974). John doe, jr.: A study of his distribution in space, time, and the social structure. *Social Forces 53*(1), 11–21.

Thomas, J., M. Doucette, D. C. Thomas, and J. Stoeckle (1987). Disease, lifestyle, and consanguinity in 58 american gypsies. *The Lancet 330*(8555), 377–379.

Tönnies, F. (1957). *Community & Society (Gemeinschaft und Gesellschaft)*. Transaction Publishers.

Uppsala Conflict Data Program (2018). *State Based Violence, India*.

Vanden Eynde, O. (2016). Targets of Violence: Evidence from India's Naxalite Conflict. *The Economic Journal 128*(609), 887–916.

Veiga, L. and F. Veiga (2007). Political Business Cycles at the Municipal Level. *Public Choice 131*(1), 45–64.

Walker, R. S. and D. H. Bailey (2014). Marrying kin in small-scale societies. *American Journal of Human Biology 26*(3), 384–388.

Weber, M. (1951). *The Religion of China: Confucianism and Taoism*, Volume 93445. Free Press.

Wirth, L. (1938). Urbanism as a way of life. *American journal of sociology 44*(1), 1–24.

Woodley, M. A. and E. Bell (2013). Consanguinity as a major predictor of levels of democracy: a study of 70 nations. *Journal of Cross-Cultural Psychology 44*(2), 263–280.

Yamin, P. (2009). The search for marital order: civic membership and the politics of marriage in the progressive era. *Polity 41*(1), 86–112.

Zhou, M. (2004). Are Asian-Americans becoming "White?". *Contexts 3*(1), 29–37.

Zhou, M. and Y. S. Xiong (2005). The multifaceted American experiences of the children of Asian immigrants: Lessons for segmented assimilation. *Ethnic and Racial Studies 28*(6), 1119–1152.

Appendix A

Appendix to Chapter 2

A.1 Randomization steps, implementation timeline and balance (identification) checks

A.1.1 Randomization steps and timeline

Each step involved in the randomization process is described in detail below.

Step 0: Determine religious composition of each section in each line

For each section of each line, first decide final number of Hindus and Muslims (typically 35%-40% Muslims in mixed sections)

s.t. $\sum Hs = \bar{H}$ *and* $\sum Ms = \bar{M}$, where $\bar{H}$ and $\bar{M}$ denote the total number of Hindus and Muslims in the line across all three cohorts.

Workers were not moved across production lines for randomization. Therefore, the religious composition of line-section-level teams was constrained by the overall number of Hindus and Muslims in the line at baseline. Since the proportion of Muslim workers in each line was very close to the overall share of Muslims in the factory, mixed sections (both HD and LD) ended up with roughly 35%-40% Muslim workers after randomization.

Step 1: Section Shifting

Suppose 2 additional Muslim workers are required in a section to achieve the desired religious composition (35%-40% Muslims in mixed teams). Then the following steps are taken:

a) Randomly order workers within section × *religion* × *skill*

b) Find a section with enough Muslims

c) Randomly pick 2 Muslim workers to shift in

d) Randomly pick 2 Hindu workers to shift out

This step is perhaps the most crucial in order to achieve the desired line-level treatment types described in Figure 2.3. At baseline, not all sections of all lines (across all 3 cohorts) had enough Muslim workers to achieve 35%-40% Muslim workers in mixed line-section-level teams post randomization. Therefore, workers were moved across sections in this manner to achieve that. This also meant that only the minimum number of workers required were moved, satisfying the firm's requirement of minimizing section-switching.

Step 2: Re-randomize

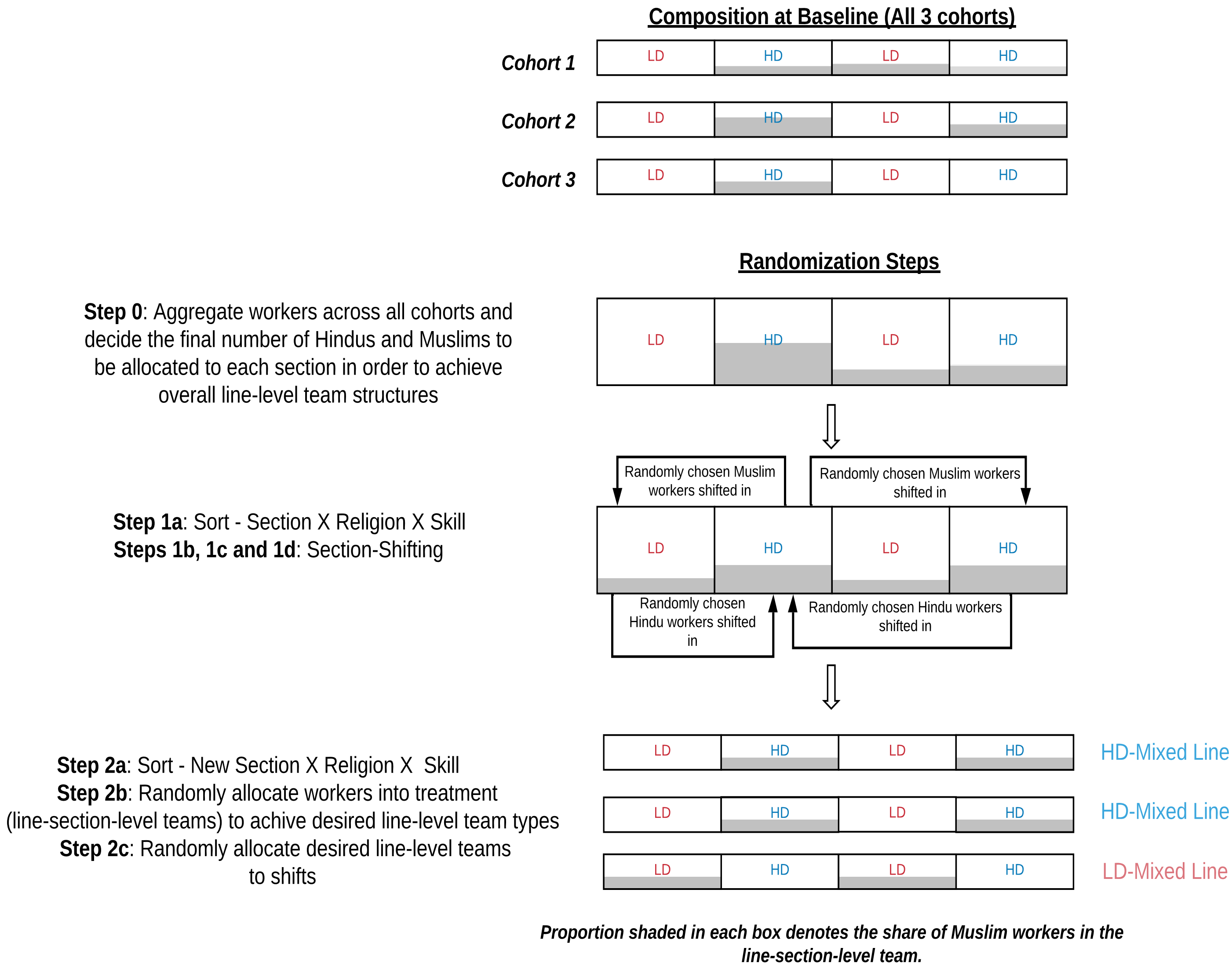
a) Randomly order within new section × *religion* × *skill level*

b) Allocate workers into mixed vs homogeneous teams as pre-specified

c) Randomly allocate teams (lines) to shifts/supervisors

In Step 2, workers were sorted by their new section (only workers who were moved in Step 1 had a different section than at baseline), religion and skill and allocated to line-section-level teams (recall that there are three teams per section in a line – one for each shift). The line-section-level teams were then aggregated to form line-level teams in accordance with two different line-level team structures (treatment types), as in Figure 2.3 (i.e HD-Mixed lines or LD-Mixed lines). Finally, the line-level teams were randomly allocated to the three shifts and the usual weekly shift rotations were introduced. Figure A.1 provides a visual representation of these steps.

Figure A.1: Randomized steps (From baseline structure to randomized teams)



Note: This figure illustrates the steps involved in the randomization process – from how given the religious composition of sections at baseline the desired line-level team types are achieved. The figure is based on the description of the steps discussed in section A.1.1. A production line with only four sections is considered for simplicity.

A.1.2 Quasi-random allocation of workers to tasks at baseline

Hiring at the factory occurs on a rolling basis as and when vacancies become available for each position on a production line. The HR manager always has a pool of job applicants at hand who are called upon on a first-come-first-served basis. As a result, workers do not have the option to choose their area of work when they join. It is possible that workers quit at a different rate across the two types of tasks (HD and LD), leading to possible selection bias. However, if that were the case, this would be reflected in the average tenure of workers in HD and LD sections. As shown in Table A.5, this is not the case - tenure is balanced between workers in HD and LD sections.

Table A.1: Dependency switches

First Job/Final Job	Low-Dependency	High-Dependency	Total
Low-Dependency	148	35	183
High-Dependency	59	344	403
Total	207	379	586

Note: This matrix reports the number of workers who, from when they first joined the factory until before the intervention, switched jobs that also involved switching dependencies. 35 workers (5.9%) switched from low- to high-dependency, while 59 workers (10%) switched from high- to low-dependency. While 15.9% of the workers switched jobs at least once, 6.85% of them held one or more job between their first and final job at the factory.

While selection into jobs is therefore unlikely at hiring, it is possible that over time, workers are able to sort into their sections of choice. In order to assess if that is the case workers were asked to report their first job at the factory and their final job immediately before the intervention began. They were also asked to report any other job that they held for a period of more than six months at the factory. Table A.1 reports a matrix of job switches between HD and LD sections. Only 94 out of 586 workers (16%) reported to be currently in jobs that involved switching *dependency* from their first job. Only 6.85% of the workers reported to switch jobs more than once, whereby majority of the workers who switched jobs did so only once. Additionally, many of these changes resulted from a closure of one production line at the factory in 2018. As a result, workers from that line were reallocated, typically to similar jobs, in the same shift, but to other existing lines and an additional line which was bought around the same time.

Table A.2: Dependency sorting

	(1) **Switched Dependency**	(2) **High to Low**	(3) **Low to High**
Age	0.0040	-0.0029	-0.0011
	(0.0020)	(0.0018)	(0.0011)
Tenure	0.0016	-0.0040	0.0023
	(0.0067)	(0.0048)	(0.0038)
Schooling (Highest grade)	0.0046	-0.0024	-0.0022
	(0.0032)	(0.0026)	(0.0019)
Muslim	-0.0070	0.0154	-0.0084
	(0.0388)	(0.0258)	(0.0279)
Worker Skill			
Semi-Skilled	-0.201	0.288	-0.0867
	(0.207)	(0.175)	(0.0509)
Operator	-0.0423	0.0946**	-0.0523
	(0.0542)	(0.0374)	(0.0317)
Line × Section F.E. (First Job)	Yes	Yes	Yes
N	579	579	579
Adj. R^2	0.068	0.094	0.284

* p<0.10, ** p<0.05, *** p<0.010. The unit of observation is an individual worker. Workers were asked to report their first job at the factory and their last job before the intervention began. Standard errors clustered at the worker's first line-section-level. "Switched Dependency" refers to whether the move between the first and last job (if any) involved changing dependency as well. Workers were also asked to report if they held any other job in between. Only 7.2 % reported that they did. Workers are categorized into the the following skill categories: unskilled, semi-skilled or operators. Unskilled workers are the omitted group.

Overall, this suggests that only a small share of workers switched jobs from when they first joined, until the time of the intervention. This rules out systematic sorting into tasks over time and possible selection bias resulting from it. Nevertheless, in Table A.2, I test whether observable characteristics of the workers are correlated with the probability of moving across task types, based on the few moves that have occurred, as shown in Table A.1. As observed, none of the factors (age, tenure, schooling, religion) which could potentially affect sorting over time, are statistically significant in Column (1). In Columns (2) and (3), I split up job switches from HD to LD and LD to HD sections. Again, the coefficients on the the usual factors are small in magnitude and not statistically significant. In Column (2) however, it can be seen that workers who are currently Operators are likely to have switched from HD to LD tasks at a higher rate than workers of other skill-levels.

Table A.3: Dependency sorting: Omitting workers shifted from shut production line

	(1) **Switched Dependency**	(2) **High to Low**	(3) **Low to High**
Age	0.0050*	-0.0035	-0.0015
	(0.0022)	(0.0021)	(0.0016)
Tenure	-0.0092	0.0035	0.0057
	(0.0059)	(0.0043)	(0.0043)
Schooling (Highest Grade)	0.0036	-0.0019	-0.0017
	(0.0034)	(0.0028)	(0.0021)
Muslim	-0.0268	0.0051	0.0217
	(0.0350)	(0.0260)	(0.0229)
Worker Skill			
Semi-Skilled	0.0455	0.0474	-0.0929
	(0.0832)	(0.0668)	(0.0536)
Operator	0.0097	0.0570	-0.0668
	(0.0628)	(0.0427)	(0.0394)
Line × Section F.E. (First Job)	Yes	Yes	Yes
N	470	470	470
Adj. R^2	0.044	0.000	0.266

* p<0.10, ** p<0.05, *** p<0.010. The unit of observation is an individual worker. Workers were asked to report their first job at the factory and their last job before the intervention began. Standard errors clustered at the worker's first line-section-level. "Switched Dependency" refers to whether the move between the first and last job (if any) involved changing dependency as well. Workers were also asked to report if they held any other job in between. Only 7.2 % reported that they did. Workers are categorized into the the following skill categories: unskilled, semi-skilled or operators. Unskilled workers are the omitted group.

Table A.3 shows that this is actually a result of the one-time move of workers from the line that was shut in 2018. If I leave this set of workers out of the analysis (as in Table A.3), Operators are no more likely to have switched from HD to LD tasks than other workers. This is understandable since Operators in the line that has now been shut were all in *Packing* sections, which was a HD section in that line.[98] However, *Packing* sections in the six production lines which are part of this experiment are a combination of both HD and LD types. As a result, some Operators mechanically moved from HD to LD jobs when this change occurred, despite continuing to be *Packing* Operators in terms of their specific role in the production line.

Table A.4 shows that the share of Muslim workers was balanced across HD and LD tasks after ran-

[98]To determine this I asked supervisors to compare the *Packing* task in the line that has been shut to *Packing* tasks in lines that are currently operative.

domization. This is important to rule out that the different effects of religious mixing in HD and LD tasks are caused by different “degrees" of mixing rather than the effects being driven by the production technology. Finally, in Table A.5, I report balance in work characteristics across treatment arms without the inclusion of line × section fixed effects. Therefore, unlike in Table 2.3, the main effect of being in HD versus LD section *is* identified. If workers were able to systematically sort into HD and LD tasks based on certain observable characteristics, then the main effect of HD versus LD should pick these differences up. This however is not the case, it can be observed that worker characteristics are balanced between HD and LD sections overall.

Table A.4: Balance in proportion Muslim

	(1) Proportion Muslim
HD vs LD mixed sections	0.0416 (0.0455)
Mean Dep. Var.	0.36
Line F.E.	Yes
N	56
Adj. R^2	0.235

* p<0.10, ** p<0.05, *** p<0.010. The unit of observation is a line-section. Standard errors clustered at the production line-level. This table shows that the proportion of Muslim workers in mixed teams was balanced across HD and LD sections after randomization.

Table A.5: Randomization check

	Panel A: Outcomes relevant at work				**Panel B: Other outcomes**				
	Tenure	Muslim co-workers *Hindus*	Taking Orders	Communicating	Age	Schooling	Trust	Altruism	Inter-religious contact outside work
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mixed	0.0224	0.0144	0.0480	0.0676	1.3496	-0.0283	0.6070	0.0765	0.0368
	(0.4650)	(0.0219)	(0.0587)	(0.0686)	(1.5003)	(0.6409)	(0.3815)	(0.2438)	(0.0520)
HD	-0.5837	0.0187	0.0012	0.0037	1.2019	-0.4931	0.2631	0.1540	0.0148
	(0.4017)	(0.0158)	(0.0618)	(0.0652)	(1.1706)	(0.4945)	(0.3774)	(0.2155)	(0.0501)
Mixed × HD	-0.0575	-0.0039	-0.0505	-0.1149	-0.6911	0.4797	-0.6483	-0.0911	-0.0290
	(0.5627)	(0.0286)	(0.0879)	(0.0881)	(1.7033)	(0.7258)	(0.4823)	(0.2781)	(0.0654)
Mean Dep Var.	4.45	0.12	0.73	0.53	34.47	7.84	3.79	6.65	0.45
Production Line F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Religion F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	586	478	586	586	586	586	586	586	586
Adj. R^2	0.089	0.046	0.008	0.025	0.024	0.021	0.002	0.003	0.109

* p<0.10, ** p<0.05, *** p<0.010. The unit of observation is an individual worker. Standard errors clustered at the line-section-level team. "Tenure" and "Schooling" are measured in years and as highest grade completed respectively. "Taking Orders" is a dummy variable coded 1 if the respondent reported to be always comfortable taking orders from non-coreligionists and 0 if they reported to be sometimes or always uncomfortable. "Communicating" is coded 1, 0.5 and 0 for the responses “Always comfortable”, “Sometimes uncomfortable” and “Always uncomfortable” when asked about being comfortable communicating with non-coreligionists. Survey questions on “Trust" and “Altruism" are used from the World Value Survey (WVS). The dependent variable "Inter-religious contact" refers to the degree of cross-religion interaction that workers had at baseline, outside of work. The variable is coded 1, 0.5 and 0 if a worker mentioned that during the daily course of their life they: 1) interact with more than 5 non-coreligionists 2) interact with 1 to 5 non-coreligionists, or 3) do not interact with anyone outside their religion, respectively.

Table A.6: Randomization check (Line-level treatment indicator)

	Panel A: Outcomes relevant at work				**Panel B: General characteristics and attributes**				
	Tenure	Muslim co-workers *Hindus*	Taking Orders	Communicating	Age	Schooling	Trust	Altruism	Inter-religious contact outside work
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
HD-Mixed Line vs	-0.0471	0.0070	-0.0178	0.0680*	-0.2390	0.0147	-0.1808	0.0111	-0.0085
LD-Mixed Line	(0.4473)	(0.0099)	(0.0265)	(0.0268)	(1.1653)	(0.2386)	(0.3178)	(0.1253)	(0.0366)
Bootstrap (Wild Cluster) C.I.	[-1.29, 2.13]	[-0.025, 0.029]	[-0.088, 0.046]	[-0.026, 0.156]	[-2.522, 3.758]	[-0.678, 0.475]	[-1.297, 0.544]	[-0.292, 0.422]	[-0.175, 0.070]
Mean Dep. Var	4.41	0.12	0.73	0.47	33.88	7.92	3.88	6.68	0.45
Production Line F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Religion F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	557	459	557	557	557	557	554	554	557

* $p<0.10$, ** $p<0.05$, *** $p<0.010$. The unit of observation is an individual worker. Standard errors clustered at the line-level. "Tenure" and "Schooling" are measured in years and as highest grade completed respectively. "Taking Orders" is a dummy variable coded 1 if the respondent reported to be always comfortable taking orders from non-coreligionists and 0 if they reported to be sometimes or always uncomfortable. "Communicating" is coded 1, 0.5 and 0 for the responses "Always comfortable", "Sometimes uncomfortable" and "Always uncomfortable" when asked about being comfortable communicating with non-coreligionists. Survey questions on "Trust" and "Altruism" are used from the World Value Survey (WVS). The dependent variable "Inter-religious contact" refers to the degree of cross-religion interaction that workers had at baseline, outside of work. The variable is coded 1, 0.5 and 0 if a worker mentioned that during the daily course of their life they: 1) interact with more than 5 non-coreligionists 2) interact with 1 to 5 non-coreligionists, or 3) do not interact with anyone outside their religion, respectively. This sample excludes individuals who work in common sections ("Egg" and "Flour") that cater to all production lines, but themselves are not part of any particular line.

