Spring 5-15-2018

Essays on Macroeconomics and Economic Development

Lijun Zhu
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Economics

Dissertation Examination Committee:
Michele Boldrin, Chair
Rodolfo Manuelli, Co-chair
Yongseok Shin
Ping Wang
Yi Wen

Essays on Macroeconomics and Economic Development

by

Lijun Zhu

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May, 2018

St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

Acknowledgements

I am deeply indebted to Michele Boldrin for his patience and continuous support during my years as both a PhD student. His guidance and insights have undoubtedly influenced the present dissertation. I also thank Rodolfo Manuelli for various discussions on the main chapter and detailed guidance on the third chapter of the dissertation. The dissertation were further enriched by comments and suggestions by Yongseok Shin, Ping Wang, Yi Wen, Zhiwu Chen, and Kaixiang Peng whom I thank for their involvement too. Finally, I thank the Graduate School of Washington University for the financial support.

Lijun Zhu

Washington University in St. Louis

May, 2018

ABSTRACT OF THE DISSERTATION

Essays on Macroeconomics and Economic Development

by

Lijun Zhu

Doctor of Philosophy in Economics

Washington University in St. Louis, 2018

Professor Michele Boldrin, Chair

Industrial Concentration and the Declining Labor Share. The labor share of national income in the United States has declined since the 1980s and especially after 2000. My paper focuses on the role played by technological change in this process. In particular, firms that adopt new technologies achieve a low labor share, and grow and take a larger market share over time. An example is online retailers such as Amazon, empowered by information technology, that have a lower labor share than traditional retailers, and have continually expanded over the last 20 years. This reallocation process drives down the aggregate labor share. I first document three facts: (i) across sectors, there is a negative correlation between change in concentration and change in the labor share; (ii) large firms usually exhibit a smaller labor share; (iii) in sectors where the labor share declines, the decline is especially strong among large firms. Specifically, gains in labor productivity are not associated with comparable increases in wages. Then I provide a rationale for these facts by assuming that capital and labor are complementary inputs and technological progress is labor saving and embodied in the capital stock. Under these assumptions, my model predicts a negative

correlation between firm size and labor share. Further, the adoption of new technologies diminishes the labor shares in large firms and increases their market share. As a consequence, the aggregate labor share declines. This technological channel is consistent with the evolution of labor productivity across sectors during the last 30 years.

Social-Economic Change and its Impact on Violence: Homicide History of Qing China. This paper constructs a quantitative history of the homicide rate in Qing China and investigates its social and economic drivers. Estimates based on historical archives indicate that this annual rate ranged between 0.35 and 1.47 per 100,000 inhabitants during the 1661-1898 period, a low level unmatched by Western Europe until the late 19th century. China?s homicide rate rose steadily from 1661 to 1821 but declined gradually thereafter until the turn of the century. Although extreme, homicide represents a random sampling of the entire distribution of interpersonal violence; hence the homicide rate serves as a proxy for overall violence, and its rise implies a decline in personal security. We use national and cross-provincial panel data to show that population density, state capacity, local self-governance, interregional grain market integration, and grain price level (which captures crop failure and other survival distress) are all statistically significant drivers of the homicide rate in 18th- and 19th-century China.

Labor Unions and the Labor Wedge: A Macroeconomic Perspective. The measured labor wedge, defined as the difference between marginal product of labor and marginal rate of substitution, is relatively stable from 1940s to 70s, and declines secularly from 1980s onwards. This paper aims to investigate the effect of a particular labor market institution, labor union, on labor wedge. Labor unions command a wage premium, which invites job application queues and job rationing in the unionized sector. This waiting values of unionized jobs creates a wedge between wages and households' willing to work. We provide sectoral evidence that supports a union-wedge connection in the manufacturing sector. A quantitative model which features two labor market, one competitive and the other unionized, is developed to estimate the effect of union power on labor wedge. Based on our quantitative results, approximately 20% of the decline in labor wedge from 1970s to 2000s is accounted for by the decrease in union densities.

Chapter 1

Industrial Concentration and the Declining Labor Share

Lijun Zhu

1.1 Introduction

The long-run constancy of labor share (LS) is typically viewed as a stylized fact of economic growth. Among other things, it justifies the widespread use of an aggregate Cobb-Douglas production function in macroeconomics. However, the recent literature documents the break of this constancy: in the United States the share of national income that goes to labor has oscillated cyclically during the whole post-WWII period and, more relevant in this context, it has clearly been declining for more than three decades. Albeit to different extents, this long-term decline is visible also in other advanced countries (Elsby et al. (2013); Karabarbounis and Neiman (2014b); Piketty and Zucman (2014)). My paper studies the role played by technological change in this process. In particular, new technologies create, or are adopted by, firms that achieve a low labor share, and grow and take a larger market share over time.

An example is online retailers such as Amazon, empowered by information technology, that have a lower labor share than traditional retailers, and have continually expanded over the last 20 years.[1] This reallocation process drives down the aggregate labor share.