A.2 Treatment effect on standard output and output gap

Supervisors keep records of standard (expected) output against actual output produced, in each shift for each line. This measure is based on inputs used in the production process; negative deviations from the standard level of output imply lower productivity and higher raw material wastage. In Table A.7 (Column 2), I use percentage deviation of actual output from standard output as the outcome variable. The formula for "Output Gap" is $\frac{ActualOutput - StandardOutput}{StandardOutput} * 100$. The coefficient estimate shows that on average, HD-Mixed lines fall short of their expected target by a greater degree than LD-Mixed lines. This happens despite the fact that on average HD-Mixed lines receive lower targets (Column 1), though this difference is not statistically significant (i.e. the treatment effect on standard output is not significant). This suggests that the treatment effects in Table 2.4 indeed result from under-performance of workers in HD-Mixed lines and are not simply a bi-product of differential target setting across treatment groups.

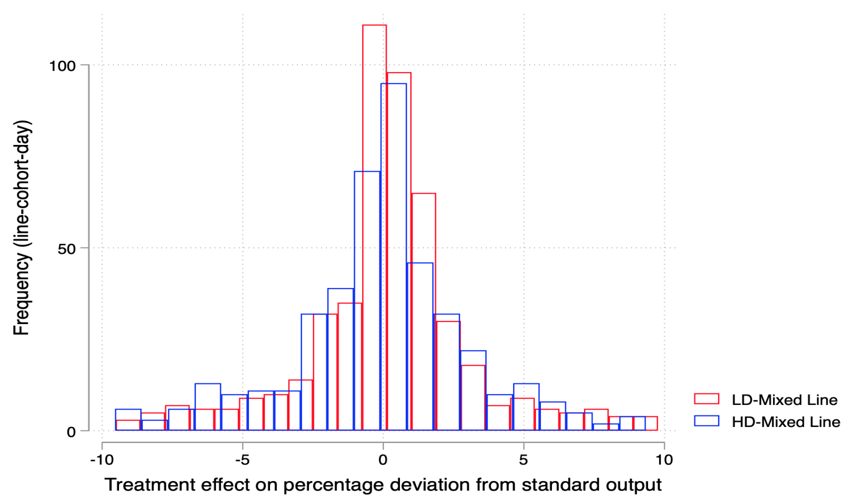
In Figure A.2, I plot histograms of deviation from standard output for each team type. There are a couple of important things to be observed in this figure. First, LD-Mixed lines have lower variance in deviation from standard output relative to HD-Mixed lines. Second, in HD-Mixed lines, negative deviations from standard output occur with greater frequency. At the same time however, large positive deviations from standard output are not completely unusual. This suggests that religious diversity leads to greater uncertainty (in terms of achieving daily targets) in HD mixed sections relative to LD mixed ones. Aditionally, this points to the fact that output in mixed teams might be more susceptible to idiosyncratic shocks (religious events, conflict etc.) in HD sections due to the tight-knit nature of intergroup contact. This is formally tested in Table A.8. I generate rolling standard deviation measures of the Output Gap variable and report that standard deviation in Output Gap is higher in HD-Mixed (LD Non-Mixed) lines. I show robustness to a range of window sizes in generating the rolling standard deviation measures. In Figure A.3, I plot cdfs of deviation of actual output from standard output (by line-level team type) and show that the probability that actual output is greater than standard output is higher throughout, in LD-Mixed lines.

Table A.7: Treatment effect on line-level standard output

	(1) **Log Standard Output**	(2) **Output Gap**
HD-Mixed vs LD-Mixed Line	-0.0223	-1.669***
	(0.0300)	(0.479)
Bootstrap Wild cluster C.I.	[-0.082, -0.0152]	[-3.833, -0.690]
Day F.E.	Yes	Yes
Shift F.E.	Yes	Yes
Production Line F.E.	Yes	Yes
N	1045	1019
Adj. R^2	0.640	0.0488

* p<0.10, ** p<0.05, *** p<0.010. Observations are at the line-cohort-day level. Standard errors clustered at the line-level team in parenthesis. Wild cluster bootstrap confidence intervals in [] brackets. HD-Mixed Line is a dummy coded 1 for a line-level team with all HD sections religiously mixed and LD sections non-mixed, and 0 for exactly the opposite line-level structure (LD-Mixed Line). Standard Output is calculated from the amount of inputs used (batches mixed) in a shift and is determined before the shift begins. Output Gap gap is a measure of deviation from standard output and is calculated as $\frac{ActualOutput - StandardOutput}{StandardOutput} * 100$.

Figure A.2: Percentage deviation from standard output



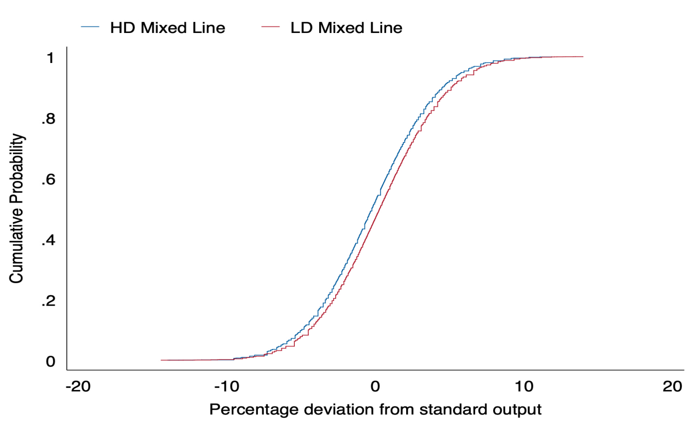
Note: This figure shows percentage deviation from standard (expected) output for HD-Mixed and LD-Mixed lines. Observations are at the line-cohort-day level.

Table A.8: Treatment effect on standard deviation of output gap

	Standard Deviation of Output Gap		
	(1)	(2)	(3)
Rolling Window	10 days	15 days	20 days
Min. Observations	5	10	15
HD Mixed vs LD-Mixed Line	0.954 (0.857)	2.784* (1.358)	3.221** (1.342)
Bootstrap Wild cluster C.I.	[-1.38, 3.227]	[-1.394, 5.993]	[-1.583, 9.907]
Shift F.E.	Yes	Yes	Yes
Production Line F.E.	Yes	Yes	Yes
Day F.E.	Yes	Yes	Yes
Mean Dep Var.	4.77 (7.01)	5.00 (6.89)	5.34 (6.84)
N	755	404	231
Adj. R^2	0.374	0.552	0.668

* p<0.10, ** p<0.05, *** p<0.010. Observations are at the line-cohort-day level. Standard errors clustered at the line-level team in parenthesis. Wild cluster bootstrap (Cameron et al., 2008) confidence intervals in square brackets. HD-Mixed Line is a dummy coded 1 for a line-level team with all HD sections religiously mixed and LD sections non-mixed, and 0 for exactly the opposite line-level structure (LD-Mixed Line). Rolling Window refers to the number of consecutive production days used to generate the standard deviation measure. Min. observations denote the lower bound on the number of observations in each window.

Figure A.3: Deviation from standard output



Note: This figure presents CDFs of the “Output Gap" measure (which is defined as the percentage deviation from expected output) by line-level team type. HD-Mixed lines fall short of expected output with higher probability than LD-Mixed lines.

A.3 Additional tables referred to in the main text

A.3.1 Summary statistics

Table A.9 presents summary statistics of key characteristics of Hindu and Muslim workers described in section 2.2.

Table A.9: Summary statistics: Hindu and Muslim workers

Variable	**Hindu**	**Muslim**	**Diff (2) - (1)**
Panel A: Dependency			
High Dependency (share of workers)	0.610	0.660	0.048
	(0.02)	(0.05)	(0.052)
Panel B: Schooling and Tenure			
Schooling (Grade)	8.08	6.83	-1.250***
	(0.16)	(0.34)	(0.370)
Tenure	4.81	2.75	-2.059***
	(0.15)	(0.28)	(0.353)
Panel C: Cross-religion interaction and attitudes			
Cross-religion interaction (outside work)	0.39	0.73	0.343***
	(0.02)	(0.03)	(0.040)
Comfortable taking orders from non-coreligionists	0.73	0.76	0.032
	(0.02)	(0.040)	(0.047)
Would live next door to non-coreligionists	0.57	0.88	0.307***
	(0.02)	(0.02)	(0.038)
Equally comfortable communicating with non-coreligionists	0.49	0.68	0.191***
	(0.02)	(0.04)	(0.047)
Panel D: Political			
Supports National Registrar of Citizens (NRC)	0.32	0.19	-0.132***
	(0.02)	(0.04)	(0.049)
N	480	106	586
Panel E: Skill			
Proportion Semi-skilled/Operator	0.22	0.16	-0.064
	(0.02)	(0.03)	(0.041)
N	575	116	691

* p<0.10, ** p<0.05, *** p<0.010. Standard errors in parentheses. "Cross-religion interaction (outside work)" is a categorical variable coded 1, 0.5 and 0 if an individual reported to come in contact with greater than 5, between 1 and 5 or 0 non-coreligionists respectively in their daily life outside of work. "Comfortable taking orders from non-coreligionists" is a dummy variable coded 1 if the respondent reported to be always comfortable taking orders from non-coreligionists and 0 if they reported to be sometimes or always uncomfortable. "Equally comfortable communicating with non-coreligionists" is coded 1, 0.5 and 0 for the responses “Always comfortable”, “Sometimes uncomfortable” and “Always uncomfortable” respectively.

The number of workers interviewed at baseline is larger than the number of workers that actually participated in the study. This is because the firm decided to lay off some workers after the baseline survey (but before the intervention began) due to low product demand in two of the production lines (which is why there are only 15 line-level teams). The table includes only those that were part of the experiment (except for the data on worker-skill).

Table A.10 presents summary statistics of key aspects of the physical environment of HD and LD sections. Please refer to section 2.2 for a detailed description of this table.

Table A.10: Summary statistics: Mean differences (physical environment)

Variable	**Low-Dependency**	**High-Dependency**	**Diff (2) - (1)**
Panel A: Interaction (Minutes out of 10)			
Direct Dependency	2.22	9.50	7.283***
	(0.63)	(0.11)	(0.688)
Non-work interaction	0.89	1.14	0.249
	(0.22)	(0.25)	(0.329)
Panel B: Noise Level (Decibels)			
Avg Noise (Db)	78.47	77.53	-0.941
	(1.42)	(1.66)	(2.170)
Max Noise (Db)	87.46	85.40	-2.055
	(1.72)	(1.65)	(2.394)
Panel C: Temperature (Celsius)			
Section Temperature (°C)	29.08	31.42	2.341*
	(0.92)	(0.72)	(1.197)
N	22	20	42

* p<0.10, ** p<0.05, *** p<0.010. This table reports mean differences in characteristics of HD and LD tasks. In some cases, certain sections can have more than one task, with their degrees of dependency highly correlated. Sections are classified based on the average dependency minutes in each section.

A.3.2 Robustness checks and additional results

Table A.11: Treatment effect on output (Line × Variety fixed effects)

	(1) **Log Output (Pieces)**	(2) **Log Output (Boxes)**
HD-Mixed vs LD-Mixed Line	-0.0473***	-0.0552***
	(0.0137)	(0.0121)
Bootstrap Wild cluster C.I.	[-0.082, -0.0152]	[-0.080, -0.021]
Day F.E.	Yes	Yes
Shift F.E.	Yes	Yes
Production Line × Variety F.E.	Yes	Yes
Mean Dep Var	10.80	6.97
	(1.24)	(0.943)
N	1045	1045
Adj. R^2	0.885	0.851

* p<0.10, ** p<0.05, *** p<0.010. Observations are daily output produced by line-level teams. Standard errors clustered at the line-level team in parenthesis. Wild cluster bootstrap confidence intervals in square brackets. HD-Mixed Line is a dummy coded 1 for a line-level team with all HD sections religiously mixed and LD sections non-mixed, and 0 for exactly the opposite line-level structure (LD-Mixed Line).

Table A.12: Treatment effect on output (Line × Day fixed effects)

	(1) **Log Output (Pieces)**	(2) **Log Output (Boxes)**
HD-Mixed Line (LD Non-Mixed)	-0.0520**	-0.0546*
	(0.0185)	(0.0264)
Bootstrap Wild cluster C.I.	[-0.092, -0.005]	[-0.111, 0.012]
Shift F.E.	Yes	Yes
Production Line × Day F.E.	Yes	Yes
Mean Dep. Var.	10.80	6.27
	(1.24)	(0.943)
N	1045	1045
Adj. R^2	0.900	0.827

* p<0.10, ** p<0.05, *** p<0.010. Observations are daily output produced by line-level teams. Standard errors clustered at the line-level team in parenthesis. Wild cluster bootstrap confidence intervals in square brackets. HD-Mixed Line is a dummy coded 1 for a line-level team with all HD sections religiously mixed and LD sections non-mixed, and 0 for exactly the opposite line-level structure (LD-Mixed Line).

Table A.13: Treatment effect on section ratings

	HD Sections		LD Sections	
	Rating	Rating > Median	Rating	Rating > Median
	(1)	(2)	(3)	(4)
Mixed	-0.0496***	-0.0499***	-0.0005	-0.0024
	(0.0184)	(0.0113)	(0.0142)	(0.0124)
Mean Dep. Var.	3.86	0.47	3.81	0.41
	(0.68)	(0.50)	(0.64)	(0.49)
Education and Tenure Controls	Yes	Yes	Yes	Yes
Day F.E.	Yes	Yes	Yes	Yes
Shift F.E.	Yes	Yes	Yes	Yes
Line × Section F.E.	Yes	Yes	Yes	Yes
N	3466	3466	3443	3443
Adj. R^2	0.609	0.385	0.595	0.324

* p<0.10, ** p<0.05, *** p<0.010. Observations are daily ratings received by line-section-level teams. Standard errors clustered at the line-section-level team. "Mixed" is a dummy variable coded 1 if the line-section-level team is religiously mixed. Education and tenure control for the mean of schooling and tenure of workers in the line-section-level team. In this table, the sample is split into HD and LD sections and it can be observed that while religious mixing leads to lower ratings in HD sections, the effects are small and not statistically significant in LD.

Table A.14: Treatment effect on section ratings (without controls collinear with religion)

	Rating (Raw)		**Rating > Median**	
	(1)	(2)	(3)	(4)
Mixed	-0.0239**		-0.0245***	
	(0.0114)		(0.00846)	
Mixed × LD		-0.0084		-0.0040
		(0.0150)		(0.0126)
Mixed × HD		-0.0394**		-0.0449***
		(0.0175)		(0.0113)
p(Mixed × HD = Mixed × LD)		0.184		0.017
Mean Dep. Var.	3.82	3.82	0.44	0.44
	(0.83)	(0.83)	(0.50)	(0.50)
Education and Tenure Controls	No	No	No	No
Day F.E.	Yes	Yes	Yes	Yes
Shift F.E.	Yes	Yes	Yes	Yes
Line × Section F.E.	Yes	Yes	Yes	Yes
N	6909	6909	6909	6909
Adj. R^2	0.600	0.600	0.358	0.358

* p<0.10, ** p<0.05, *** p<0.010. Observations are daily ratings received by line-section-level teams. Standard errors clustered at the line-section-level team. "Mixed" is a dummy variable coded 1 if the line-section-level team is religiously mixed. Line × Sections fixed effects are included in the all specifications; as a result the main effect of HD versus LD is not separately identified in columns (2) and (4). Education and tenure control for the mean of schooling and tenure of workers in the line-section-level team.

Table A.15: Treatment effect on section ratings: Event study

	Raw Ratings			
	HD Sections		LD Sections	
	(1)	(2)	(3)	(4)
Mixed	-0.0496***		-0.0005	
	(0.0184)		(0.0142)	
Mixed × 0-25 days		-0.1050*		0.0525
		(0.0675)		(0.0615)
Mixed × 26-50 days		-0.0716**		-0.1030
		(0.0355)		(0.0724)
Mixed × 51-75 days		0.0279		-0.0134
		(0.0340)		(0.0400)
Mixed × 76-100 days		-0.0647**		0.0579*
		(0.0319)		(0.0286)
Mixed × 101-120 days		-0.0247		-0.0542
		(0.0532)		(0.0446)
Mean Dep. Var.	3.85	3.85	3.80	3.80
	(0.68)	(0.68)	(0.64)	(0.64)
Education and Tenure Controls	Yes	Yes	Yes	Yes
Day F.E.	Yes	Yes	Yes	Yes
Shift F.E.	Yes	Yes	Yes	Yes
Line × Section F.E.	Yes	Yes	Yes	Yes
N	3466	3466	3443	3443
Adj. R^2	0.609	0.609	0.595	0.596

* p<0.10, ** p<0.05, *** p<0.010. Observations are daily ratings received by line-section-level teams. Standard errors clustered at the line-section-level team. “Mixed" is a dummy variable coded 1 if the line-section-level team is religiously mixed. Line × Sections fixed effects are included in the all specifications; as a result the main effect of HD versus LD is not separately identified. Education and tenure control for the mean of schooling and tenure of workers in the line-section-level team. This table is based on specification 2.2; interactions of “Mixed" with day bins are added.

Table A.16: Treatment effect on worker interactions: Hindus respondents only

	Identified teammate as contributing low effort		Blamed by teammate		Unwilling to give up relief time	
	(1)	(2)	(3)	(4)	(5)	(6)
Mixed	0.0399***		0.0393**		0.0565	
	(0.0139)		(0.0165)		(0.0369)	
Mixed × LD		0.0309		0.0656**		0.0317
		(0.0274)		(0.0260)		(0.0636)
Mixed × HD		0.0421***		0.0330*		0.0631
		(0.0159)		(0.0191)		(0.0441)
p(Mixed × LD = Mixed × HD)		0.665		0.282		0.573
Mean. Dep. Var.	0.14	0.14	0.08	0.08	0.24	0.24
Worker skill F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Line × Section F.E.	Yes	Yes	Yes	Yes	Yes	Yes
N	3020	3020	3009	3009	3056	3056
Adj. R^2	0.015	0.015	0.013	0.013	0.079	0.079

* p<0.10, ** p<0.05, *** p<0.010. Observations are at the worker-teammate level for line-section-level teams i.e. there are (N-1) observations per worker, where N denotes the number of workers in the section. Standard errors clustered at the line-section-level team. Workers were asked to choose teammates who they: (1) think have not contributed sufficient effort at any point during the intervention (2) have been blamed by during the intervention and (3) would give (or already have) up their relief time for.

Table A.17: Treatment effect on section ratings: Adding key controls

	Rating (Raw)			
	(1)	(2)	(3)	(4)
Mixed × LD	-0.0068 (0.0144)	-0.0210 (0.0179)	-0.122** (0.0507)	-0.173*** (0.0603)
Mixed × HD	-0.0349** (0.0185)	-0.0609** (0.0264)	-0.164*** (0.0522)	-0.216*** (0.0614)
Mixed × Group Size		0.0040 (0.0034)	0.0078** (0.0038)	0.0067* (0.0040)
Mixed × Tenure			0.0169** (0.0075)	0.0119 (0.0085)
Mixed × Schooling				0.0102 (0.0072)
p(Mixed × HD = Mixed × LD)	0.229	0.112	0.069	0.075
Mean Dep. Var.	3.82 (0.83)	3.82 (0.83)	3.82 (0.83)	3.82 (0.83)
Education and Tenure Controls	Yes	Yes	Yes	Yes
Day F.E.	Yes	Yes	Yes	Yes
Shift F.E.	Yes	Yes	Yes	Yes
Line × Section F.E.	Yes	Yes	Yes	Yes
N	6909	6909	6909	6909
Adj. R^2	0.600	0.600	0.600	0.600

* p<0.10, ** p<0.05, *** p<0.010. Observations are daily ratings received by line-section-level teams. Standard errors clustered at the line-section-level team. "Mixed" is a dummy variable coded 1 if the line-section-level team is religiously mixed. Line × Sections fixed effects are included in the all specifications; as a result the main effect of HD versus LD is not separately identified. Education and tenure control for the mean of schooling and tenure of workers in the line-section-level team.