My research is motivated by three empirical facts documented in the paper. First, there is a negative correlation between change in market concentration—measured as the share of total revenue in a sector attributable to its largest firms—and change in labor share, in that sector. Even though the aggregate labor share has declined, the decline is not universal across sectors. The decline concentrates in several sectors, namely, manufacturing, retail and wholesale trade[2], and transportation & warehousing. These are also sectors where concentration has increased the most. From 1997[3] to 2012, the revenues share of the 50 largest firms increased from 20.3% to 27.6% in wholesale trade, from 25.7% to 36.9% in retail trade, and from 30.7% to 42.1% in transportation. This negative correlation between change in concentration and change in labor share also holds across more disaggregated manufacturing sectors (i.e. at various digits North American Industry Classification System (NAICS) sectors), and is robust for different periods, for a range of cutoffs for what constitutes a "large" firm, and to adding other control variables and sector fixed effect. Concentration ratio for the whole manufacturing sector is relatively stable from 1960s to around 1980, and increases thereafter. The manufacturing labor share shows a similar but opposite pattern: relatively stable until around 1980, and declines in the last 3-4 decades. In almost all the other service and fi-

[1] The payroll-sales ratio in "Electronic shopping and mail-order houses" subsector is 30% lower than that in the retail trade sector in 2012. From 1992 to 2014, the share of nonstore sales (the majority of which are electronic shopping) in retail sales increases from about 6% to 13%. Of course, electronic commerce increases labor demand in the transportation sector. However, the increases in labor productivity in retail trade and in the distribution sector which contains trade and transportation are almost the same, suggesting this spillover effect is secondary.

[2] To avoid confusion I should point out that the term "trade" is used to denote wholesale trade and retail trade in the text.

[3] The NAICS system replaced the old Standard Industrial Classification (SIC) system in 1997, so the data used for non-manufacturing sectors are from 1997 onward.

nance sectors, the labor share did not fall and concentration increased at a much slower pace.

Second, the *relative* labor share of large firms, defined as the ratio between the labor share in these firms to the average in the sector, tend to be smaller than that of other firms. In 2002, labor share for the 50 largest manufacturing firms[4] was 67% of that for the manufacturing sector as a whole. In the same year, the relative labor share for 50th to 100th, 101st to 150th, 151st to 200th largest firms, and 201st largest and smaller firms in the manufacturing sector were 73%, 82%, 97%, and 121%, respectively. Large firms typically offer a higher wage than do small firms. However, wage differentials compensate only partially for the even wider gap in labor productivity between large and small firms, which results in a lower labor share in the former.

Third, the relative labor share for large manufacturing firms has been declining since around 1980 and this coincides with the time when the labor share in the manufacturing sector began to decline. The relative labor share for 50 largest firms in manufacturing was 98% in 1967, 97% in 1977, 92% in 1987; it then declined steadily to 72% and 59% in 1997 and 2012, respectively. The implication is that, in comparison with other firms, the decline of labor share in large manufacturing firms has been especially pronounced. From 1967 to 1977, the relative labor productivity of top-50 firms increased from 128% to 143% and their relative wage increased at about the same rate: from 127% to 139%[5]. However, there has been an increasing divergence between the two series since the late 1970s. From 1977 to 2012, the relative labor productivity of top-50 firms increased from 143% to 242%; yet at the same time, their relative wage was essentially stable: 139% in 1977 and 144% in 2012.

[4] These Census data give concentration ratios for the 50, 20, 8, and 4 largest firms. In the baseline case, I use the 4 largest firms for 6 digit NAICS sectors, and the 50 largest firms for 2 digit sectors. As shown in the text, the empirical pattern is robust to different choices in this regard.

[5] Relative labor productivity (resp. wage) of large firms is defined as the ratio of labor productivity in these firms to labor productivity (resp. wage) in their sector.

Increasing concentration, i.e. the rising market share of large firms that have a lower labor share and the decline of labor shares within large firms are the driving forces behind the overall decline of aggregate labor share in manufacturing.

Large firms in non-manufacturing sectors also have a lower labor share. However, their relative labor share exhibits a different time pattern across sectors. In particular, it declines in the trade and transportation sectors, but does not exhibit a clear trend in most service and finance sectors. For trade and transportation, just as for manufacturing, the relative labor share of large firms decreases over time owing to a combination of these firms' increasing relative labor productivity and a fairly stagnant relative wage. From 1997 to 2012, the relative labor productivity of the top-50 firms in wholesale trade increased by 95.9% yet their relative wage increased only by 16.2%. In transportation, the increase in relative labor productivity of top-50 firms was 32.8%—significantly greatly than the 1.6% increase in relative wage. In retail trade, the increase in relative labor productivity of large firms is 4.2% greater than the relative wage.

I provide a rationale for these empirical facts that is based on two assumptions. First, for a given technology—embodied in machines—capital and labor are complementary inputs. In a constant elasticity of substitution (CES) production function, this is equivalent to an elasticity of substitution that is less than 1. Second, technological progress is labor saving in this sense: new technology embodied in new machines allows for less labor input per unit of output. I start from a static model featuring N vintages of capital. Under these two assumptions, my model predicts a negative correlation between firm size[6] and labor share. In particular, more advanced technologies increase output and decrease the labor share. The

[6] In the static case, each technology is interpreted as a firm. This presupposition will be relaxed when I discuss the general equilibrium.

intuition is as follows. Technology complements capital and increases its productivity, which further increases demand for effective labor.Therefore firms using more advanced technologies produce more output with less labor, reducing the labor share of income.

I then extend the static model to a dynamic general equilibrium one by incorporating heterogeneous firms and capital accumulation. Each firm is endowed with a level of own-productivity, and each firm optimally choose to adopt one technology from all feasible ones. The fixed cost of installing a machine is assumed to be an increasing function of the technological vintage it embodies. Firms with low own-productivity find it optimal to avoid the large fixed cost of advanced technologies by adopting less advanced and cheaper ones. On the other hand, firms with high own-productivity find it convenient to pay the fixed cost of adopting advanced technologies and therefore use them. Due to the negative effect of technology on labor share, more productive firms exhibit a lower labor share.

Given a fixed number of technological vintages, capital-labor complementarity and fixed labor supply, the economy arrives at a steady state. When a new and more advanced technological vintage becomes exogenously available, the most productive firms find it profitable to switch to that new technology. This response increases their market share because, as established for the static case, technology has a positive effect on firm size; in other words, concentration increases with technological change. Furthermore, the more advanced technology reduces the labor share as firms increase in size. As a consequence, the aggregate labor share declines. Next, to illustrate the dynamics of the model in the simplest possible case, I calibrate a steady-state economy in which, initially, there are two technological vintages and a third more advanced one becomes available.

For a CES production function in which capital and labor are complementary inputs, any labor saving technological change that reduces the labor share will simultaneously increase labor productivity. Hence we should observe a faster increase in labor productivity in sectors, or in periods, characterized by declining labor share. From 1987 to 1997, the labor share in manufacturing declined 10.3% while labor productivity increased 34.7%. Since the late 1990s, both decreases in labor share and increases in labor productivity have accelerated. From 1997 to 2007, the labor share fell 20.2% while labor productivity rose 59.0%. From 1987 to 2016, the economy-wide labor productivity has increased by 72.7%; during the same period, labor productivity increases in manufacturing, wholesale trade and retail trade amounted to (respectively) 146.4%, 123.5%, and 128.8%. This evidence strongly supports the technological channel - as opposed to monopoly power - as the main driver of the negative correlation between concentration and labor share. Labor productivity is measured as the ratio of real value added (net of price changes) to hours worked. An increase in monopoly power, alone, would have driven up prices but would have not increased labor productivity.

Two recent papers, Barkai (2016) and Autor et al. (2017), also independently document a negative correlation between change in concentration and change in labor share. Barkai (2016) focuses on the fact that an increase in monopoly power decreases both labor and capital shares while increasing the share of profits. Autor et al. (2017) conjectures that the concentration-LS correlation is driven by the rise of superstar firms, which have a lower labor share since the fixed overhead labor cost is distributed over a larger output base. My paper differs from both in that I emphasize both increasing concentration, which accounts for approximately 30% of the decline, and decrease of labor share in large firms are critical to the decline of the aggregate labor share, and argue that both are driven by the labor saving technological advancement.

Koh et al. (2016) claim that the decline in the U.S. labor share is accounted for by the growing share of Intellectual Property Products (IPP) in total capital stock: two thirds of the decline is driven, in the data, by the higher depreciation rate of IPP capital. Karabarbounis and Neiman (2014a) finds that capital depreciation explains about 45% of the 4.7% decrease in gross labor share in the U.S. corporate sector, and that both net and gross labor shares have declined 'meaningfully', worldwide, since 1975. In light of these facts I have adjusted the labor shares I use by netting out the depreciation of capital (but not the net return to IPP capital as in Koh et al. (2016)), both traditional and IPP, and found that the patterns studied in my paper are robust to these adjustments.

This paper relates most closely to the macroeconomic literature on the determinants of the aggregate labor share, which dates back to at least the early-to-middle 20th century (e.g. Kaldor (1957); Solow (1958)), and the recent studies of its decline in the United States and in other developed countries. Karabarbounis and Neiman (2014b) attributes the decline of labor share to decreases in the relative price of investment goods while Piketty and Zucman (2014) attributes it entirely to the process of capital accumulation. In a CES production function, these channels lead to a lower LS if capital and labor are substitutes, i.e. if the elasticity of substitution between capital and labor is greater than 1. My paper departs from these approaches by featuring a production function for which the elasticity of substitution between capital and labor ranges between 0 and 1, which is consistent with the majority of empirical estimations[7]. In addition, at the aggregate level the return to capital, measured using National Income and Product Accounts (NIPA) tables while accounting for changes in

[7]For example, Brown and DeCani (1963) estimates that the elasticity of substitution ranged from 0.08 to 0.44 over the period from 1890 to 1958. David and van de Klundert (1965) estimates an elasticity of 0.32, and the estimate in Wilkinson (1968) is 0.5. Most recent estimates also obtain values between 0 and 1. The estimated elasticity of substitution between capital and (skilled) labor is 0.67 in Krusell et al. (2000). In Antras (2004), the estimated elasticity of substitution between capital and labor ranges between 0.6 and 0.9. Klump et al. (2007) estimates the elasticity to be 0.51. Herrendorf et al. (2015) estimates the elasticity of substitution to be 0.80 in manufacturing, and 0.84 for the whole economy. Oberfield and Raval (2014) use plant level data, and they estimate the elasticity of 0.51 at that level and 0.71 for the Manufacturing sector.

relative prices of investment goods, has not declined over the last three to four decades. See Gomme et al. (2011) and the discussion in Section 4.2 of this paper.