Table A.18: Proportion of old teammates

	Proportion of old teammates	
	(1)	(2)
Mixed	-0.0115 (0.0163)	
Mixed × LD		-0.0312 (0.0321)
Mixed × HD		-0.0006 (0.0193)
p(Mixed × HD = Mixed × LD)		0.442
Mean Dep. Var.	0.34	0.34
Religion F.E.	Yes	Yes
Line × Section F.E.	Yes	Yes
N	577	577
Adj. R^2	0.599	0.600

* p<0.10, ** p<0.05, *** p<0.010. The unit of observation is an individual worker. Standard errors clustered at the line-section-level team. The outcome variable in these regressions is the share of co-workers in each individual's line-section-level team that were also in their team at baseline. “Mixed" is a dummy variable coded 1 if the line-section-level team is religiously mixed. Line × Sections fixed effects are included in the all specifications; as a result the main effect of HD versus LD is not separately identified in column (2).

Table A.19: Section change and treatment status

	Changed Section			
	(1)	(2)	(3)	(4)
Mixed	-0.0338 (0.0277)	-0.0288 (0.0249)		
Mixed × LD			-0.0353 (0.0538)	0.0188 (0.0351)
Mixed × HD			-0.0329 (0.0361)	-0.0551* (0.0322)
Mean Dep. Var.	0.079	0.079	0.079	0.079
Religion F.E.	Yes	Yes	Yes	Yes
Line × Section F.E.	No	Yes	No	Yes
Line × Old Section F.E.	Yes	No	Yes	No
N	586	586	586	586
Adj. R^2	0.043	0.030	0.041	0.033

* p<0.10, ** p<0.05, *** p<0.010. The unit of observation is an individual worker. Standard errors clustered at the line-section-level team. “Mixed" is a dummy variable coded 1 if the line-section-level team is religiously mixed. Line × Sections fixed effects are included in the all specifications; as a result the main effect of HD versus LD is not separately identified in columns (3) and (4). These are individual worker-level regressions. The outcome variable is a dummy coded 1 if after the randomization process the worker was in a different section (task) than their section of work at baseline.

Table A.20: Heterogeneous attenuation by characteristics of Hindus at baseline (LD section ratings)

Sample:	**Full**	**Contact at Baseline**		**Tenure at Baseline**		**Support for BJP and NRC**	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Above Median	Below Median	Above Median	Below Median	Above Median	Below Median
Mixed × 0-60 days	-0.0157	-0.0034	-0.0542	-0.0020	-0.0436	-0.0286	0.0086
	(0.0376)	(0.0542)	(0.0538)	(0.0598)	(0.0511)	(0.0499)	(0.0459)
Mixed × 61-120 days	0.0117	-0.0499	0.0467	0.0038	-0.0079	0.0088	-0.0194
	(0.0233)	(0.0413)	(0.0360)	(0.0318)	(0.0326)	(0.0289)	(0.0382)
Mean Dep. Var	3.81	3.74	3.87	3.86	3.74	3.8	3.82
Education and Tenure Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Shift F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Line × Section F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3443	1607	1836	1945	1498	2430	1013
Adj. R^2	0.595	0.622	0.563	0.584	0.606	0.588	0.614

* p<0.10, ** p<0.05, *** p<0.010. Observations are daily ratings received by line-section-level teams. Standard errors clustered at the line-section-level team. “Mixed" is a dummy variable coded 1 if the line-section-level team is religiously mixed. In column (2), the sample consists of all line-section-level teams in which the share of Muslim teammates, that Hindus in that team had at baseline, is above median. In column (3), the sample consists of all line-section-level teams in which the share of Muslim teammates, that Hindus in that team had at baseline, is below median. In columns (4) and (5) teams are split by median tenure of Hindus at baseline. In column (6), the sample consists of all line-section-level teams with above median support for the BJP or the NRC (averaged across all Hindu workers in the team). In Column (6), the sample consists of all line-section-level teams with below median support for the BJP or the NRC (averaged across all Hindu workers in the team).

Table A.21: Treatment effect on worker interactions: Decomposition (Mixed teams by dependency)

	Identified teammate as contributing low effort		**Blamed by teammate**		**Unwilling to give up relief time**	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: HD Sections						
Target Muslim	0.0711***	0.0987***	-0.0139	0.0056	0.0534**	0.0325
	(0.0185)	(0.0276)	(0.0145)	(0.0360)	(0.0215)	(0.0410)
Respondent Muslim	-0.0202	0.0053	-0.0147	0.0081	-0.0423	-0.0620
	(0.0274)	(0.0328)	(0.0255)	(0.0252)	(0.0408)	(0.0438)
Target Muslim × Respondent Muslim		-0.0829		-0.0656		0.0624
		(0.0609)		(0.0495)		(0.0815)
Mean Dep. Var.	0.16	0.16	0.09	0.09	0.31	0.31
N	1576	1576	1584	1584	1568	1568
Adj. R^2	0.023	0.025	0.019	0.021	0.107	0.107
Panel B: LD Sections						
Target Muslim	0.0008	0.0502*	0.0392*	0.0672*	0.0290	-0.0635*
	(0.0276)	(0.0284)	(0.0205)	(0.0371)	(0.0291)	(0.0321)
Respondent Muslim	-0.0080	0.0462	-0.0196	0.0055	0.0505	-0.1538***
	(0.0336)	(0.0382)	(0.0250)	(0.0357)	(0.0386)	(0.0290)
Target Muslim m × Respondent Muslim		-0.1443**		-0.0726		0.2665***
		(0.0512)		(0.0628)		(0.0754)
Mean Dep. Var	0.13	0.13	0.12	0.12	0.19	0.19
N	457	457	445	445	467	467
Adj. R^2	0.01	0.02	0.01	0.01	-0.01	0.02

* p<0.10, ** p<0.05, *** p<0.010. Observations are at the worker-teammate level for line-section-level teams i.e. there are (N-1) observations per worker, where N denotes the number of workers in the section. Standard errors clustered at the line-section-level team. Workers were asked to choose teammates who they: (1) think have not contributed sufficient effort at any point during the intervention (2) have been blamed by during the intervention and (3) would give (or already have) up their relief time for. All regressions include Worker skill F.E. and Line × Section F.E.

Table A.22: Religious violence and section ratings

	HD Sections		LD Sections	
	(1)	(2)	(3)	(4)
Mixed	-0.0496*** (0.0184)		-0.0006 (0.0142)	
Mixed × No Violence		-0.0466** (0.0190)		0.0112 (0.0164)
Mixed × Violence		-0.0749** (0.0376)		-0.0951** (0.0434)
p(Mixed × No Violence = Mixed × Violence)		0.457		0.038
Mean Dep. Var.	3.86 (0.68)	3.86 (0.68)	3.81 (0.64)	3.81 (0.64)
Education and Tenure Controls	Yes	Yes	Yes	Yes
Day Effects	Yes	Yes	Yes	Yes
Shift Effects	Yes	Yes	Yes	Yes
Line × Section Effects	Yes	Yes	Yes	Yes
N	3466	3466	3443	3443
Adj. R^2	0.609	0.609	0.595	0.595

* p<0.10, ** p<0.05, *** p<0.010. Observations are daily ratings received by line-section-level teams. Standard errors clustered at the line-section-level team. "Mixed" is a dummy variable coded 1 if the line-section-level team is religiously mixed. Line × Sections fixed effects are included in the all specifications; as a result the main effect of HD versus LD is not separately identified. Education and tenure control for the mean of schooling and tenure of workers in the line-section-level team. Between 13th-18th December 2019, immediately after the passing of the Citizenship Amendment Act (CAA) violent protests erupted in the district of West Bengal where the factory is located. Hindu-Muslim riots occurred in Delhi between 23rd-28th Feb 2020 during protests against the CAA as well. These days are coded as violent days in these regressions.

Table A.23: Treatment effect on line-level performance (aggregated section ratings)

	Raw Rating (1)	**Rating > Median** (2)
HD-Mixed vs LD-Mixed Line	-0.0126 (0.0200)	-0.0224** (0.0118)
Bootstrap (Wild Cluster) C.I.	[-0.072, 0.038]	[-0.061, 0.005]
Day F.E.	Yes	Yes
Shift F.E.	Yes	Yes
Production Line F.E.	Yes	Yes
Mean Dep. Var	3.87	0.47
N	1012	1012
Adj. R^2	0.734	0.621

* p<0.10, ** p<0.05, *** p<0.010. The dependent variable is daily line-section-level team ratings aggregated to line-level teams (averaging across all sections). Standard errors clustered at the line-level team in parenthesis. Wild cluster bootstrap (Cameron et al., 2008) confidence intervals in square brackets. HD-Mixed Line is a dummy coded 1 for a line-level team with all HD sections religiously mixed and LD sections non-mixed, and 0 for exactly the opposite line-level structure (LD-Mixed Line).

Table A.24: Inter-religious contact (outside work) and section ratings

	HD Sections		LD Sections	
	(1)	(2)	(3)	(4)
Mixed	-0.0496***	-0.0236	-0.0006	0.0121
	(0.0184)	(0.0409)	(0.0142)	(0.0217)
Inter-religious contact (outside work)		0.0092		0.0165
		(0.0749)		(0.0466)
Mixed × Inter-religious contact (outside work)		-0.0621		-0.0365
		(0.0948)		(0.0567)
Mean Dep. Var.	3.86	3.86	3.81	3.81
Education and Tenure Controls	Yes	Yes	Yes	Yes
Day Effects	Yes	Yes	Yes	Yes
Shift Effects	Yes	Yes	Yes	Yes
Line × Section Effects	Yes	Yes	Yes	Yes
N	3466	3466	3443	3368
Adj. R^2	0.609	0.609	0.595	0.595

* p<0.10, ** p<0.05, *** p<0.010. Observations are daily ratings received by line-section-level teams. Standard errors clustered at the line-section-level team. "Mixed" is a dummy variable coded 1 if the line-section-level team is religiously mixed. Line × Sections fixed effects are included in the all specifications; as a result the main effect of HD versus LD is not separately identified. Education and tenure control for the mean of schooling and tenure of workers in the line-section-level team. "Inter-religious contact" refers to the degree of cross-religion interaction that workers had at baseline, outside of work. The variable is coded 1, 0.5 and 0 if a worker mentioned that during the daily course of their life they: 1) interact with more than 5 non-coreligionists 2) interact with 1 to 5 non-coreligionists, or 3) do not interact with anyone outside their religion, respectively. For these regressions the average value of this variable across all Hindus in a line-section-level team is used.

Table A.25: Attrition

	Attrited (1)	Attrited (2)
Mixed	-0.0164	
	(0.0223)	
Mixed × LD		0.0069
		(0.0279)
Mixed × HD		-0.0292
		(0.0296)
p(Mixed X HD = Mixed X LD)		0.35
Mean Dep. Var	0.05	0.05
Religion Effects	Yes	Yes
Line × Section Effects	Yes	Yes
Observations	586	586

* p<0.10, ** p<0.05, *** p<0.010. The unit of observation is an individual worker. The outcome variable is coded 1 for individuals who left the firm before the end of the experiment. Note that the total number of workers who left the firm is actually lower than the number interviewed at endline. A handful of workers could not be reached by phone during the endline survey.

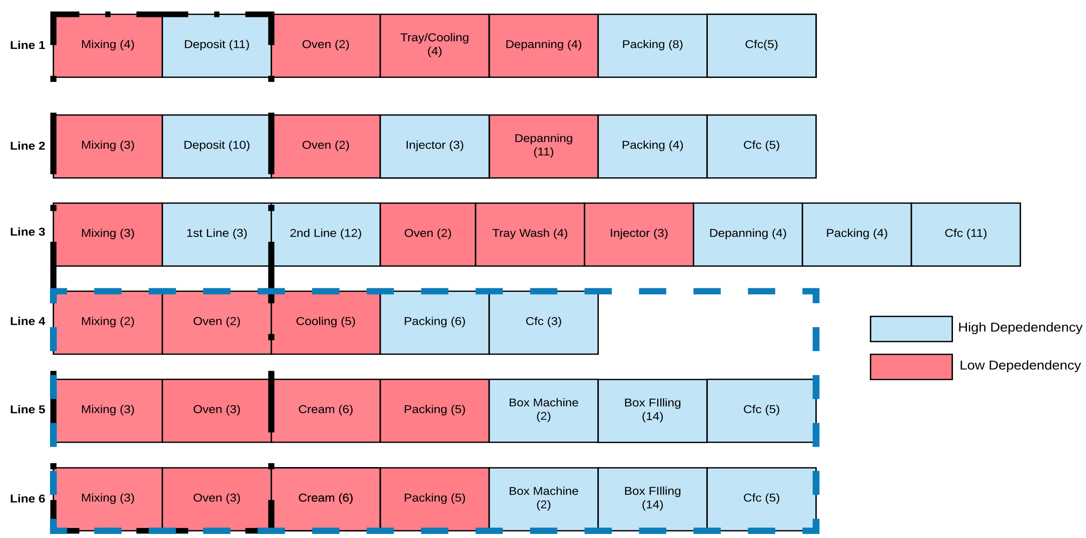
A.3.3 Spillovers

One concern with the analysis of the treatment effects on output is that there could be spillover effects from upstream to downstream sections, potentially biasing my estimates (even though supervisors tried to each section based solely on it's performance). To understand how this could affect the main findings, I restrict attention to the following two sub-samples (as shown in Figure A.4)

1. Only the first two sections of every line (black dashed-dotted portion)

2. Lines where all HD sections come after all LD sections (blue dashed portion)

Figure A.4: Sub-sample analysis



Note: This figure shows all sections of all lines at the factory split into HD and LD types. Direct Dependency is measured as described in section 2.2.3. The two relevant sub-samples used for analysis in this section are highlighted by the black dashed-dotted lines and blue dashed lines.

With sub-sample 1, I first show that there is no effect of religious mixing in the first section which is always a low-dependency section (and by definition cannot be affected by spillovers) (see columns (1) and (2) of Table A.26.). This is consistent with the main results. Furthermore, this suggests that religious mixing is unlikely to cause differential spillover effects from the first section to the second based on treatment status. Finally, once the second section of each line is added to the sample, the main section level results (Table 2.5) are replicated – the magnitude of the effects are also very similar.

Table A.26: Treatment effect on line-section-level ratings

	Only first section (Mixing, only LD)		First two sections	
	Ratings	Ratings > Median	Ratings	Ratings > Median
	(1)	(2)	(3)	(4)
Mixed × LD	-0.0033 (0.0396)	0.0085 (0.0260)	-0.0082 (0.0299)	-0.0032 (0.0206)
Mixed × HD			-0.0318 (0.0373)	-0.0461*** (0.0131)
p(Mixed X HD = Mixed X LD)			0.63	0.09
Mean Dep. Var	3.74	3.74	3.74	3.74
Day Effects	Yes	Yes	Yes	Yes
Shift Effects	Yes	Yes	Yes	Yes
Line × Section Effects	Yes	Yes	Yes	Yes
Observations	964	964	1929	1929

* p<0.10, ** p<0.05, *** p<0.010. Observations are daily ratings received by line-section-level teams. Standard errors clustered at the line-section-level team. "Mixed" is a dummy variable coded 1 if the line-section-level team is religiously mixed. Line × Sections fixed effects are included in the all specifications; as a result the main effect of HD versus LD is not separately identified in columns (2) and (4). Education and tenure control for the mean of schooling and tenure of workers in the line-section-level team.

In Table A.27, I show that the main results are replicated with sub-sample 2 as well. This sample is unique in the sense that it only has production lines where all the HD sections come after the LD sections. I once again find that there is no effect of religious mixing in LD sections. The HD sections at the end of the line are therefore unlikely to be affected differentially by spillovers from LD sections (based on whether they are religiously mixed or not). In HD sections, a negative, large and statically significant effect of religious mixing can still be observed. Taken together, the sub-sample analysis is re-assuring in that they convey the same findings as the core results — which is that religious mixing leads to lower team performance but only in HD tasks.

How should we expect line-section-level spillovers to affect the overall line-level treatment effect estimates (Table 2.4)? Notice that on average production lines have LD sections earlier in the line while HD sections come later. For the line-level effects to be overestimated (in other words the difference in output between HD-Mixed lines and LD-Mixed lines to be more negative than it actually is), it must be the case that there are larger negative spillovers from LD Non-mixed sections to HD-Mixed sections than from LD-Mixed sections to HD Non-Mixed sections, which is unlikely. Therefore, if anything, the line-level results are likely to be underestimated.

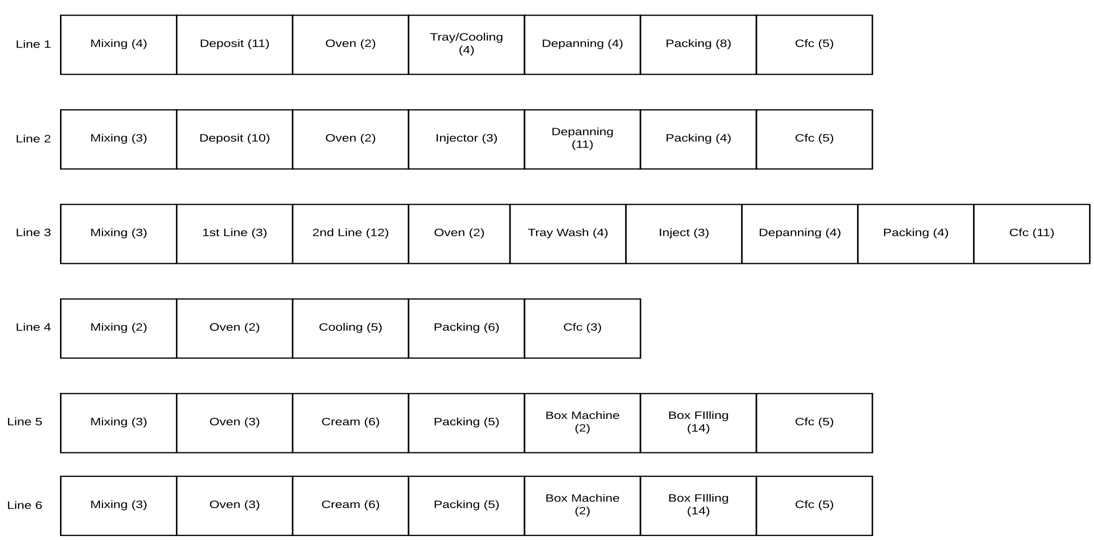
Table A.27: Treatment effect on section ratings (HD after LD)

	Ratings (1)	**Ratings** (2)
Mixed	-0.0391* (0.0195)	
Mixed × LD		-0.0042 (0.0263)
Mixed × HD		-0.0898*** (0.0317)
p(Mixed X HD = Mixed X LD)		0.05
Mean Dep. Var.	3.93	3.93
Day Effects	Yes	Yes
Shift Effects	Yes	Yes
Line × Section Effects	Yes	Yes
Observations	1799	1799

* p<0.10, ** p<0.05, *** p<0.010. Observations are daily ratings received by line-section-level teams. Standard errors clustered at the line-section-level team. “Mixed" is a dummy variable coded 1 if the line-section-level team is religiously mixed. Line × Sections fixed effects are included in the all specifications; as a result the main effect of HD versus LD is not separately identified in columns (2) and (4). Education and tenure control for the mean of schooling and tenure of workers in the line-section-level team.