Lawrence (2015) shows that rapid progress in labor-augmenting technological change in the United States reduces the effective capital-labor ratio and labor share under an aggregate production function where capital and labor are complementrary inputs. Elsby et al. (2013) decompose declines of aggregate labor share into sectors, and they identify the offshoring of labor-intensive tasks as a potential driver. My paper complements that research because offshoring, no differently than other forms of technological change, such as automation, enables reduced (domestic) labor input per unit of output and thus is viewed as a labor-saving technology. The increase of offshoring is heavily concentrated in Manufacturing[8]. The LS declines in both trade and transportation are also substantial, but the extent of offshoring in these sectors has not changed significantly during the last 15-20 years. The technological channel proposed in my paper has the potential to explain the decline of labor share in a wider context.

Several papers use firm level data in seeking to understand the decline of aggregate labor share. Loecker and Jan (2017) document a rise in the average markups across public firms since 1980, which they argue could account for the reduction in labor share. Kehrig and Vincent (2017) also report a reallocation of market share towards hyper-productive manufacturing plants, which arrive at a low labor share by gradually increasing value added while keeping employment and compensation unchanged. These authors show how this pattern can be explained: the concave response of hiring to total factor productivity shocks is becoming more concave over time. My proposed explanation is consistent with the firm-level evidence, yet it also comports with observed sector heterogeneity in terms of concentration,

[8]See Figure A in the Appendix.

labor share and labor productivity.

The rest of paper is organized as follows. The empirical facts are documented in section 2, and Section 3 develops a model that rationalizes those facts and presents a quantitative exercise to illustrate the mechanism. I discuss several related issues in Section 4, including the return to capital, the evolution of labor productivity across sectors, and firm level capital/labor ratios. Section 5 concludes.

1.2 Empirical Facts

Several empirical facts are presented in this section. First, I document a negative correlation—across manufacturing sectors—between change in concentration (measured as the value-added share of large firms)[9] and change in labor share. The baseline definition of "large" firms is the 4 largest firms in an NAICS 6-digit sector or the 50 largest firms at the 2-digit level.[10] If large firms have a lower labor share, then increases in concentration (i.e., an increased market share for large firms) lead to a lower labor share for the sector. I introduce the concept of *relative labor share*, defined as the ratio of the labor share in a subset of firms to sector labor share, and find that the relative labor share for large manufacturing firms is less than that for other firms. Note also that the relative labor share of large firms is stable from the 1960s to the late 1970s but declines thereafter. The relative LS for the 50 largest manufacturing firms was 98% in 1967 and 97% in 1977 but then declined steadily to 59% in 2012.

This empirical pattern observed in manufacturing holds also in other sectors. Declining

[9] Value added is used whenever the necessary data are available; otherwise, I use the share of revenue as the measure of concentration. Appendix Figure A.6 compares these two measures vis-à-vis manufacturing and shows that the difference is negligible.

[10] The choice of 4 and 50 reflect, inter alia, the availability of data. However, the described pattern is robust to other definitions of a large firm.

labor share is characteristic not only of manufacturing but also of retail trade, wholesale trade, and transportation & warehousing. In these sectors, concentration has risen significantly over the same period while the relative labor share of large firms has declined. In contrast, the labor share does not decline in most finance and services sectors, where the relative labor share of large firms does not exhibit a clear time trend.

1.2.1 Manufacturing

Figure 1.1 plots the labor share in manufacturing from 1947 to 2015.[11] The manufacturing LS is relatively stable from the 1940s to the early 1980s and declines steadily thereafter. The magnitude of decline from 1980 to 2015 is nearly 20%. From 1997 to 2007, the baseline period for which I have disaggregated data, the labor share in manufacturing decreases by about 6%.

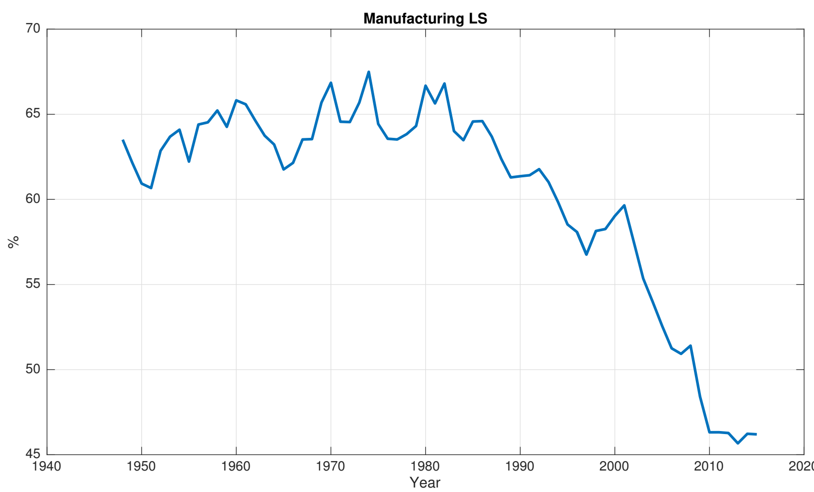


Figure 1.1: Labor Share in Manufacturing, 1947-2015

Note: Labor share (LS) is measured as the fraction of compensation in value added.

[11]Labor share in Figure 1.1 is calculated as compensation of employees *divided by* manufacturing value added. The general trend—stable then declining—is robust to adjustments for proprietor income and depreciation. The nonadjusted series is chosen as the benchmark measure for consistency with respect to the measure used in manufacturing *sub*sectors, for which data on proprietor income and depreciation are not available.

The disaggregated data are from the Annual Survey of Manufactures (ASM), a survey of manufacturing establishments with one or more paid employees. A summarized and simplified version of the ASM data is provided in the NBER-CES data set. Included in the data are payroll and value added for manufacturing sectors at various NAICS digit levels. The labor share is the fraction of payroll in value added, and concentration is measured as the share of value added due to a sector's 4 largest firms (*Share04*). The concentration data are available every five years from the Economic Census, for which 2012 is the most recent publication.

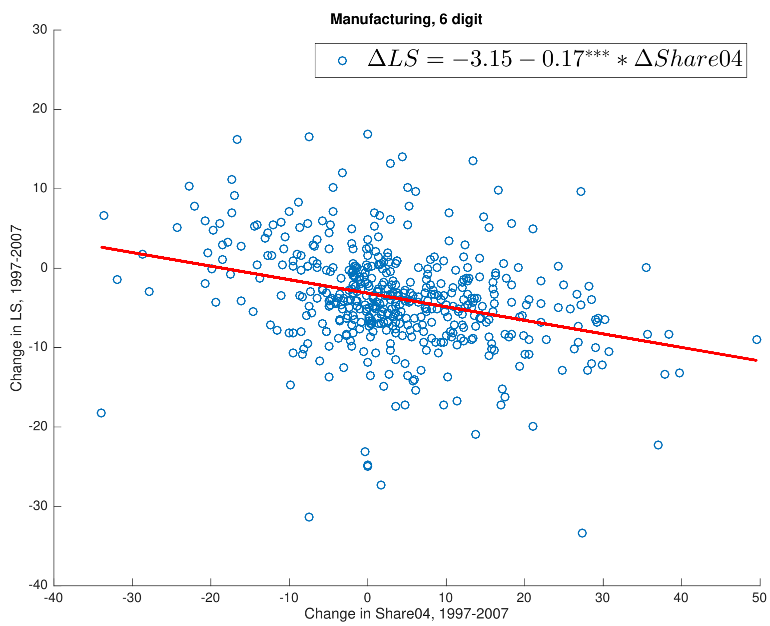


Figure 1.2: Change in Labor share versus Change in Concentration, MFG 1997-2007

Note: Labor share (on the vertical axis) is the fraction of payroll in value added; concentration (on the horizontal axis) is the value added share of the 4 largest firms. Each circle represents an NAICS 6-digit manufacturing sector.

Figure 1.2 plots change in labor share against change in concentration (*share04*), across 6 digit manufacturing sectors, from 1997 to 2007[12]. This graph reveals the negative and

[12]NAICS has been in use since 1997. I chose 2007 as the ending year because the 6-digit code underwent

significant correlation between change in concentration and change in labor share. When a sector becomes more concentrated, its labor share tends to decline. From 1997 to 2007, ASM data shows that the manufacturing LS decreased by 5.61%, due mostly (73%[13] to within-sector declines. Over this period, concentration has increased in nearly two thirds of the 465 6-digit manufacturing sectors.

The result of a single variable regression, where the dependent variable is change in the labor share and the independent variable is change in *Share04*, is[14]

$$\Delta\text{LS} = -3.15 - \underbrace{0.17^{***}}_{(0.02)} \times \Delta \mathit{Share04} + \epsilon, \qquad R^2 = 0.1025,\ N = 464;$$

As shown in the Appendix (see Table A.1 and A.3 and Figure A.3, the negative correlation between concentration and labor share also holds across manufacturing sectors at the 3-, 4-, and 5-digit levels and across different periods. Moreover, the results are robust to measuring concentration by the value added share of the 8, 20, and 50 largest firms (instead of 4 largest firms) within a sector. The result is not affected by the omission of fringe benefits when measuring labor share. Since 2005, the publicly available ASM tables have provide information on payroll as well as benefits. "Compensation" is the sum of payroll and total fringe benefits, which include the employer's cost for health insurance, defined benefit pension plans, other defined contribution plans, and other fringe benefits. Appendix Table A.4 reports a similar negative and significant correlation between change in concentration and change in the revised labor share.[15] Last, as shown in Appendix Table A.2, this correlation

a major revision in 2012. A similar result is obtained if concentration is measured in terms of revenue rather than value added.

[13]This number is based on a standard within-between decomposition.

[14]Blank magnetic and optical recording media manufacturing (NAICS code 334613), has a labor share exceeding 100% in 2007, so that subsector is excluded in my calculation

[15]In 2012, the total numbers of NAICS 6-digit manufacturing sectors was reduced from 467 to 362. I therefore use data at the 5-digit level, which are fairly stable: there were 184 5-digit manufacturing sectors in 2007 and 180 in 2012, of which 175 were unchanged.

is robust in panel regressions with sector fixed effect and other control variables.