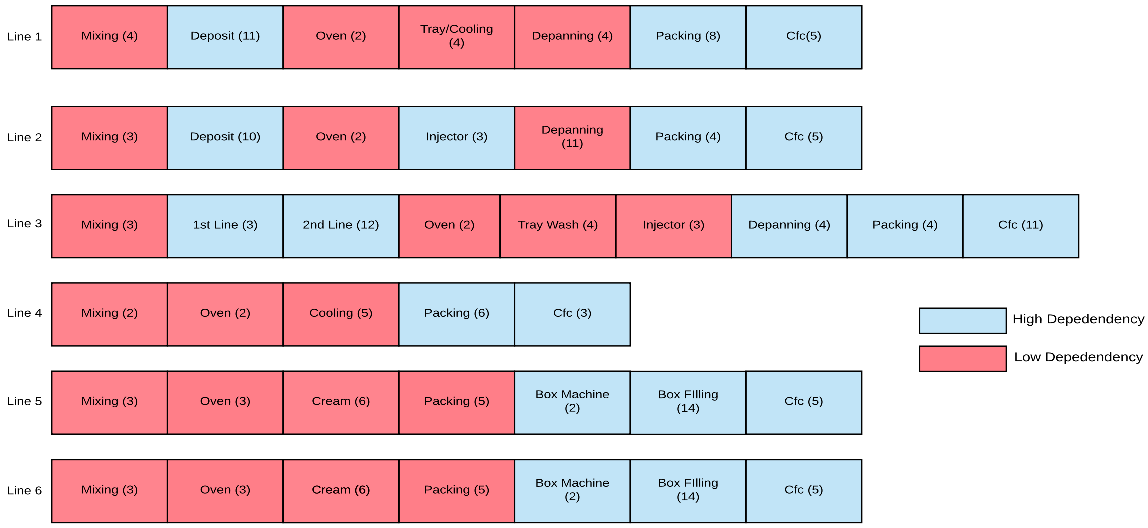
A.4 Additional figures referred to in the main text

Figure A.5: Structure of production lines



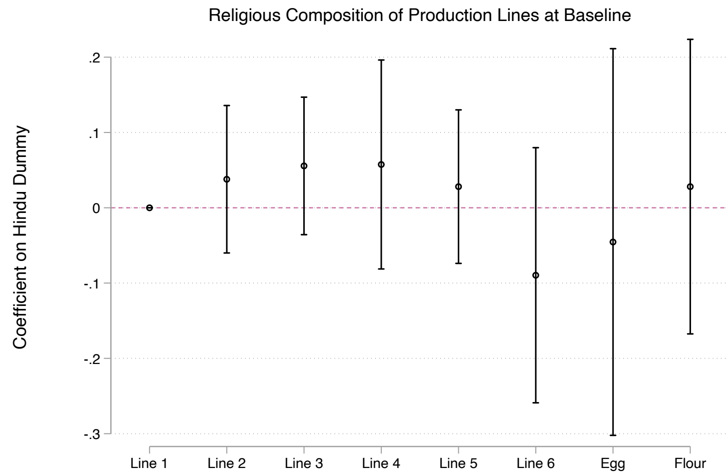
Note: This figure shows the structure of all six production lines in the factory. The numbers in parentheses denote the count of workers in each section per cohort. Each production line has three cohorts working on it in each of the three shifts in a day.

Figure A.6: High- and Low-Dependency sections

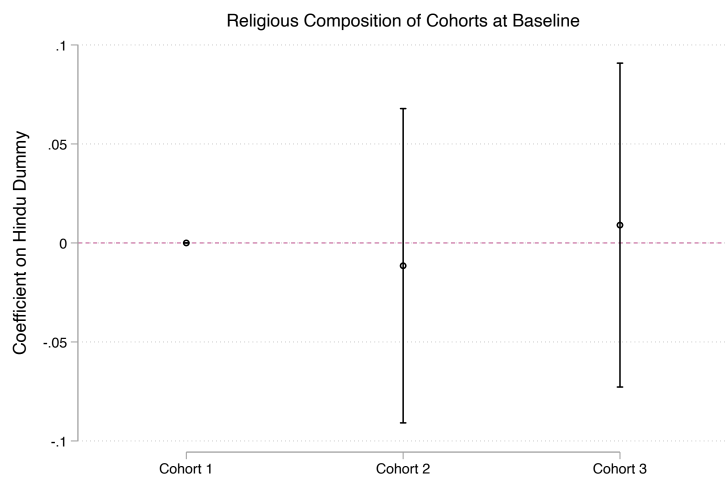


Note: This figure shows all sections of all lines at the factory split into HD and LD types. Direct Dependency is measured as described in section 2.2.3.

Figure A.7: Religious composition of lines and cohorts at baseline



(a) This figure plots coefficients from worker-level regressions. The outcome variable is a dummy coded 1 if the religion of the worker is Hindu and the independent variables are a set of dummy variables denoting each production line. “Egg" and “Flour" refer to production areas where raw materials (eggs and flour) are processed. These production areas are common to all production lines.



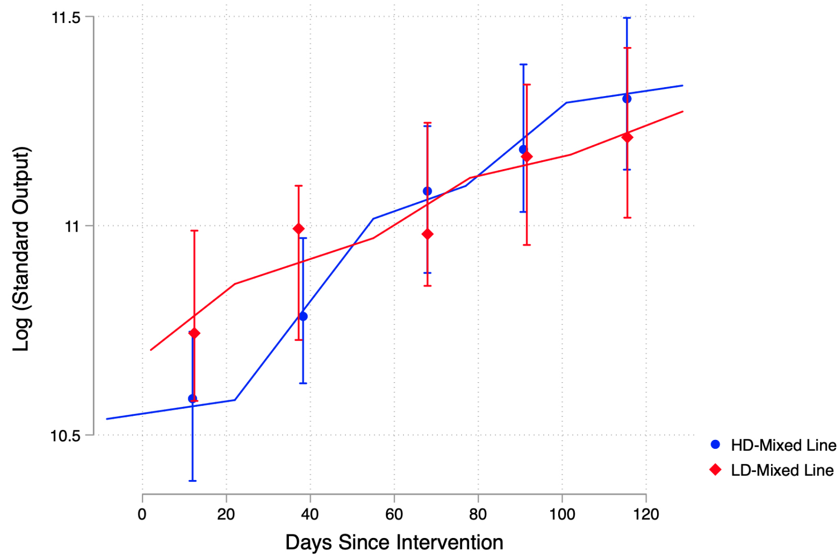
(b) This figure plots coefficients from worker-level regressions. The outcome variable is a dummy coded 1 if the religion of the worker is Hindu and the independent variables are dummies denoting cohorts (groups of workers who work at the factory at the same time).

Figure A.8: High- and Low-Dependency tasks

(a) High-Dependency (b) Low-Dependency

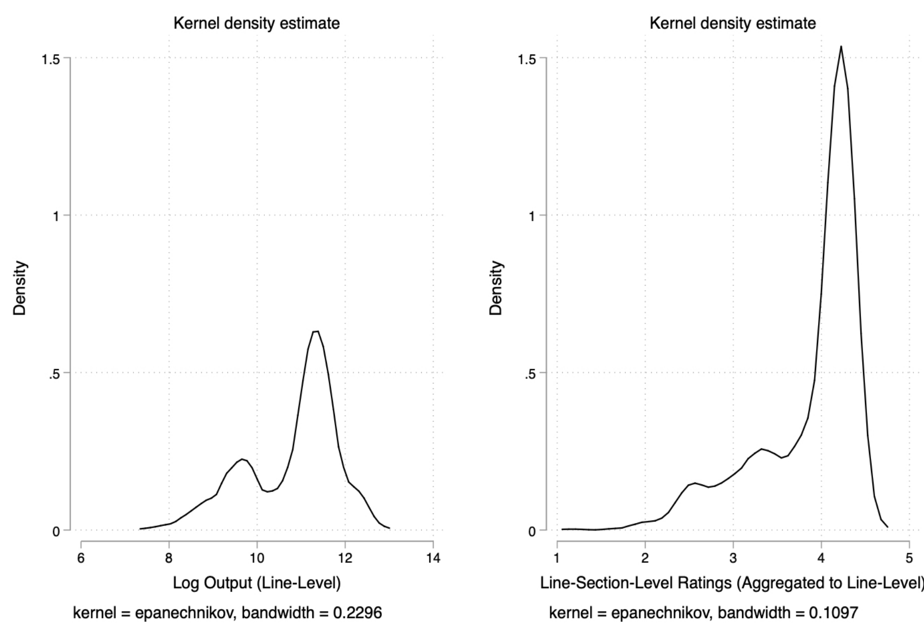
Note: This figure illustrates some key differences between HD and LD tasks. Sub-figure (a) is an example of a HD task. Workers are stood next to each other beside a fast moving conveyor belt. As a group they have to ensure each individual product is put into small packets before they go onto the next stage of production. If the team cannot coordinate and ensure the same, the supervisor has to reduce the speed of the belt to prevent wastage, which in turn would reduce output. Sub-figure (b) is a picture from a mixing room, which is a LD task. One worker is using the weighing scale to weigh raw materials, a second worker is arranging flour buckets and finally a third worker is operating the mixing machine. These workers have to coordinate as well to complete the process, but the frequency of interaction is intermittent and the degree of coordination is much lower relative to the HD task.

Figure A.9: Treatment effect on standard output (Event study)



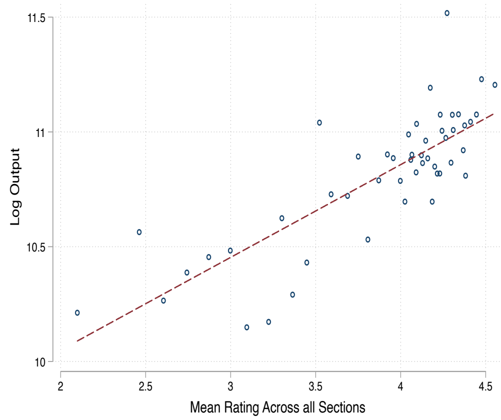
Note: This figure is generated from binned regressions using exactly the same controls variables as in Table 2.4. The treatment period is divided into 5 equal sized bins. The outcome variable is standard output (logged), which denotes the expected amount of output given inputs used.

Figure A.10: Distribution of actual line output and section ratings



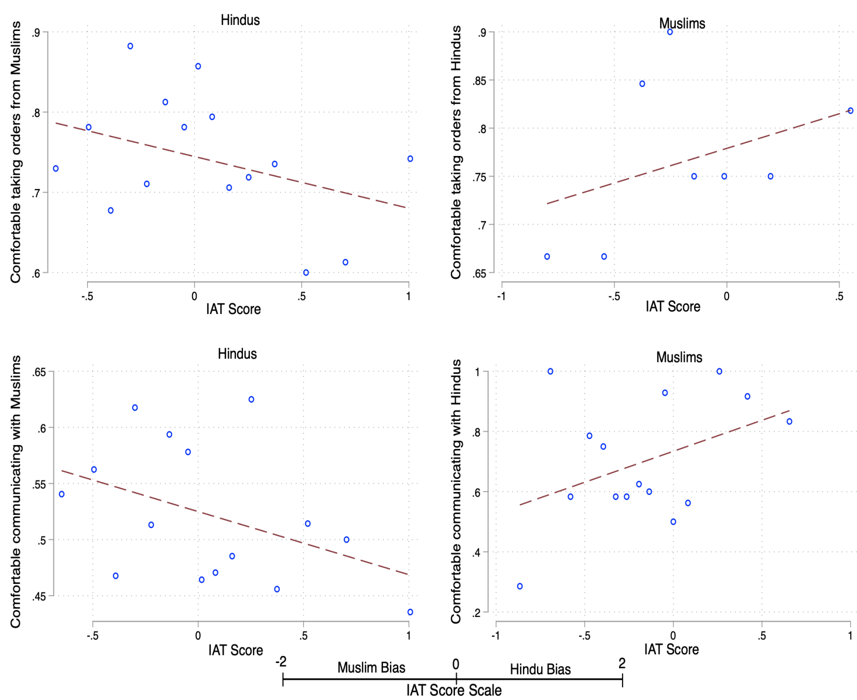
Note: This figure presents the distribution of raw ratings given by production supervisors to line-section-level teams aggregated up to the line-level as well as actual log output at the line-level.

Figure A.11: Line output and section ratings



Note: Production line fixed effects are included in this binscatter plot. The variable on the y-axis is daily output (logged) produced by a line-level team, and on the x-axis it is the average value of supervisor ratings received by sections in that line-level team.

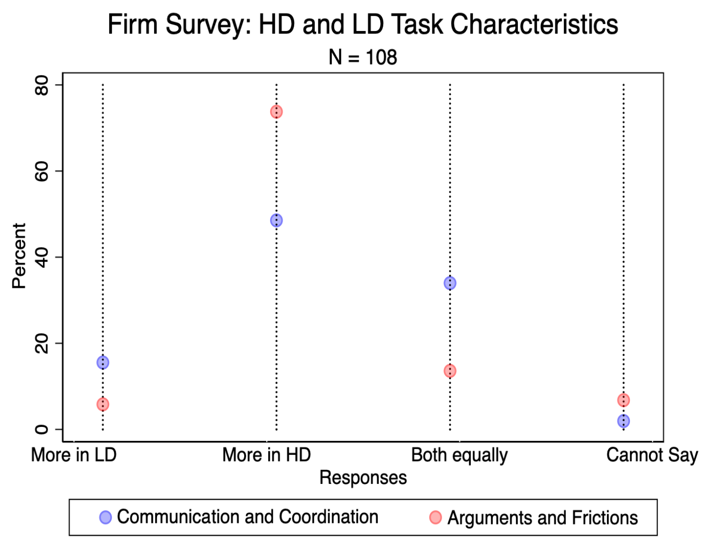
Figure A.12: Correlating IAT scores with survey responses



Note: This figure correlates Implicit Association Test scores (where workers were asked to associate Hindu and Muslim names with positions in the firm's hierarchy) with self-reported survey outcomes of the workers. In the top two figures, the outcome variable on the y-axis is willingness to take orders from non-coreligionists, while in the bottom two figures it is the workers' reported level of comfort in communicating with non-coreligionists. Positive IAT scores denote a bias towards having Hindus in higher positions while a negative value denotes a bias towards Muslims.

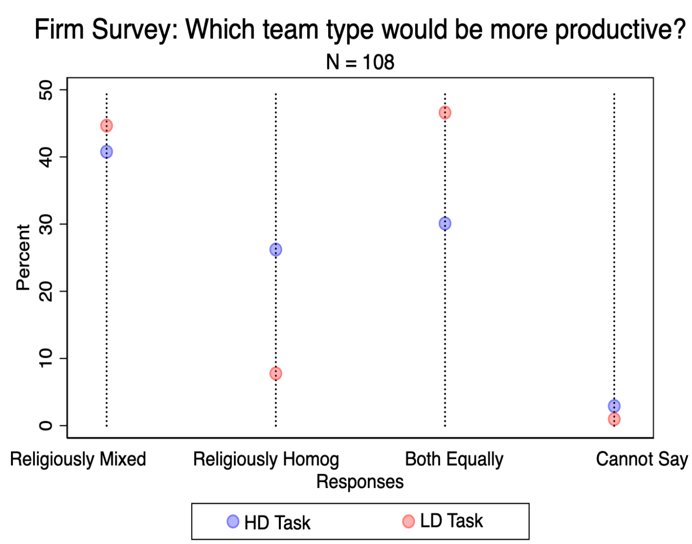
A.4.1 Figures from firm survey

Figure A.13: Characteristics of HD and LD tasks



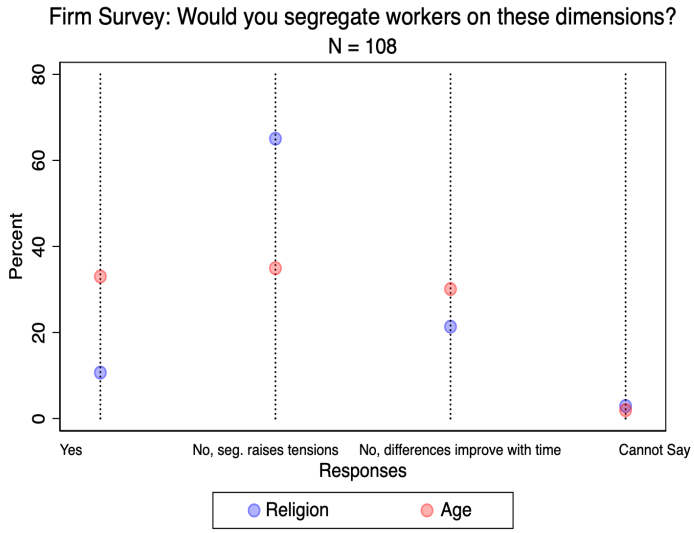
Note: This figure reports the percentage of respondents who picked each option for the following questions: (1) respondents were asked to pick the task that they thought requires greater continuous coordination and communication amongst workers (blue dots), and (2) they were to pick the one that is likely to cause more frictions and arguments amongst workers (pink dots).

Figure A.14: Religious mixing and productivity by task type



Note: This figure reports supervisors' perception of which type of team (religiously homogeneous or mixed) would be more productive in HD vs LD tasks. It reports percentage of respondents who picked each option when they were asked whether a religiously mixed or a homogeneous team would be more productive separately for HD (blue dots) and LD tasks (pink dots).

Figure A.15: Willingness to segregate workers by religion/age



Note: This figure presents responses of supervisors when asked if they are willing to segregate workers based on certain demographic dimensions. Percentage of respondents who chose each option for age and religion are denoted by pink dots and blue dots respectively.

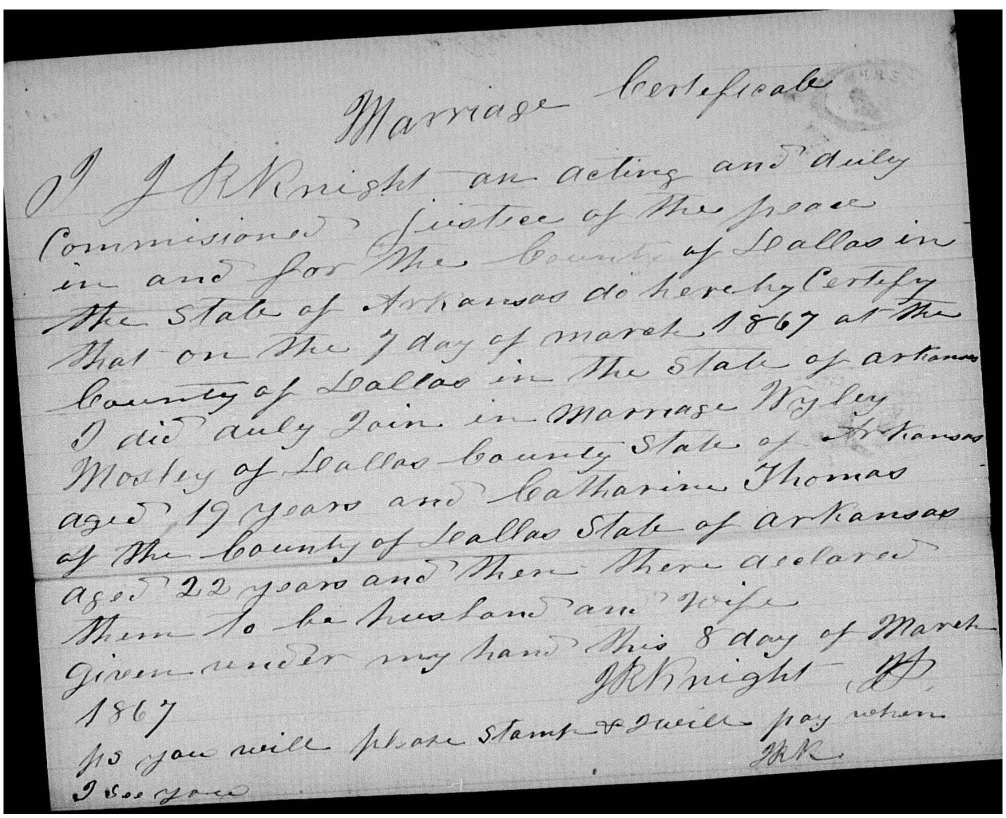
Appendix B

Appendix to Chapter 3

B.1 Marriage records

Our marriages dataset consists of marriage records, mostly from local governments or church parishes, which have been transcribed and made available online by Family- Search. We retrieved this data for all US states between 1800 and 1940. Most states that banned cousin marriage did so over this period. An example of a marriage certificate is shown in Figure B.1.