Concentration in Manufacturing has been increasing. For industrial classifications, the Census of Manufactures (published every five years) used SIC codes before 1992 and switched to NAICS codes in 1997. Although each system underwent revision, the total numbers of SIC 4-digit manufacturing sectors (444 in 1977) is comparable to the number of NAICS 6-digit sectors (467 in 2007).[16] Manufacturing concentration—measured as the average *Share04* (and the average *Share08*) across SIC 4-digit sectors until 1992 and across NIACS 6-digit sectors starting in 1997—weighted by revenue[17]. Results from the 1963–2012 period[18] are presented in Figure 1.3.

[16]For details, see Appendix Table A.8.

[17]The concentration measure is given in terms of revenue (rather than value added) because that is the measure until 1992 for disaggregated manufacturing sectors. Yet the pattern is robust to various measures. See Appendix Table A.8.

[18]The original full reports of the Census of Manufactures could not be found for 1947, 1954 and 1958, so those years are excluded from the date set. However, the 1958 summary report indicates that concentration is stable in most sectors for the 1947–1958 period.

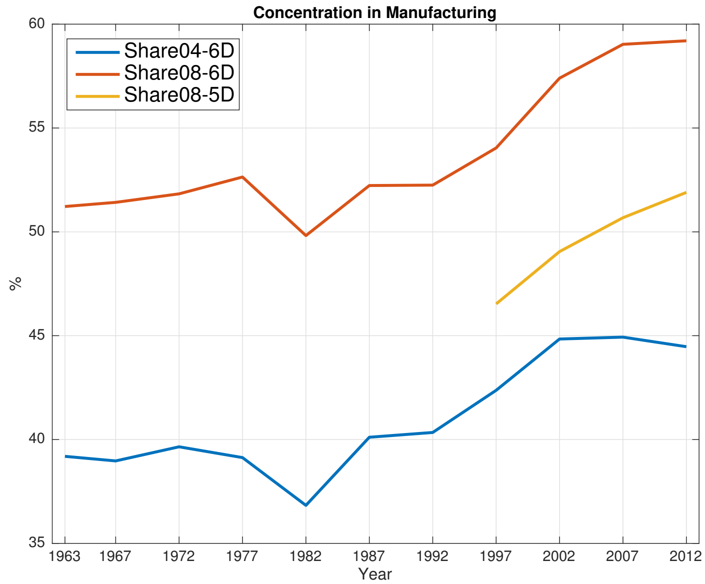


Figure 1.3: Concentration in Manufacturing, 1963-2012

Note: The blue (resp. red) line plots the average revenue share of the 4 (resp. 8) largest firms across 6-digit manufacturing sectors; the yellow line is the average revenue share of the 8 largest firms across 5-digit manufacturing sectors. All values are weighted by revenue.

Concentration in Manufacturing was relatively stable from the 1960s to the early 1980s, but it has increased steadily over the past three decades. Before 2000, the increasing concentration was due mainly to the 4 largest firms in each sector (the increase in average share of the 8 largest firms was similar to that of the 4 largest). After 2000, the average *share04* was relatively stable but the average *share08* continued to increase, which suggests that the expansion of large firms—though not the very largest—has driven up concentration in the past decade. From 2007 to 2012, the total number of NAICS 6-digit manufacturing sectors was reduced from 467 to 364. The figure also plots *Share08* at the 5-digit level, which is consistently defined across the period shown.

The increase in concentration means that market share has been transferred from small to large firms. If the labor share is lower in large firms, then an increase in concentration reduces sector LS in a mechanical way. The labor shares of different firms can be compared using concentration data from the Economic Census. Toward that end, I define the relative labor share (RLS) for a subset of firms (e.g. 50 largest firms)—where firm size is measured by value added or revenue—as the ratio of LS for these firms to the sector's LS. So if the top-50 firms have an RLS that is less than 100%, then the labor share of these firms is lower than the sector average.

The relative labor share for the top-50 firms is calculated as follows:

$$\text{RLS-top50} \equiv \frac{\text{LS-top50}}{\text{LS-Sector}} \times 100\% = \frac{\text{Payroll-top50 / Payroll-Sector}}{\text{Vadd-top50 / Vadd-Sector}} \times 100\%$$

That is, the relative labor share for top-50 firms equals their payroll share in the sector *divided by* their value added share; the relative labor productivity and relative wage are defined similarly. Labor productivity (LP) is measured as value added per worker[19]. The relative labor productivity (resp. wage) for top-50 firms is measured as the ratio of labor productivity (resp. wage) for those firms to that of sector. It follows that the relative LS of top-50 firms equals the ratio of their relative wage to their relative labor productivity.

Table 1.1 presents relevant statistics for manufacturing in 2002. The repirted values for share of employees, payroll, and value added are from original (sourced) tables; the valued for relative LP, relative wage, and RLS are calculated as just described.

[19]The actual labor productivity is output—not value added which is equal to output times price—per hours worked. As defined in the text, relative LP reflects the differences in actual LP across firms *provided that* those firms do not exhibit significant differences in prices and average working hours per worker

Table 1.1: Share of Industry Statistics (%) for the Manufacturing Sector, 2002

Firm groups	Emp.	Payroll	Vadd.	Rel. LP	Rel. Wage	Rel. LS
50 largest	12.1	16.9	25.3	209	140	67
50th to 100th largest	5.3	6.1	8.4	158	115	73
101st to 150th largest	3.9	4.1	5.0	128	105	82
151st to 200th largest	3.1	3.6	3.7	119	116	97
201st and smaller	75.5	69.3	57.5	76	92	121

Note: Firms/companies are ranked by value added, and a firm/company is defined as a business organization consisting of one establishment or more under common ownership or control.