Figure B.1: Marriage certificate



Marriage Certificate
I J R Knight an acting and duly
Commissioned Justice of the peace
in and for the County of Dallas in
the State of Arkansas do hereby Certify
that on the 7 day of march 1867 at the
County of Dallas in the State of arkansas
I did duly Join in Marriage Wyley
Mosley of Dallas County State of Arkansas
aged 19 years and Catharine Thomas
of the County of Dallas State of arkansas
aged 22 years and then there declared
them to be husband and wife
Given under my hand this 8 day of March
1867 JRKnight JP
If you will please stamp & I will pay when
I see you JRK

Source: familysearch.com

The transcribed marriage records typically include the following information: names of both spouses, their age, and location and date of marriage. The transcribed records often contain duplicates. We considered observations to be duplicates if the names of both spouses, location and date of marriage were exactly the same and dropped these from our analysis. However, despite this, our dataset is likely to include some duplicate marriages due to misspelled names, as well as discrepancies in the date and location of marriages within duplicate observations. In many of these cases it is difficult to be certain that two records refer to the same marriage, making it difficult to systematically remove all duplicates. As an alternative conservative approach, we dropped multiple marriages where both spouses have the same name and the marriage took place in the same state and year, and re-ran our analysis. Our results are robust to this restriction. These results are available upon request.

B.1.1 Genealogical records

To complement our measure of cousin marriage using marriage records, we present an alternative dataset that includes a direct measure of kinship ties between spouses. This comes from anonymised, publicly available family trees from the Familinx database (Kaplanis et al., 2018), downloadable at `familinx.org`. This dataset provides exact ancestral links of individuals derived from family trees that have been created and managed by users of the `geni.com` website, who are mostly amateur genealogists researching their own family trees. The website allows independent trees to be merged and automatically suggests doing so when it detects sufficient overlap. Kaplanis et al. (2018) have cleaned this dataset and removed obvious errors (such as someone having three parents or being both the descendent and ancestor of another individual). The largest tree in this dataset, once cleaned, includes 13 million individuals.

This individual-level dataset includes the following variables for a large share of entries: gender, year and place of birth and death. However, we do not directly observe marriages. We instead simply assume that the two parents of any individual were married. Since the date and place of marriage are almost always absent, we proxy for these using the year and birthplace of the firstborn child of a given pair. This introduces some error in our measure of consanguinity over time and place, since we cannot measure the delay between marriage and the birth of the first child. It also means that we ignore childless marriages, and treat unmarried parents as having been married. Note that since we wish to focus on US marriages, we keep only couples where the first child was born in the US. Their ancestors, of course, may have been born abroad, which potentially allows us to identify cousin marriage among first or second generation immigrants.

A limitation of this dataset is that family trees are often incomplete – few individuals in the data have the full complement of eight great-grandparents. This is problematic since finding common ancestors of a husband and wife requires going back at least two generations in the case of first cousins (great-grandparents of the couple's child), and three generations for second cousins (great-great-grandparents). Figure B.2 presents summary statistics describing the nature and degree of non-missing links for individuals in the Familinx data. Each bar's horizontal length represents the number of people who have at least as many ancestors in the dataset as the bar's rank within its generation (vertically from top). For example, for grandparents, the top bar shows the number of individuals who have at least one grandparent link, the second bar shows the number of individuals with at least 2 grandparents, and so on. The black horizontal bars separate generations: grandparents, great grandparents and great great grandparents. Despite the large number of missing links, this genealogical data includes many ancestral links for a large number of individuals, which allows us to compare this data with isonymy rates from the marriage records, and also confirm the impact of bans on cousin marriage rates.

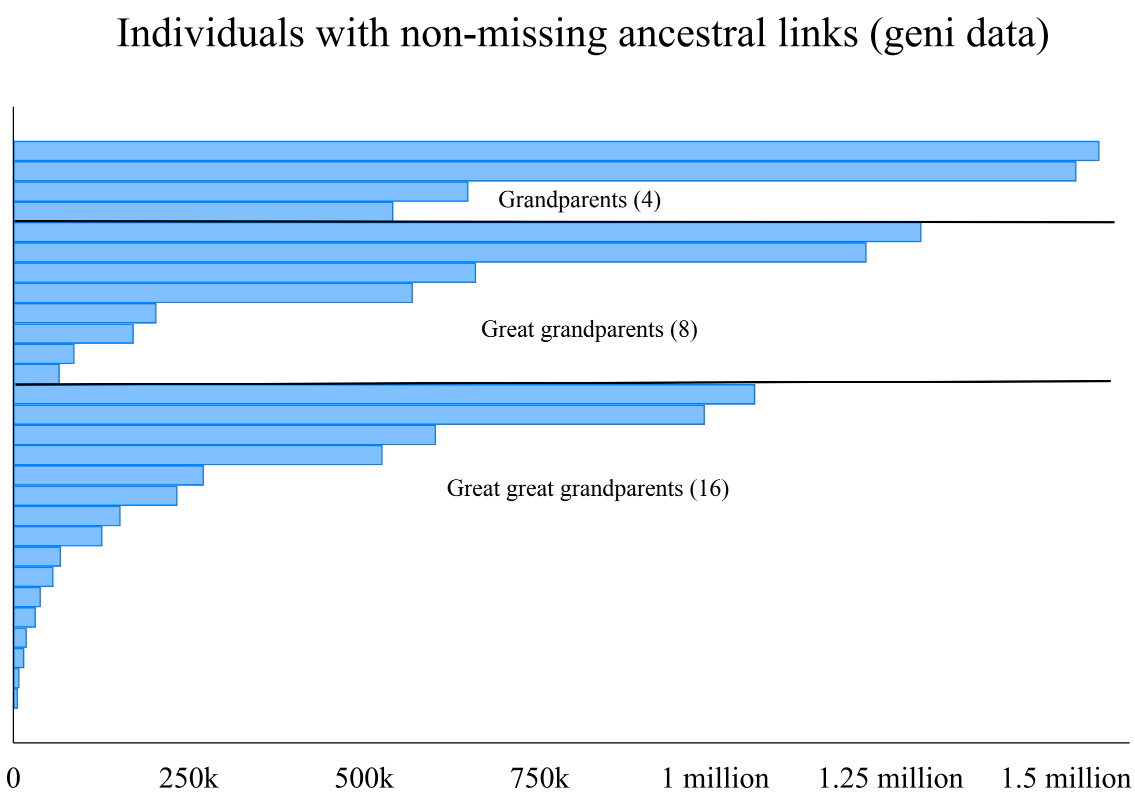
B.1.2 Types of first-cousin marriages, and implications for measures of isonymy

In this section, we discuss the type of first-cousin marriages that isonymy rates proxy for, the types it cannot capture and what this means for the type of effects identified in our main results. We also directly compare the distribution of cousin marriage rates in the 19th century with isonymy rates over the same period of time.

Recall, from section 2, that there are four different types of first-cousin marriage from a male's perspective:

1. Marrying father's brother's daughter (Type 1)
2. Marrying father's sister's daughter (Type 2)
3. Marrying mother's brother's daughter (Type 3)

Figure B.2: Individuals with non-missing ancestral links (genealogical data)



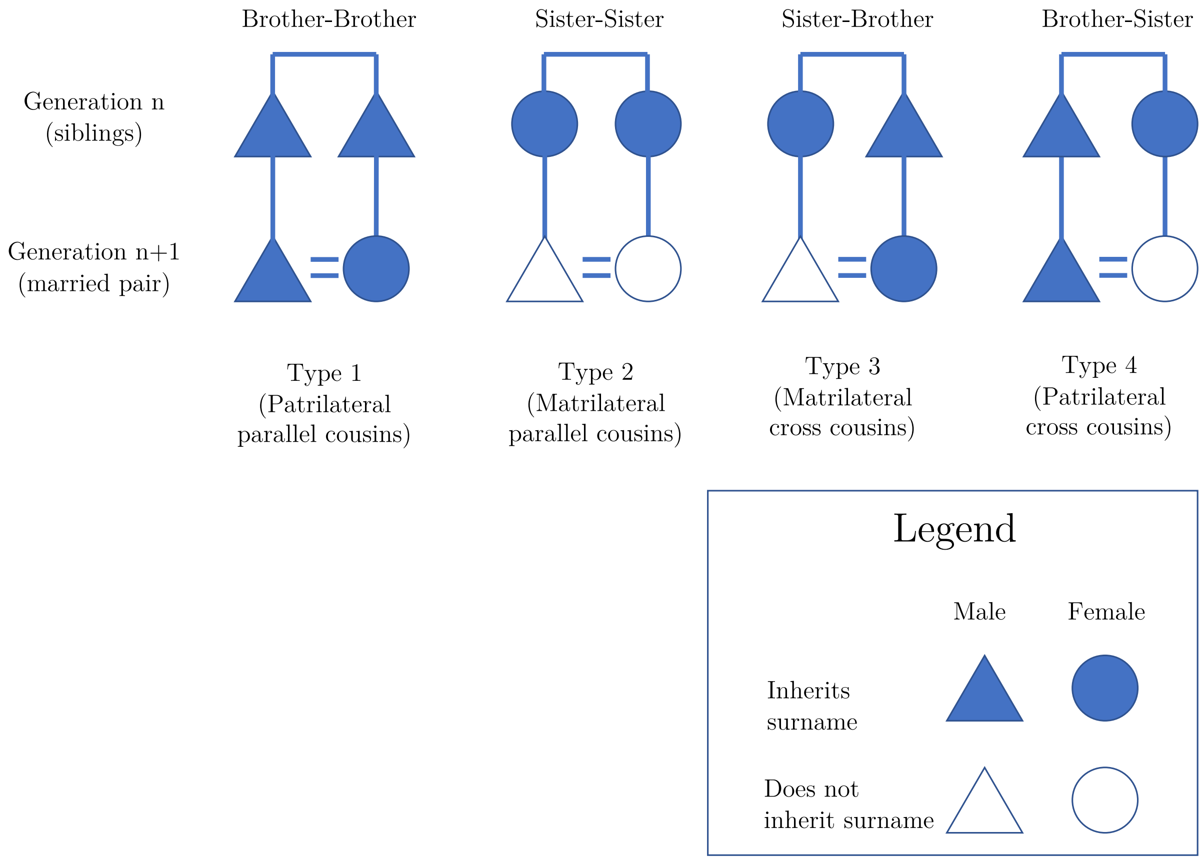
Note: This figure shows the number of observations (individuals) in the Familinx data who have a given number of ancestral links. For example, for grandparents, the first bar (vertically top) shows the number of individuals who have at least one grandparent link, the second bar shows the number of individuals with at least 2 grandparents so on and so forth. The pattern is similarly followed for great grandparents and great great grandparents.

4. Marrying mother's sister's daughter (Type 4)

In societies with patrilineal naming systems, surnames are inherited along the male line. This implies isonymy only captures the first type from the above list. This is because only in the first case will the female in the marriage be given the common family surname, resulting in an isonymous marriage. While each of the other cases are also first-cousin marriages, isonymy rates by construction will not be able to capture those. This is illustrated in figure B.3.

This implies some measurement error in our rate of cousin marriage. If all four types of first-cousin marriages are equally preferred, we would be capturing a quarter of the true cousin marriage rates, in expectation. Fortunately, the genealogical data, by allowing us to identify exact ancestral links, can speak to

Figure B.3: Cousin marriage and isonymy

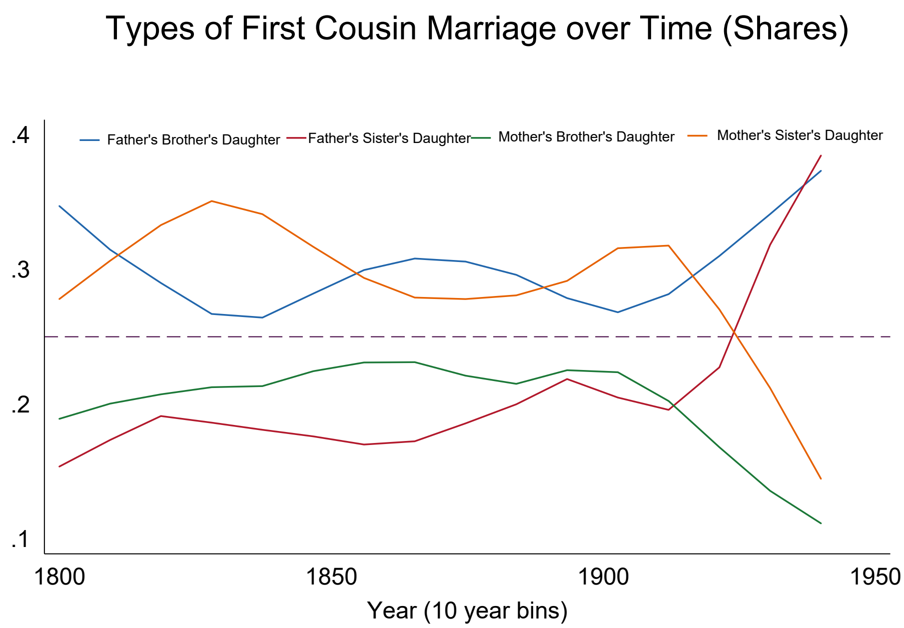


Note: This figure illustrates why only one of the four types of cousin marriage leads to isonymy. The solid shapes represent the offspring of someone from generation *n* who carries their surname (in a patrilateral society, where children inherit the surname of their father). Hollow shapes do not. Hence only marriage between the first type of cousins (the offspring of brothers) leads to isonymy.

this. We calculate rates of all four types of first-cousin marriage using Familinx data and plot their shares between 1800 and 1940 in Figure B.4. The type of first-cousin marriages that isonymy rates capture is denoted by the blue line in the figure—it hovers between 25% to 35% throughout the period. Overall, while there does seem to have been a slight preference for parallel cousin marriage, the share of each type is roughly constant over time. We further show in the next section that bans on cousin marriage did not affect the proportion of marriages that are isonymous.

We next compare the rate of isonymous marriages from marriage records to the rate of cousin marriages from Familinx. Table B.1 reports results from regressions at the state-year level—the outcome

Figure B.4: Types of first-cousin marriages



variable is the proportion of isonymous marriages in the marriage records. The Familinx-derived independent variables are, in column (1), the proportion of marriages with any overlap in family trees, and in column (2), the proportion of marriages between first cousins. We restrict the sample to state-year pairs in the period 1800-1940 with at least 100 marriages in both datasets. Both columns report a positive and statistically significant correlation between isonymy and cousin marriage rates.

Table B.1: Genealogical data and Isonymy

	Isonymous Marriage *Source: Marriage records* (1)	(2)
Consanguineous marriage	0.0421*	
Source: Genealogical records	(0.0224)	
First-cousin marriage		0.0563*
Source: Genealogical records		(0.0340)
N	2314	2314
Adj. R^2	0.256	0.256

* p<0.10, ** p<0.05, *** p<0.010. Robust standard errors in parentheses. Data on consanguineous marriages from genealogical records (Familinx). Isonymous marriages from marriage records. Observations are at the state-year level. Sample for state-year level observations is restricted to those years with at least 100 marriages both in the Familinx data and in the marriage records.

We now compare isonymy rates and Type 1 first-cousin marriage rates (as a percentage of all marriages in the U.S.) over time. If isonymy rates capture Type 1 first-cousin marriages we should expect these rates to evolve in a similar pattern over time. We indeed observe this in Figure B.5 throughout the 120 year period between 1800 to 1940.

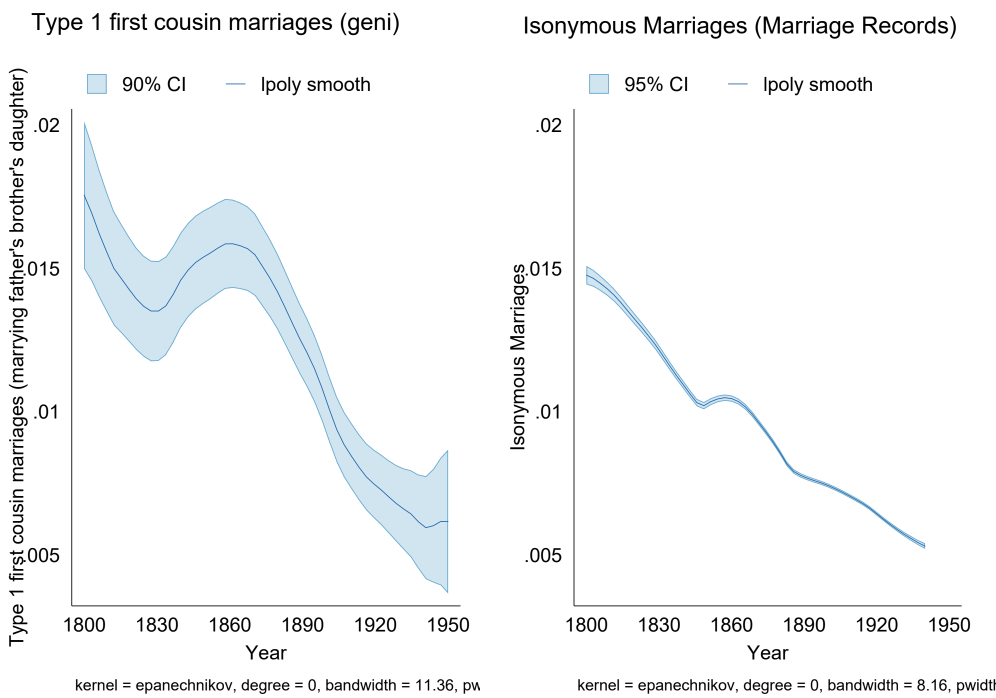
Both isonymy and Type 1 first-cousin marriage rates decreased gradually from 1800 to 1940 (from 1.5%-2% to about 0.6%). There is however an increase in both of these rates from around 1840 until about 1860, after which they continued to fall throughout the rest of the period.

To summarize, in this section we have specified the relationship between isonymy and first-cousin marriages. Then, using genealogical data, we inferred rates over time for the particular type of first-cousin marriage identified by isonymy. We have shown these rates are closely correlated with isonymy rates calculated from the marriage records data. These data patterns strengthen the case for using isonymy as a measure of consanguinity.

B.1.3 Cousin marriage bans: Evidence from genealogical records

In table B.2, we provide evidence that bans on cousin marriage led to the reduction in consanguineous unions, using genealogical data. Marriage-level observations are collapsed to create a state-year panel of

Figure B.5: Type 1 first-cousin marriages (Familinx) and isonymy (marriage records)



cousin marriage rates, where state refers to the couple's state of residence (inferred from the birth state of their first child). The main regressor is a dummy coded 1 if the state already had a ban in place in that year, 0 otherwise. We include state and year fixed effects, whereby the coefficients can be interpreted as difference-in-difference estimates. The coefficients show that cousin marriage rates fell by about half in states with a ban on cousin marriage, which is consistent with our main results.

Table B.2: Effect of bans on cousin marriage rates (genealogical records)

	Consanguineous marriages
State-year ban	-0.0385***
	(0.0101)
State F.E.	Yes
Year F.E.	Yes
N	1335
Adj. R^2	0.338
Mean Dep Var.	0.074

* p<0.10, ** p<0.05, *** p<0.010. Robust standard errors in parentheses. Data on consanguineous marriages from genealogical records (Familinx). Observations are at the state-year level. Any state-year with fewer than 25 marriages are dropped. Regressions include state and year fixed effects. The sample includes all individuals with at least one non-missing great-grandparent from each parent's side.

Using this genealogical dataset we also test a key assumption that underlies our results using marriage records: that bans on cousin marriage did not change the fraction of cousin marriages that are isonymous. Recall that our calculation of cousin marriage rates assumes that a quarter of cousin marriages are between the children of two brothers. These are the only such marriages that are isonymous (since women do not pass on their surname), and hence the only type we can observe through marital isonymy. However, this type of cousin marriage may be easier to observe, given the shared surnames, and therefore may be disproportionately affected by a ban on cousin marriage. This would lead us to overstate the effect of such a ban since our marriage records only allow us to measure this type of ban. The genealogical data allows us to test for this directly.

Specifically, table B.3 tests whether the ratio of isonymous (FBD, or "Father's brother's daughter") cousin marriages relative to the other three types changes post-ban. We find no evidence for a differential effect on this type of cousin marriage, which suggests that our ability to infer overall rates of cousin marriage using isonymy is consistent pre- and post-ban.