Data Source: "*Concentration Ratios: 2002 Economic Census, Manufacturing, Subject Series*", in *Census of Manufactures*, 2002.

Both labor productivity and wage are higher in large firms, but the labor share in large firms is lower. The RLS for the 50 largest manufacturing firms (as ranked by valued added) is 67% of the entire sector. Although the wage in small firms is lower, it accounts for a larger portion of the per-capita value added in these firms—as reflected in their 121% of relative labor share. The pattern that large firms have smaller labor shares also holds for other years. Appendix Table A.9 reports results for 1997, 2007 and 2012.

The trend of relative labor share for 50 the largest manufacturing firms(again ranked by value added) from 1967 to 2012 is plotted in Figure 1.4. That RLS remains fairly constant from the 1960s to the 1980s, but the figure shows a clear downward trend thereafter. The relative labor share for the 50 largest manufacturing firms was 98% in 1967 and 97% in 1977, but it declined steadily to 59% in 2012.

Recall that the relative LS equals the ratio of relative wage to relative labor productivity.

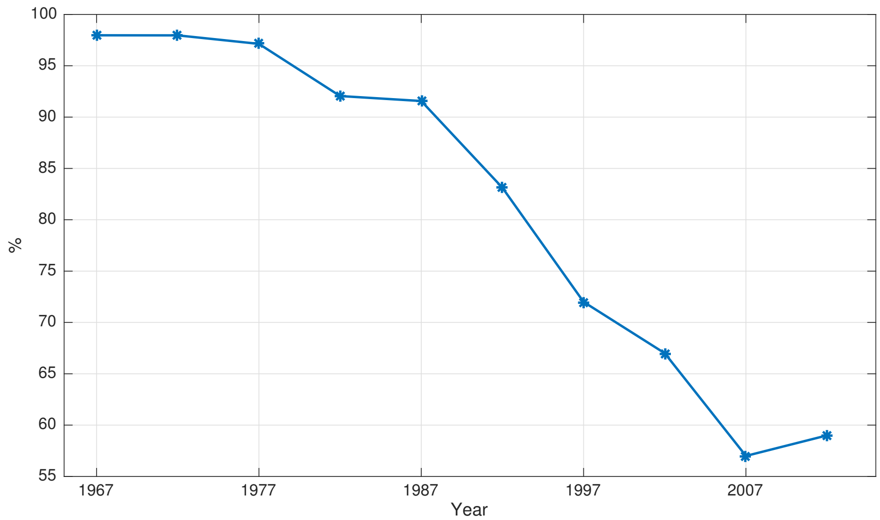


Figure 1.4: Relative Labor Share of the 50 largest MFG firms

Note: The relative labor share in this graph is calculated as LS in the 50 largest manufacturing firms *divided by* LS in the entire manufacturing sector.

Figure 1.5 displays the trends of relative LP and relative wage for the 50 largest manufacturing firms. From 1967 to 1977, the relative LP of these top-50 firms increased from 128% to 143% while their relative wage increased at about the same rate: from 127% to 139%. From 1977 to 2012, however, the relative labor productivity of those firms increased from 143% to 242%. Yet their relative wage was practically stable, increasing only to 144% in 2012. The clear divergence between relative LP and relative wage since the 1980s suggests that marginal workers in large and small firms are easily substitutable with each other. The Census of Manufactures public data also allows me to calculate relative labor shares for the 100 and 200 largest firms; Appendix Table A.10 establishes that the overall patterns are similar.[20]

[20]The manufacturing sector comprises many heterogeneous subsectors, and the top-50 constitute but a small subset of total firms. From 1997 to 2012, the value added share of the top-50 firms increased from 24.5% to 26.1%. This small increase reflects the concomitant modest increases for large (but not top-50) firms in the sector. For instance, the value added share of the 200 largest manufacturing firms increased by 2.9% for the same period; and the average share of the 50 largest firms across 3-digit manufacturing sectors increases 5.7%. Even so, the RLS of the top-50 firms is still informative—as evidenced by the strongly similar RLS pattern of the 100 and 200 largest firms. Data for top-50 firms are available also for non-manufacturing