Table B.3: DID Regressions: Impact of bans on ratio of cousin marriage types

	(1)	(2)	(3)	(4)
	$\frac{FBD}{FSD}$	$\frac{FBD}{MBD}$	$\frac{FBD}{MSD}$	$\frac{FBD}{FSD+MSD+MBD}$
State-year ban	-0.00345	0.00261	-0.0132	-0.0187
	(0.00543)	(0.00201)	(0.0111)	(0.0124)
State F.E.	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes
N	5429	5429	5429	5429
Adj. R^2	0.026	0.006	0.052	0.057

* p<0.10, ** p<0.05, *** p<0.010. This table estimates the impact of bans on first-cousin marriages on the frequency type-1 first-cousin marriages (FBD marriages) as a proportion of other types (FSD, MBD and MSD). Standard errors are clustered at the state-level in parentheses. All regressions include state and year fixed effects. The ratio in each column is coded to be 0 if neither type of marriage is recorded in the state in that year.

B.1.4 Census variable definitions

In this section, we provide definitions of the main census outcome variables, from the US census, as obtained from IPUMS. We also describe how (in some cases) we construct other variables from these. The census rounds from which we use a particular variable are mentioned below its description. The variable names in uppercase letters refer to the IPUMS labels, and the ones in italics are those used in the paper.

Income

1. *Log Occupational Income* — OCCSCORE is a constructed variable that assigns each occupation an income score. It assigns each occupation in all years a value that represents median total income (in hundreds of 1950 dollars) of all persons with that particular occupation in 1950. We use log of this variable as our outcome. We drop observations with missing values from our analysis.

   Census Rounds: 1850, 1870, 1900, 1920, 1930, 1940

2. *Log Wage Income* — INCWAGE is a respondent's total pre-tax wage and income from salary - in other words it is money received as an employee in the previous year. The following are included in INCWAGE: wages, salaries, commissions, cash bonuses, tips, and any other money income re-

ceived from an employer. In-kind payments or reimbursements for business expenses are not included in this. We use log of this variable as outcome dropping observations with missing values.
Census Rounds: 1940

Schooling

1. *Grade* — HIGRADE is a continuous variable denoting the highest grade of school attended or completed by the respondent. The general code for this variable denotes the highest grade that has been completed. We use this variable to denote the highest level of schooling attained by an individual.
   Census Rounds: 1940

2. *In School* — SCHOOL is a binary variable that indicates whether a person attended school during a specified time period in the past. The period varies across censuses, but is typically within 3 months to a year from when the person was being surveyed.
   Census Rounds: 1850, 1870, 1900 (5%), 1920, 1940

Genetic Effects

1. *Living in Hospital, Mental Institution, or Home for Physically Handicapped* — This is a dummy variable constructed from the GQTYPE variable, which denotes the type of group-quarter within which the respondent resided in. Most respondents reside in private households, rather than group quarters. The variable takes the value 1 if the respondent resided in a hospital, in a mental institution or in homes, hospitals or schools for the physically handicapped.
   Census Rounds: 1930, 1940

Urbanization

1. *Urban* — URBAN indicates whether the location of residence of a household is urban or rural. It is widely used but it does have some problems. Definitions of "urban" vary slightly from year to year - but it typically denotes all cities and incorporated places which have more than 2500 inhabitants.
   Census Rounds: 1850, 1870, 1900, 1920, 1930, 1940

2. *Farm Household* — FARM identifies all farm households. Based on the census round such households are identified either by the occupation of the household members or whether their house was located on a farm. All group quarters are coded as non-farm. We use a dummy variable coded 1 to indicate a farm household.
   Census Rounds: 1850, 1870, 1900, 1920, 1940

3. *Log Population Size* — SIZEPL reports the population of a municipality in bins of different sizes. Unidentified locations are grouped as "Under 1,000 or unincorporated". We use the log of the mid-point of each bin as our outcome, Log Population Size.
   Census Rounds: 1850, 1870, 1900, 1920, 1930, 1940

Labor Supply

1. *Weeks worked* — WKSWORK1 in the IPUMS data is the total number of weeks that a respondent worked during the previous year for profit, pay, including unpaid work.
   Census Rounds: 1940

2. *Weeks Worked > 0* — A dummy variable which takes a value of 1 if WKSWORK1 > 0
   Census Rounds: 1940

B.1.5 Predictors of state bans on cousin marriage

This section explores differences across states in the timing of bans on first-cousin marriage. We do not argue that the timing of these bans was random. This is in part why we rely on variation within states, comparing surnames with high versus low rates of cousin marriage in the pre-period. However it may still be that predictors of earlier bans on cousin marriage have differential effects on high versus low cousin marriage families.

Table B.4 presents differences between early and late ban states for a range of factors which historians have pointed to as possible drivers of these bans on cousin marriage. We categorize states that banned cousin marriage pre-1902 as 'early' in order to equalize the number of early versus late ban states. Results in Table B.4 do not suggest major differences between these states, though the small number of observations means we have limited power to detect such differences.

Table B.4: Early versus late bans and state characteristics

Variable	Early ban (Pre-1902) (1)	Late ban (1902-onward) (2)	Diff (2) - (1)	N
Cousin marriage rate (1800-1850)	0.05	0.05	-0.001	24
	(0.00)	(0.00)	(0.001)	
Year of Union	1851.12	1847.12	-4.00	32
	(9.39)	(9.78)	(13.564)	
Year of Compulsory Schooling	1887.31	1892.88	5.563	32
	(3.53)	(3.35)	(4.869)	
Share Foreign Born	0.08	0.13	0.051	19
	(0.02)	(0.03)	(0.043)	
Proportion urban	0.11	0.08	-0.034	18
	(0.03)	(0.02)	(0.037)	
Literacy Rate	0.86	0.89	0.026	18
	(0.03)	(0.02)	(0.038)	
Baptist churches per 10k ppl	3.98	3.46	-0.517	16
	(0.60)	(0.88)	(1.134)	
Methodist churches per 10k ppl	5.38	5.82	0.618	16
	(0.92)	(0.96)	(1.362)	
Presbyterian churches per 10k ppl	1.96	2.02	0.070	16
	(0.44)	(0.31)	(0.523)	
Roman Catholic churches per 10k ppl	0.63	0.73	0.097	16
	(0.13)	(0.20)	(0.256)	

Significance levels: * <10% ** <5% *** <1%. Standard errors in parentheses. "Year of Union" denotes the year in which the state became a part of the Union. "Year of Compulsory Schooling" denotes the year in which compulsory schooling laws were passed in the state. Cousin marriage rates are calculated from marriage records as described in section 3.2.2. Proportion urban, share foreign born and literacy rates are measured from the 1850 Census. The number of churches is measured from the 1870 Census. The number of churches are per 10,000 inhabitants.

Another potential threat to identification could be differential pre-trends in cousin marriage rates between states that banned cousin marriage early versus those that banned it late. In other words, it might be the case that states that banned cousin marriage early are the ones where it was declining faster anyway. To rule this out, Table B.5 uses a state-year panel of 1800-1858 isonymy rates to show that changes over time in rates of cousin marriage are not correlated with the (future) timing of cousin marriage bans. Specifically, we regress state-year rates of observed isonymy on the year (as a continuous variable) interacted with the timing of a cousin marriage ban in that state. In absence of pre-trends we should see a null effect on this interaction term, which is what we find.

Table B.5: Pre-trends in isonymy rates

	(1) Observed isonymy rates
Years Banned by 1940 (Fraction)	-0.00006
× Year (1800-1858)	(0.0003)
State F.E.	Yes
Year F.E.	Yes
N	1239
Adj. R^2	0.092

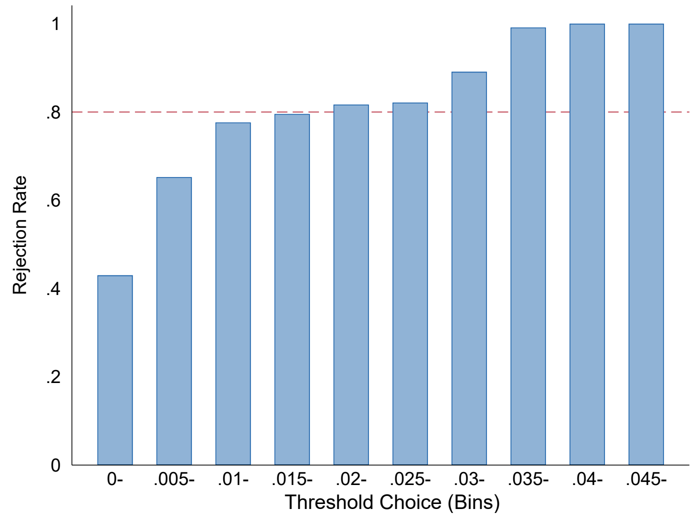
* p<0.10, ** p<0.05, *** p<0.010. State-year level observations. A marriage is considered to be isonymous if both individuals involved in the marriage share the same surname. Marriages are collapsed at the state-year level to obtain isonymy rates. Standard errors clustered at the state level in parentheses.

B.1.6 Non-random isonymy censoring

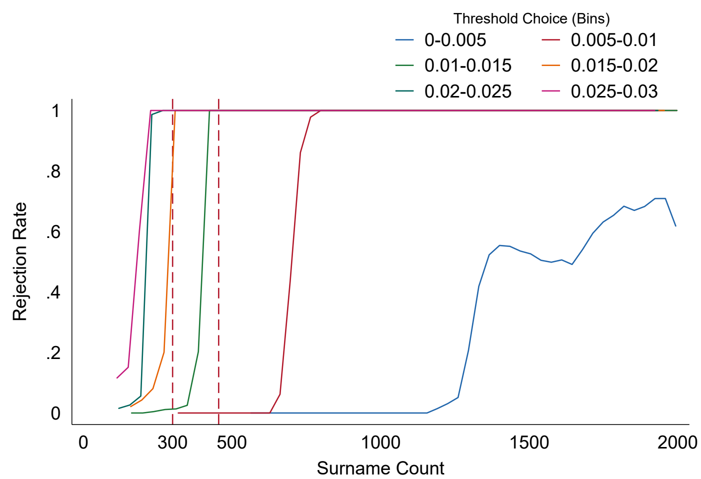
Throughout our analysis, we use log values of non-random isonymy as our measure of cousin marriage. Before taking the log, we replace values of non-random isonymy below ϵ with ϵ. We do this for two reasons. The simplest is that non-random isonymy takes a negative value whenever isonymous marriages are less frequent than predicted under random mating. Replacing these values with zero would still not allow us to convert them into logs, hence a positive value. The second reason is that most surname-state cells have rates of non-random isonymy near zero, and we cannot statistically distinguish these rates from zero (or each other). Since small level changes near zero are large proportional changes, values of ϵ that are too small result in most of the variation in the log of non-random isonymy being driven by measurement error. To attenuate this concern, we choose a value of ϵ sufficiently large that rates of non-random isonymy above this threshold are typically distinguishable from zero.

We use a threshold of $\epsilon = 0.015$ in our main analysis. In this section we justify our threshold choice and show robustness to other values of ϵ. First, we analyze how the total number of individuals with a surname (sample size) and its isonymy rate, in a marriage pool, affect the ability to reject zero non-random isonymy for that group. This motivates our threshold choice. To do this, we first test for each surname whether its rate of non-random isonymy in the pre-period can be distinguished from zero. We split non-random isonymy values into a range of threshold choice bins and present the average rejection probability (rate) across all surnames in each bin, in Figure B.6a. Clearly, for low threshold values, one needs a very large sample size to reject zero non-random isonymy, whereby the rejection rate is low. For bins with higher values, the rejection rate increases. A value of $\epsilon = 0.015$ is the lower bound of the bin 0.015-0.02. As can be observed in Figure B.6a, this gives a rejection rate around 80%.

Figure B.6: Surname Frequency and Zero NR Isonymy Rejection



(a) Threshold Choice and Rejection Rate



(b) Surname Count and Rejection Rate

Note: In Figure B.6a non-random isonymy values are split into a range of threshold choice bins and average rejection rates across all surnames in each bin is presented. In Figure B.6b, we show how rejection rates vary, given a threshold choice bin, as the number of records for a surname increases. Each line represents a separate threshold bin.

In Figure B.6b, we show how rejection rates vary, given a threshold choice bin, as the number of records for a surname increases. With lower threshold values a larger number of records are required to reject zero isonymy. On average, around 300 records are required for surnames with non-random isonymy in the range 0.015-0.02, to achieve a 80% rejection probability. The median surname count in our data is 455. The choice of 0.015 as the threshold therefore leads to a rejection rate of greater than 80% for more than half of the surnames in our sample.

We test robustness of our results to alternative choices of thresholds in Table B.8 in the Appendix. We re-run our analysis using various values of ϵ. As expected, low values of ϵ attenuate the coefficients, consistent with our interpretation that small differences in non-random isonymy between surnames with low rates of cousin marriage are mostly measurement error.

B.2 Supplementary tables and figures

Table B.6: Impact of bans on cousin marriage rates (Year Bins)

	(1) Log (CousinMarr^{post})
Treatment intensity (pre-1940 bans)	
Year banned (1-16) × Log (CousinMarr^{pre})	–
Year banned (17-33) × Log (CousinMarr^{pre})	-0.141
	(0.0907)
Year banned (34-50) × Log (CousinMarr^{pre})	-0.212***
	(0.0820)
Year banned (51-67) × Log (CousinMarr^{pre})	-0.245***
	(0.0782)
Year banned (68-85) × Log (CousinMarr^{pre})	-0.287***
	(0.0773)
N	1,207,523
Adj. R^2	0.450

* p<0.10, ** p<0.05, *** p<0.010. Standard errors clustered at the origin state (father's birth state)-surname level in parentheses. All regressions include age effects, origin-state, residence state, and surname fixed effects. Log (CousinMarr^{pre}) and Log (CousinMarr^{pre}) are constructed as described in section 3.2.2. Cousin marriage rates are calculated at the state level in the pre-period and surname-state level in the post period and linked at the surname-state (of father's birth) level with the 1940 Census. Sample includes White males aged 18 to 50 in 1940 unless otherwise specified.

Table B.7: Cousin marriage rates in levels

	(1)	(2)	(3)	(4)
	Log Occup. Income	**Highest Grade**	**Weeks Worked**	**Log Urbanization**
			Females	
Years Ban (Fraction) × CousinMarr^{pre}	0.150*	0.839*	5.534**	1.004**
	(0.0890)	(0.478)	(2.718)	(0.439)
Mean Dep. Var.	2.87	10.17	14.95	8.28
Surname F.E.	. Yes	Yes	Yes	Yes
State of Origin F.E.	Yes	Yes	Yes	Yes
State of Residence F.E.	Yes	Yes	Yes	Yes
N	976,414	1,187,725	942,605	1,207,523
Adj. R^2	0.122	0.120	0.055	0.132

* p<0.10, ** p<0.05, *** p<0.010. Standard errors clustered at the surname-origin state (father's birth state) level in parentheses. Cousin marriage rates are calculated (as described in section 3.2.2) at the surname-level in the pre-period and surname-state level in the post period and linked at the surname-origin state (of father's birth) level with the 1940 Census records. Years Ban (Fraction) refers to the number of years each state had a ban on cousin marriage in place by the year 1940, divided by 82, which is the number of years that had passed since the first ban in Kansas in 1858. Sample includes White males aged 18 to 50 in 1940 unless otherwise specified.

Table B.8: Robustness to threshold choice

	ϵ = **0.005**	ϵ = **0.010**	ϵ = **0.015**	ϵ = **0.020**
	(1)	(2)	(3)	(4)
		Log Occupational Income		
Years Ban (Fraction) × Log(CousinMarr^{pre})	0.00788	0.0148*	0.0392***	0.0522**
	(0.00533)	(0.00875)	(0.0140)	(0.0218)
Mean Dep. Var.	2.87	2.87	2.87	2.87
N	976,414	976,414	976,414	976,414
Adj. R^2	0.122	0.122	0.122	0.122
	Log population size of locality of residence (Urbanization)			
Years Ban (Fraction) × Log(CousinMarr^{pre})	0.0391	0.109***	0.303***	0.492***
	(0.0243)	(0.0421)	(0.0733)	(0.122)
Mean Dep. Var.	8.29	8.29	8.29	8.29
N	1,207,523	1,207,523	1,207,523	1207523
Adj. R^2	0.132	0.132	0.132	0.132
Surname F.E.	Yes	Yes	Yes	Yes
State of Origin F.E.	Yes	Yes	Yes	Yes
State of Residence F.E.	Yes	Yes	Yes	Yes

* p<0.10, ** p<0.05, *** p<0.010. Standard errors clustered at the surname-origin state (father's birth state) level in parentheses. Cousin marriage rates are calculated (as described in section 3.2.2) at the surname-level in the pre-period and surname-state level in the post period and linked at the surname-origin state (of father's birth) level with the 1940 Census records. Years Ban (Fraction) refers to the number of years each state had a ban on cousin marriage in place by the year 1940, divided by 82, which is the number of years that had passed since the first ban in Kansas in 1858. Sample includes White males aged 18 to 50 in 1940 unless otherwise specified.

Table B.9: Dropping top and bottom 5% of common surnames

	Log Occupational Income (1940)		**Highest Grade Completed (1940)**		**Weeks Worked** *Females*		**Log (Urbanization)**	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years Ban (Fraction) × Log (CousinMarr^{pre})	0.0354**	0.0301*	0.223***	0.178**	1.005**	0.906*	0.298***	0.231***
	(0.0142)	(0.0157)	(0.0776)	(0.0905)	(0.4377)	(0.4777)	(0.0730)	(0.0774)
Surname F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State of Origin F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State of Residence F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Surname-state pre-treatment controls	No	Yes	No	Yes	No	Yes	No	Yes
Mean Dep. Var.	2.88	2.88	10.17	10.10	14.98	15.04	8.29	8.30
N	924,512	837,719	1,124,534	1,013,575	893,306	808,537	1,143,195	1,028,614
Adj. R^2	0.122	0.122	0.120	0.120	0.055	0.101	0.132	0.135

* p<0.10, ** p<0.05, *** p<0.010. Standard errors clustered at the surname-origin state (father's birth state) level in parentheses. Cousin marriage rates are calculated (as described in section 3.2.2) at the surname-level in the pre-period and surname-state level in the post period and linked at the surname-origin state (of father's birth) level with the 1940 Census records. Years Ban (Fraction) refers to the number of years each state had a ban on cousin marriage in place by the year 1940, divided by 82, which is the number of years that had passed since the first ban in Kansas in 1858. Surname-state pre-treatment controls include occupational income (logged) and population of locality of residence (logged) measured at the surname-state level from the 1850 census records. Sample includes White males aged 18 to 50 in 1940 unless otherwise specified.

Table B.10: Including all states (including states that never banned)

	Log Occupational Income (1940)		**Highest Grade Completed (1940)**		**Weeks Worked** *Females*		**Log (Urbanization)**	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years Ban (Fraction) × Log (CousinMarr^{pre})	0.0270***	0.0306***	0.0465	0.0733	0.505*	0.625*	0.136**	0.163***
	(0.00930)	(0.0116)	(0.0494)	(0.0656)	(0.2907)	(0.3462)	(0.0565)	(0.0624)
Surname F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State of Origin F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State of Residence F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Surname-state pre-treatment controls	No	Yes	No	Yes	No	Yes	No	Yes
Mean Dep. Var.	2.88	2.87	9.97	9.85	15.47	14.84	8.34	8.18
N	1,665,794	1,208,303	2,018,176	1,460,251	1,658,089	1,256,009	2,057,490	1,486,457
Adj. R^2	0.129	0.124	0.138	0.138	0.0629	0.0576	0.189	0.156

* p<0.10, ** p<0.05, *** p<0.010. Standard errors clustered at the surname-origin state (father's birth state) level in parentheses. Cousin marriage rates are calculated (as described in section 3.2.2) at the surname-level in the pre-period and surname-state level in the post period and linked at the surname-origin state (of father's birth) level with the 1940 Census records. Years Ban (Fraction) refers to the number of years each state had a ban on cousin marriage in place by the year 1940, divided by 82, which is the number of years that had passed since the first ban in Kansas in 1858. Surname-state pre-treatment controls include occupational income (logged) and population of locality of residence (logged) measured at the surname-state level from the 1850 census records. Sample includes White males aged 18 to 50 in 1940 unless otherwise specified.

Table B.11: 1930 Outcomes

	Log Occupational Income		**Log (Urbanization)**	
	(1)	(2)	(3)	(4)
Years Ban (Fraction) × Log (CousinMarr^{pre})	0.0230***	0.0173**	0.0852**	0.0379
	(0.00732)	(0.00781)	(0.0463)	(0.0481)
Surname-state pre-treatment controls				
Log Occupational Income (1850)		0.0541***		0.259***
		(0.00766)		(0.0444)
Log Urbanization (1850)		0.0263***		0.310***
		(0.00181)		(0.0137)
Mean Dep. Var.	3.03	3.02	8.59	8.59
Surname F.E.	Yes	Yes	Yes	Yes
State of Origin F.E.	Yes	Yes	Yes	Yes
State of Residence F.E.	Yes	Yes	Yes	Yes
N	3,161,458	2,989,111	4,021,199	3,790,663
Adj. R^2	0.133	0.131	0.145	0.144

* p<0.10, ** p<0.05, *** p<0.010. Standard errors clustered at the surname-origin state (father's birth state) level in parentheses. Cousin marriage rates are calculated (as described in section 3.2.2) at the surname-level in the pre-period and surname-state level in the post period and linked at the surname-origin state (of father's birth) level with the 1930 Census records. Years Ban (Fraction) refers to the number of years each state had a ban on cousin marriage in place by the year 1940, divided by 82, which is the number of years that had passed since the first ban in Kansas in 1858. Sample includes White males aged 18 to 50 in 1930.

Table B.12: Robustness to years of compulsory schooling

	Log Occupational Income (1940)		**Highest Grade Completed (1940)**		**Weeks Worked** *Females*		**Log (Urbanization)**	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years Ban (Fraction) × Log(CousinMarr^{pre})	0.0466***	0.0408***	0.270***	0.215**	0.904**	0.934*	0.307***	0.249***
	(0.0144)	(0.0156)	(0.0832)	(0.0942)	(0.452)	(0.494)	(0.0717)	(0.0764)
Years Compulsory Schooling × Log(CousinMarr^{pre})	-0.000733*	-0.00104**	-0.00235	-0.00214	0.0127	0.00177	-0.000355	-0.00145
	(0.000377)	(0.000410)	(0.00186)	(0.00213)	(0.01123)	(0.01262)	(0.00191)	(0.00221)
Surname F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State of Origin F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State of Residence F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Surname-state pre-treatment controls	No	Yes	No	Yes	No	Yes	No	Yes
Mean Dep. Var.	2.87	2.88	10.17	10.10	14.95	14.95	8.28	8.29
N	976,414	885,324	1,187,725	1,071,133	942,605	853,731	1,207,523	1,087,102
Adj. R^2	0.122	0.122	0.120	0.119	0.0549	0.0545	0.132	0.134

* p<0.10, ** p<0.05, *** p<0.010. Standard errors clustered at the surname-origin state (father's birth state) level in parentheses. Cousin marriage rates are calculated (as described in section 3.2.2) at the surname-level in the pre-period and surname-state level in the post period and linked at the surname-origin state (of father's birth) level with the 1940 Census records. Years Ban (Fraction) refers to the number of years each state had a ban on cousin marriage in place by the year 1940, divided by 82, which is the number of years that had passed since the first ban in Kansas in 1858. Surname-state pre-treatment controls include occupational income (logged) and population of locality of residence (logged) measured at the surname-state level from the 1850 census records. Sample includes White males aged 18 to 50 in 1940 unless otherwise specified.

Table B.13: Robustness to years of statehood

	Log Occupational Income (1940)		Highest Grade Completed (1940)		Weeks Worked *Females*		Log (Urbanization)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years Ban (Fraction) × Log(CousinMarr^{pre})	0.0402***	0.0345**	0.241***	0.189**	0.9762**	0.8516*	0.340***	0.262***
	(0.0144)	(0.0159)	(0.0793)	(0.0917)	(0.4531)	(0.5002)	(0.0790)	(0.0842)
Years Statehood × Log(CousinMarr^{pre})	0.0000776	0.0000532	-0.000346	-0.000792	-0.00598	-0.00742	0.00287**	0.00166
	(0.000175)	(0.000223)	(0.000856)	(0.00113)	(0.00537)	(0.00709)	(0.00119)	(0.00124)
Surname F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State of Origin F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State of Residence F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Surname-state pre-treatment controls	No	Yes	No	Yes	No	Yes	No	Yes
Mean Dep. Var.	2.87	2.88	10.17	10.10	14.95	14.95	8.28	8.29
N	976,414	885,324	1,187,725	1,071,133	942,605	853,731	1,207,523	1,087,102
Adj. R^2	0.122	0.122	0.120	0.119	0.050	0.101	0.132	0.134

* p<0.10, ** p<0.05, *** p<0.010. Standard errors clustered at the surname-origin state (father's birth state) level in parentheses. Cousin marriage rates are calculated (as described in section 3.2.2) at the surname-level in the pre-period and surname-state level in the post period and linked at the surname-origin state (of father's birth) level with the 1940 Census records. Years Ban (Fraction) refers to the number of years each state had a ban on cousin marriage in place by the year 1940, divided by 82, which is the number of years that had passed since the first ban in Kansas in 1858. Surname-state pre-treatment controls include occupational income (logged) and population of locality of residence (logged) measured at the surname-state level from the 1850 census records. Sample includes White males aged 18 to 50 in 1940 unless otherwise specified.

Table B.14: Robustness to percent native population

	Log Occupational Income (1940)		**Highest Grade Completed (1940)**		**Weeks Worked** *Females*		**Log (Urbanization)**	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years Ban (Fraction) × Log(CousinMarr^{pre})	0.0415***	0.0339**	0.247***	0.200**	0.938**	0.946*	0.295***	0.242***
	(0.0151)	(0.0155)	(0.0876)	(0.0921)	(0.4760)	(0.4890)	(0.0824)	(0.0779)
Percent Native Born (1850)× Log(CousinMarr^{pre})	0.000659	0.000370	0.000721	-0.00164	0.00163	-0.01312	0.0108***	0.00989***
	(0.000604)	(0.000631)	(0.00305)	(0.00341)	(0.00165)	(0.01817)	(0.00351)	(0.00361)
Surname F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State of Origin F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State of Residence F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Surname-state pre-treatment controls	No	Yes	No	Yes	No	Yes	No	Yes
Mean Dep. Var.	2.87	2.88	10.11	10.10	15.06	15.00	8.29	8.29
N	914,220	885,324	1,106,857	1,071,133	882,015	853,731	1,123,479	1,087,102
Adj. R^2	0.123	0.122	0.119	0.119	0.099	0.101	0.132	0.134

* p<0.10, ** p<0.05, *** p<0.010. Standard errors clustered at the surname-origin state (father's birth state) level in parentheses. Cousin marriage rates are calculated (as described in section 3.2.2) at the surname-level in the pre-period and surname-state level in the post period and linked at the surname-origin state (of father's birth) level with the 1940 Census records. Years Ban (Fraction) refers to the number of years each state had a ban on cousin marriage in place by the year 1940, divided by 82, which is the number of years that had passed since the first ban in Kansas in 1858. Surname-state pre-treatment controls include occupational income (logged) and population of locality of residence (logged) measured at the surname-state level from the 1850 census records. Sample includes White males aged 18 to 50 in 1940 unless otherwise specified.

Figure B.7: Enforcement of cousin marriage bans in the news

UNWITTING OFFENDERS.—The Governor of Arkansas has pardoned Alexander Pruitt and Mary Tinsley of Green county. They were first cousins, and not knowing that the State law forbade marriages between first cousins, were united in matrimony near Gainesville about eighteen months ago. They were arrested, tried and convicted in the Green County Circuit Court last week, the wife appearing in the court-room with her little baby. The sentence was three years each, the minimum punishment. The Judge, prosecuting attorney, jury, and many citizens, united in the petition for pardon.—*New Orleans Picayune, March 23d.*

(a) Arkansas 1884

First Cousins Marry in Indiana.

Special Dispatch to the Globe-Democrat.

TERRE HAUTE, IND., February 25.—A novel case is attracting the attention of attorneys here, and it is likely to be brought into the courts before a settlement is reached. One day this week Howard Holmes, aged 17, and Ella Tucker, 22, were married, and it was discovered soon after that they were first cousins, and under the State law such a marriage is illegal. They have been stopping with Holmes' father in the country near here, and yesterday the bride's friends took her from her husband and refused to allow her to return. Holmes was in the city to-day consulting with lawyers, and will probably bring suit to get his wife back, as she desires to live with him. The law provides for a fine of $500 for a minister or justice who solemnizes a marriage between first cousins, but fixes no penalty for the parties to such a marriage. The marriage in this case was performed by a country justice, who claims to have had no knowledge of the relationship.

(b) Indiana 1887

Appendix C

Appendix to Chapter 4

C.1 Tables and figures

Table C.1: Election cycle and fatal mining accidents (State-Level) (OLS)

	(1) **Log Cases**	(2) **Log Injuries**	(3) **Log Deaths**
Four Years	-0.0495	0.236	0.234**
	(0.116)	(0.156)	(0.136)
	[0.671]	[0.133]	[0.04]
Three Years	0.0231	0.206	0.173
	(0.177)	(0.157)	(0.137)
	[0.903]	[0.207]	[0.252]
Two Years	-0.0962	0.0155	0.0246
	(0.138)	(0.0909)	(0.0717)
	[0.478]	[0.855]	[0.711]
One Year	-0.247**	-0.237*	-0.181
	(0.111)	(0.123)	(0.115)
	[0.036]	[0.058]	[0.102]
State F.E.	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes
State Time Trends	Yes	Yes	Yes
N	350	350	350
R^2	0.898	0.550	0.495

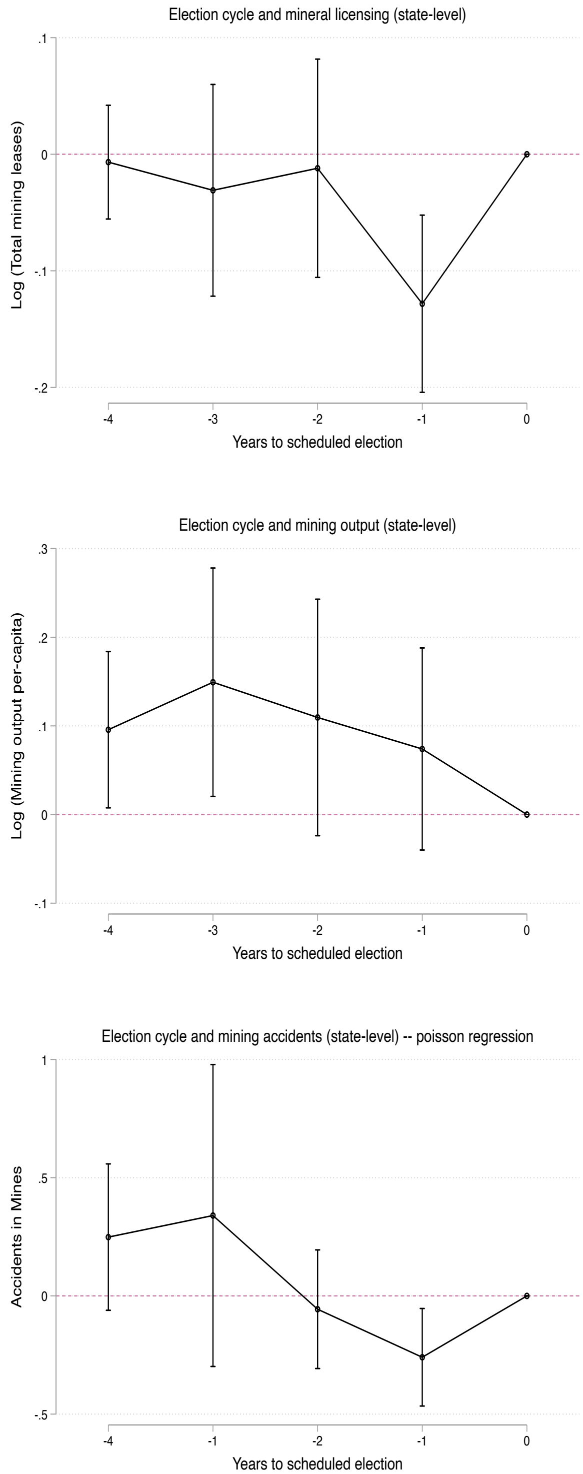
Notes : * p<0.10, ** p<0.05, *** p<0.010. State-Year Level observations between 1998-2015. Dependent variables are the natural logarithm of the number of events in a year in a given state. Standard Errors clustered at the State Level in parenthesis. P-values from Wild cluster Bootstrap with 1000 replications presented in square [] brackets. All regressions include state effects, year effects and state time trends. Outcome variables are subject to a log $(x+1)$ transformation.

Table C.2: Election cycle and mining fatalities (district-level)

	Any Bad Event	Accidents		Deaths	
	(1)	(2)	(3)	(4)	(5)
Election_{-1}	-0.0899	-0.132*	-0.0262	-0.167**	-0.321*
	(0.0607)	(0.0667)	(0.0859)	(0.0687)	(0.169)
Election	-0.176***	-0.151***	0.0467	-0.0814*	-0.083
	(0.0610)	(0.0523)	(0.0502)	(0.0448)	(0.130)
Election_{+1}	0.0779	0.121***	0.325***	0.161***	0.420***
	(0.0691)	(0.0397)	(0.0715)	(0.0491)	(0.149)
District F.E.	Yes	Yes	Yes	Yes	Yes
Region-Year F.E.	Yes	Yes	Yes	Yes	Yes
N	412	412	412	412	412
$R^2/Log-Likelihood$	0.409	0.857	-543	0.623	-382
Estimation	OLS	OLS	Poisson	OLS	Poisson

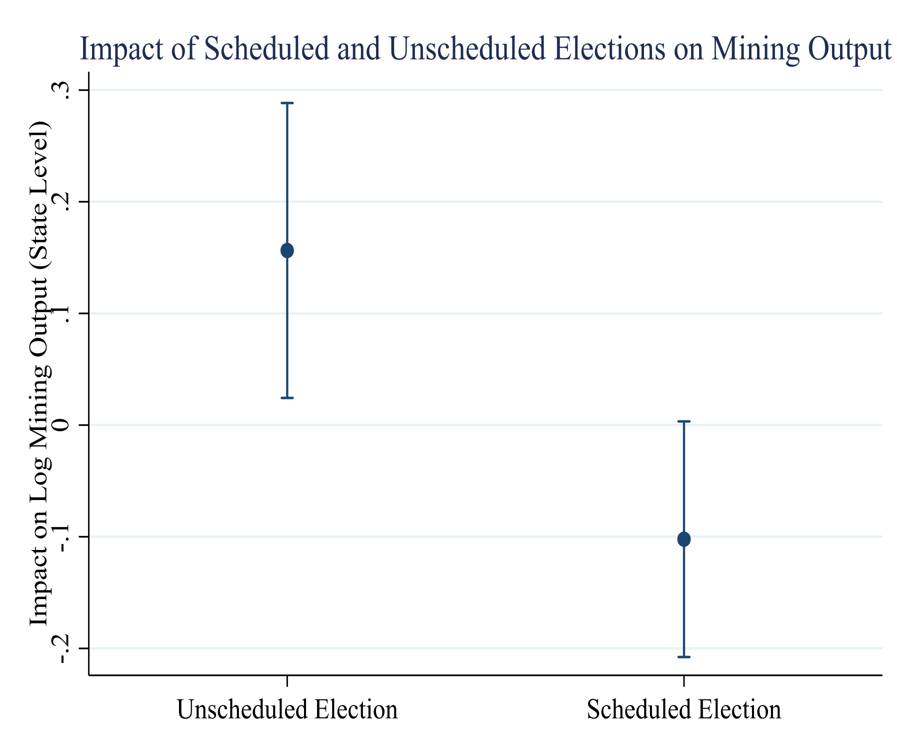
Notes : * p<0.10, ** p<0.05, *** p<0.010. District-Year level observations covering 104 districts across 15 states between 2010-2015. Standard errors clustered at the state-year level reported in parenthesis. Standard deviations for outcome means are reported in parenthesis. “Any Bad Event" is a dummy variable coded 1 if there was at least one serious or fatal accident in a mining field. “Accidents" refer to the sum of both fatal and serious accident and “Deaths" refer to the total number of peopled killed in mining fields in a district in a given year. Election_{-1}, Election and Election_{+1} are dummies for the year immediately before, the year of and the year immediately after a scheduled election respectively. All regressions include district and region-year effects. Outcomes Deaths and Accidents for OLS regressions are subject to a log $(x+1)$ transformation.

Figure C.1: Election cycle and mining intensity (state-level coefficient plots)



Note: This figure presents coefficient plots from results in Tables 4.2, 4.3 and 4.4. The first plot (licensing) is based on Column (1) of Table 4.2, the second (output) is based on Column (1) of Table 4.3 and the third (fatal accidents) is based on Column(2) of Table 4.4.

Figure C.2: Scheduled vs unscheduled elections and mining output – placebo test



Note: This figure shows the impact of scheduled and unscheduled elections on mining output from the state-level analysis. The y-axis plots coefficient values from a regression of log mining output on a election dummy and an interaction between the election dummy and a dummy denoting whether the election was scheduled or not. Unscheduled elections are therefore all elections that occurred before the end of the 5 year political term.

This figure plots coefficient estimates from the following regression,

$$Y_{st} = \alpha_s + \beta_1 Election_{st} + \beta_2 Election_{st} \times Scheduled_{st} + \delta_t + \epsilon_{st} \tag{C.1}$$

where Y_{st} is log mining output in state s at year t, α_s is a state fixed effect and δ_t is a year fixed effect and $Scheduled_{st}$ denotes whether the election in state 's' is scheduled or not. The effect of an unscheduled election is therefore given by β_1 and that of a scheduled election by $\beta_1 + \beta_2$. Unscheduled elections include all elections that took place before the end of a 5 year political term. The coefficient estimates are relative to the average of all non-election years.

Table C.3: Mining cycles, literacy rates and Scheduled Tribe population

	Log Accidents			**Log Casualties**		
	(1)	(2)	(3)	(4)	(5)	(6)
Election	-0.176***	-0.126**	-0.588**	-0.174***	-0.120	-0.767***
	(0.0499)	(0.0624)	(0.234)	(0.059)	(0.074)	(0.233)
Election × Proportion ST		-0.370			-0.500	
		(0.340)			(0.490)	
Election × Proportion Literate			0.720*			1.00***
			(0.380)			(0.370)
District F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Region-Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes
N	605	605	605	605	605	605
R^2	0.878	0.878	0.879	0.696	0.697	0.698

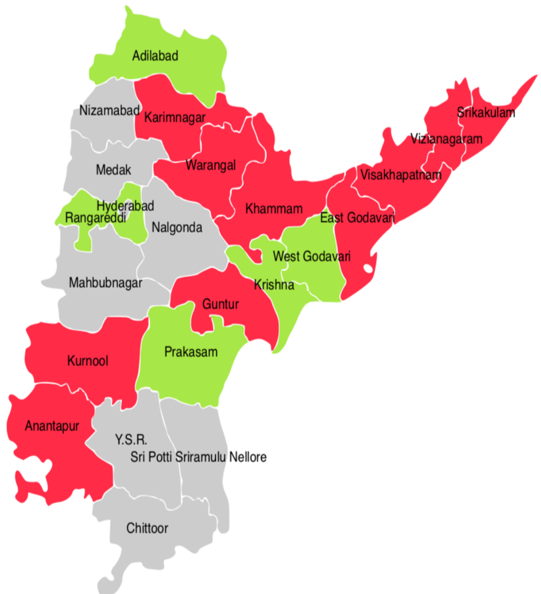
* p<0.10, ** p<0.05, *** p<0.010. District-Year Level observations covering 104 districts between 2010-2015. Each column represents a separate regression. Standard errors clustered at the state-year level reported in parenthesis. All regressions include District Fixed Effects and Region-Year Fixed Effects. Log Accidents refer to the natural logarithm of all mining accidents in a district in a given year. Log Casualties refer to the natural logarithm of the sum of deaths and injuries in mining fields in a district in a given year. Percentage of Scheduled Tribe in each district have been obtained from the 2011 Census of India. Outcomes Log Accidents and Log Casualties are subject to a log $(x+1)$ transformation.

Table C.4: Correlates of mining fatalities and Naxalite conflict

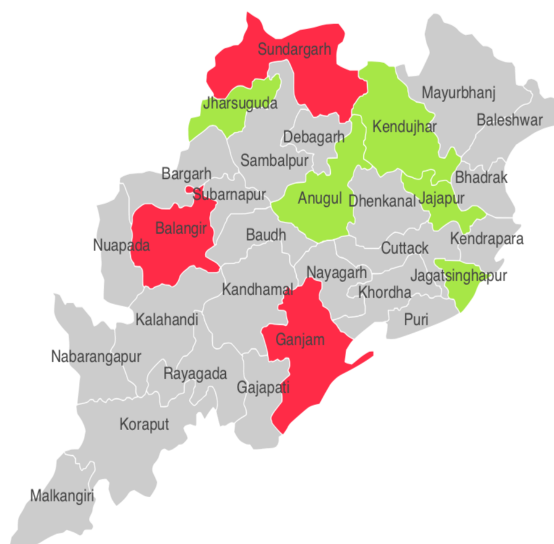
	Onset		**Log Rebel Deaths**		**Log Security Forces Deaths**	
	(1)	(2)	(3)	(4)	(5)	(6)
Log Accidents	0.0790***	0.0217	0.121***	-0.0113	0.033	-0.0575
	(0.0288)	(0.0393)	(0.0446)	(0.0710)	(0.0344)	(0.0593)
Log Accidents × Proportion ST		0.280*		0.622**		0.430*
		(0.151)		(0.309)		(0.260)
Log Accidents × Proportion ST × Election		-0.231**		-0.221**		-0.143**
		(0.095)		(0.095)		(0.065)
District F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Region-Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes
N	684	684	684	684	684	684
R^2	0.624	0.632	0.475	0.489	0.475	0.682

* p<0.10, ** p<0.05, *** p<0.010. District-Year Level observations between 2010-2015 covering 114 districts. Standard errors clustered at the state-year level reported in parenthesis. All regressions include District Fixed Effects and Region-Year Fixed Effects. Log Accidents refer to the natural logarithm of all mining accidents in a district in a given year. Percentage of Scheduled Tribe in each district have been obtained from the 2011 Census of India. “Onset" is a dummy variable coded 1 if at least one death occurred from a clash between state forces and rebels. Outcomes Log Rebel and Log Security forces deaths are subject to a log $(x+1)$ transformation.

Figure C.3: Red Corridor in Andhra Pradesh and Orissa



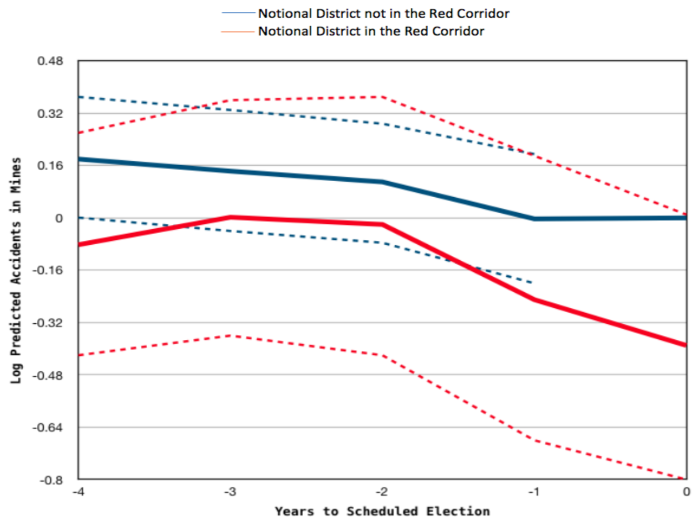
(a) Andhra Pradesh (before split)



(b) Orissa

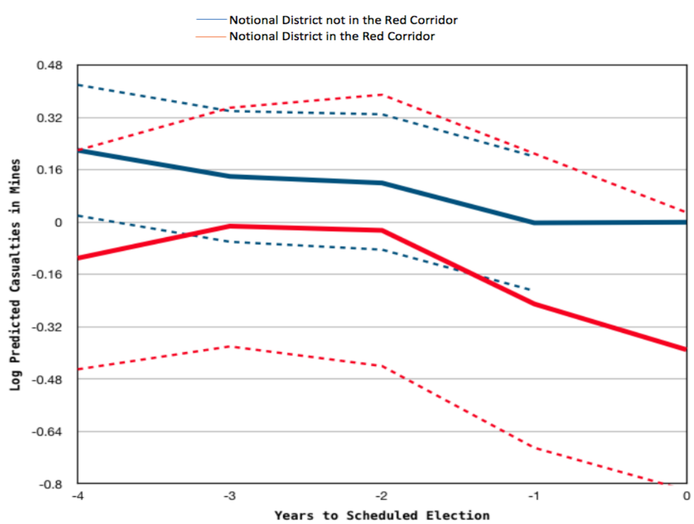
Note: This figure marks all districts in the states of Andhra Pradesh and Orissa that form a part of the sample for mining fatalities data. Within each state Naxal affected districts as classified by the Union Ministry of Home Affairs in 2008 have been marked in red and the rest in green. This figure illustrates that when studying mining cycles, within-state variation can be used to identify the overall effect of being in the Red Corridor on mining intensity.

Figure C.4: Fatality cycles and Red Corridor (state fixed effects)



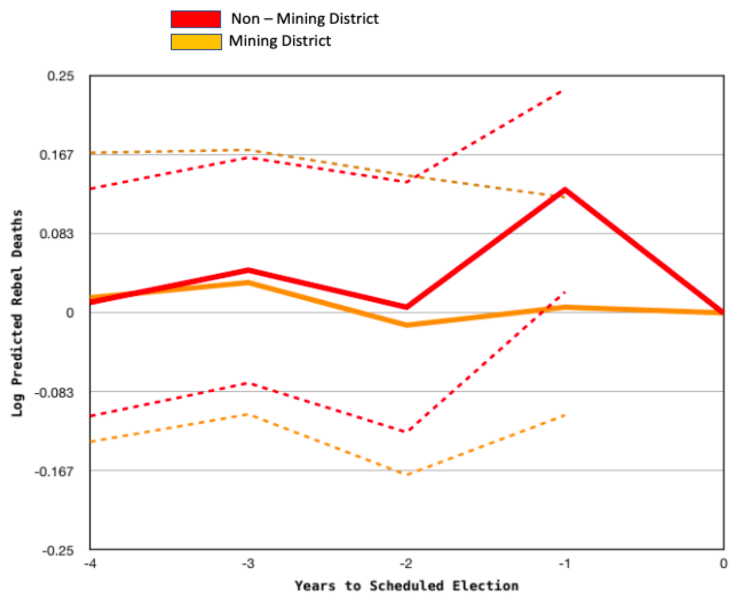
Note: This figure plots log predicted accidents over the electoral cycle comparing districts in the Red Corridor (red) and those not in the Red Corridor (blue) with state-fixed Effects . I use pre-sample classification of districts into the Red Corridor whereby it is a fixed district characteristic. Since there is variation within states with respect to districts that are classified as part of the Red Corridor and those that are not, the main effect of being in the Red Corridor is captured here. This figure is based on results from Column (2) of Table 4.7. Dotted lines show 95% confidence intervals.

Figure C.5: Fatality cycles and Red Corridor (state fixed effects)



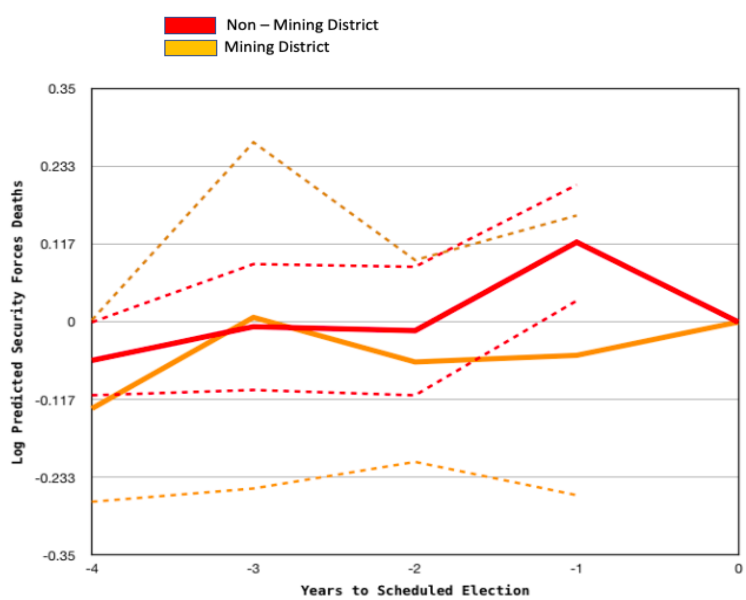
Note: This figure plots log predicted casualties over the electoral cycle comparing districts in the Red Corridor (red) and those not in the Red Corridor (blue) with state-fixed Effects. I use pre-sample classification of districts into the Red Corridor whereby it is a fixed district characteristic. Since there is variation within states with respect to districts that are classified as part of the Red Corridor and those that are not, the main effect of being in the Red Corridor is captured here. This figure is based on results from Column (5) of Table 4.7. Dotted lines show 95% confidence intervals.

Figure C.6: Naxal rebel deaths over the election cycle



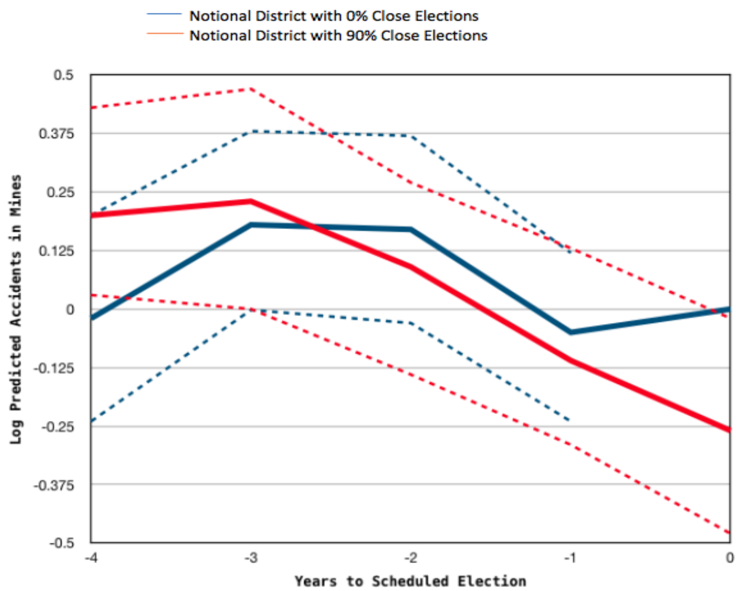
Note: This diagram plots log predicted naxal rebel deaths from clashes between naxal rebels and security forces over the electoral cycle in mining and non-mining districts. The value in the election year is normalized to 0 in each case. It is based on specifications of the form in Table 4.8 Columns (3), (6) and (9) but from a OLS model with log rebel deaths as the outcome variable. The dotted lines represent 95% confidence intervals.

Figure C.7: Security forces deaths over the election cycle



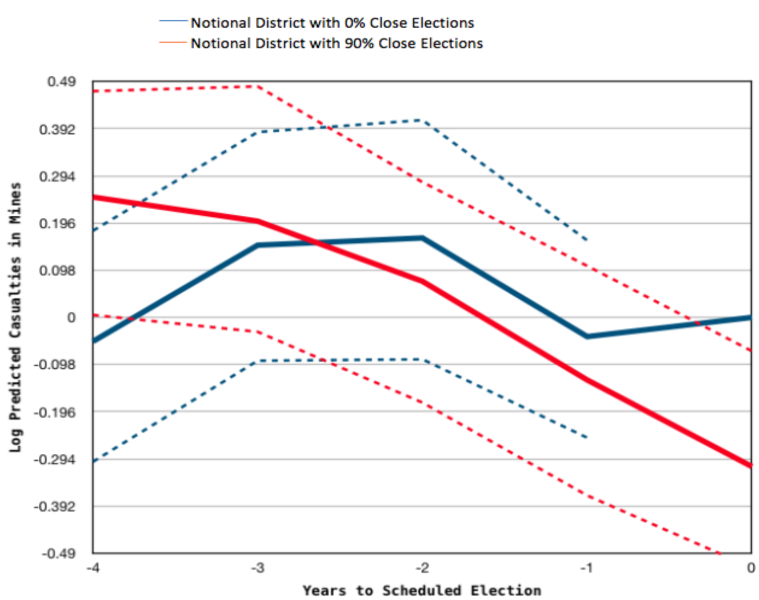
Note: This diagram plots log predicted security forces deaths from clashes between naxal rebels and security forces over the electoral cycle in mining and non-mining districts. The value in the election year is normalized to 0 in each case. It is based on specifications of the form in Table 4.8 Columns (3), (6) and (9) but from a OLS model with log security forces deaths as the outcome variable. The dotted lines represent 95% confidence intervals.

Figure C.8: Accident cycles and electoral competition



Note: This diagram plots log predicted accidents in mines over the next electoral cycle for notional districts with 0% and 90% close elections in the previous elections respectively. It is based on Column (3) of Table 4.5. Note that the value in election year is 0 by construction for a district with no close elections. Dotted lines represent 95% confidence intervals.

Figure C.9: Fatality cycles and electoral competition



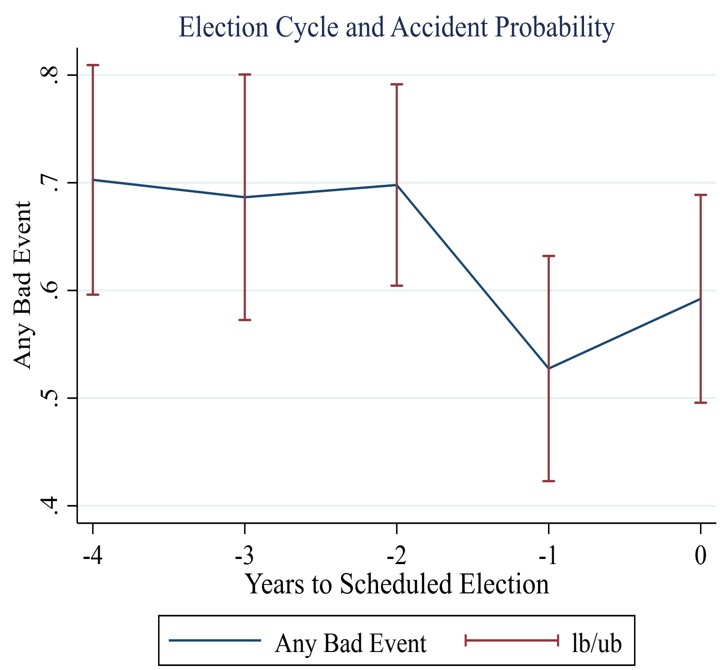
Note: This diagram plots log predicted casualties in mines over the next electoral cycle for notional districts with 0% and 90% close elections in the previous elections respectively. It is based on Column (6) of Table 4.5. Note that the value in election year is 0 by construction for a district with no close elections. Dotted lines represent 95% confidence intervals.

C.2 Data appendix

Table C.5: Primary data sources

Data	Source
Mining Output (State -Level)	EOPP Indian States Data and Planning Commission of India
Mineral Licensing (State Level)	Indian Bureau of Mines
Mining Fatalities (State and District Level)	Directorate General of Mines Safety, Ministry of Labour and Employment
Naxalite Conflict (District Level)	Uppsala Conflict Data Program (UCDP) and South Asia Terrorism Portal (SATP)
State Level Election Outcomes	Election Commission of India

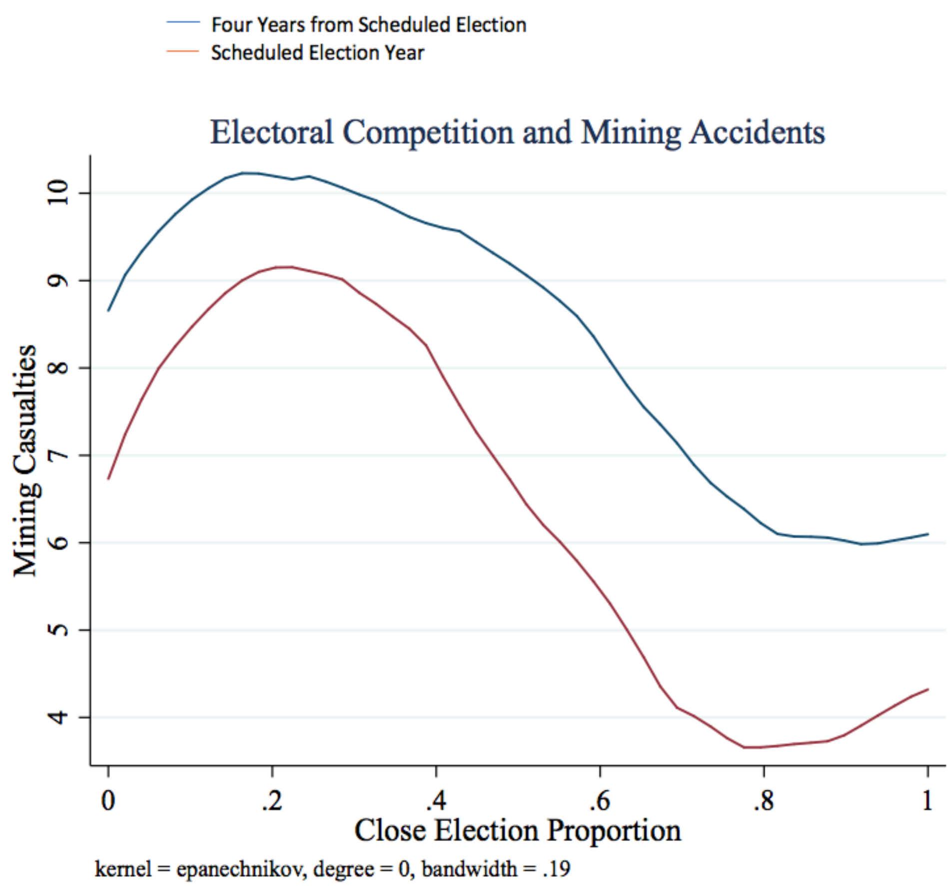
Figure C.10: Electoral cycle and accident probability (raw data)



Note: This figure presents district-level averages of the probability of a casualty ("Any Bad Event") over the electoral cycle. The variable "Any Bad Event" is a dummy coded 1 if at least one person was either seriously injured or killed in a mining accident in a district in a given year. Data from a total of 104 districts across 15 Indian states have been used to generate this graph. The 95 % confidence intervals are shown.

This figure plots district-level averages of the probability of a mining field accident in each year of the electoral cycle. There are no fixed effects included in this analysis.

Figure C.11: Electoral cycle, close elections and mining accidents (local polynomial)



Note: This figure presents a kernel weighted local polynomial fit of mining accidents against close election proportions conditional on elections being four years away, and in the election year respectively. Each polynomial in a different colour represents a different year in the electoral cycle. There are no fixed effects included in this analysis.

Figure C.8 presents plots from kernel-weighted local polynomial regressions of mining casualties on the proportion of close elections across all districts, in the year immediately after the election and in the next scheduled election year respectively. There are no fixed effects included in generating these graphs. The disproportionately large drop in fatalities in politically competitive districts during election years is evident in the raw data itself as presented above. Notice that the polynomial in red (i.e. the plot in the election year) has a steeper downward slope and has a extended flatter portion compared to the polynomial in blue (the plot four years from a scheduled election year). This implies that the difference between accidents in competitive and noncompetitive districts increase leading up to elections.