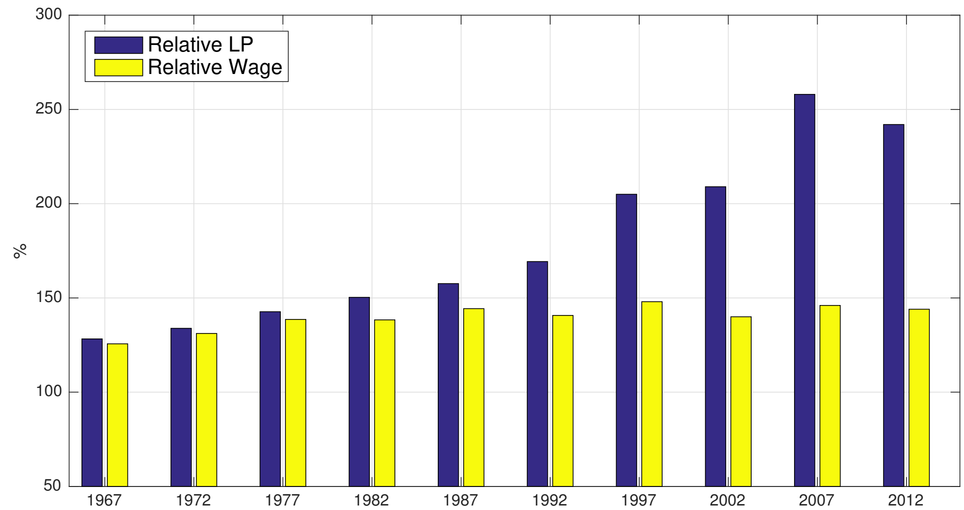


Figure 1.5: Relative Labor Productivity and Relative Wage of the 50 Largest MFG Firms

Note: Relative labor productivity (, resp. relative wage) is calculated as the LP (resp. wage) in the 50 largest manufacturing firms *divided by* the LP (resp. wage) in the manufacturing sector.

In the 1990s and 2000s, the relative labor share of large firms was significantly smaller than for other firms; an increase in concentration thus leads to a decline in sector labor share. In the 1960s and early 1970s there was not much difference between the labor shares of large versus small firms, so one would expect that increased concentration would lead to a relatively smaller decline in labor share. These statements are confirmed in Appendix Table A.3, which reports results from single-variable regressions (with changes in the labor share and in concentration as the dependent and independent variable, respectively) from 1963 to 1966 and also from 1972 to 1977. The coefficients for concentration are significantly smaller than these two periods than those for the 1980s, 1990s, and 2000s.

The pattern that large firms have lower labor shares continues to hold in more disaggregated manufacturing sub-sectors. I find public data that enables calculation of subsector-level

sectors and are therefore used as the baseline (to ensure consistency across sectors).

Table 1.2: Relative Labor Share in the Manufacturing Sector, 1977

SIC	Index	1-4	5-20	21-50	≥51
4 digit	average	91.0%	99.4	108.8	119.5
4 digit	weighted average	91.6	97.2	108.6	119.8

Note: "1-4" denotes the 4 largest firms, (weight=value added).

relative labor shares in the Census of Manufactures report for 1977 but not for more recent years. The original data include payroll and value added for the 4, 8, 20, and 50 largest firms in each SIC 4-digit sector. Firms are classified into four groups: "1-4", the 4 largest firms in terms of value added; "5-20", the 5th to 20th largest firms; "21-50", the 21st to 50th, and "≥51", the 51th and smaller firms. Table 1.2 summarizes the calculated RLS values, and it shows the same pattern as before: the relative labor share of large firms is less than that for other firms.

Compustat data can be used to investigate the evolution of relative labor shares of large firms within more disaggregated manufacturing sectors. Among Compustat firms, about 85% report sale numbers but only 22% report total staff expenses. The latter include wages, salaries, pension costs, profit sharing and incentive compensation, payroll taxes and other employee benefits; this expense category corresponds closely to NIPA's "compensation of employees". Because value-added data for firms are not directly available, I use the the fraction of compensation in sales as a proxy for firm-level labor share. Sector dummies are assigned for each SIC 2-digit sector (e.g. textile mill products; electronic & other electronic equipment). From 2010 to 2014, there were, on average 18 firms per sector.

The following regression is performed for manufacturing firms (i.e., SIC $\in [2000, 3999]$)

Table 1.3: Correlation between Size and Labor Share in the Manufacturing Sector

	70-74	75-79	80-84	85-89	90-94	95-99	00-04	05-09	10-14
Size	0.41***	0.48***	0.22**	−0.22**	0.01	−0.88***	−1.61***	−1.31***	−1.93***
	(0.12)	(0.10)	(0.12)	(0.12)	(0.12)	(0.15)	(0.15)	(0.15)	(0.13)
Sector D.	Yes	–	–	–	–	–	–	–	–
Year D.	Yes	–	–	–	–	–	–	–	–
R^2	0.46	0.40	0.37	0.24	0.23	0.22	0.27	0.26	0.32
Obs.	1674	1847	1428	1222	1166	1145	1035	957	1528

Note: $*** : p < 1\%$, $** : p < 5\%$; $* : p < 10\%$. Size is measured as assets (in log). LS is the share of compensation in revenue. Results are qualitatively similar if using employments.
Data Source: Compustat, Manufacturing firms ($SIC \in [2000, 3999]$).

for each five-year period:

$$\text{LS}_i = \beta_0 + \beta_1(\text{Size}_i) + \text{Sector dummies} + \text{Year dummies} + \varepsilon_i.$$

Size is measured by (log) assets. Table 1.3 presents the coefficients β_1 for different periods. The correlation coefficient here also shows a downward trend[21] consistent with the aggregate outcomes. The implication is that the trend seen in Figure 1.4 reflects actual changes between large and small firms within sectors—that is, rather than simply a shift of large firms towards less labor intensive sectors.[22]

1.2.2 Non-Manufacturing sectors

I now turn from manufacturing to non-manufacturing sectors, where the negative correlation between changes in concentration and changes in labor share holds as well. Furthermore, in non-manufacturing sectors where the labor share declines, the concentration increases and

[21] The coefficient is positive in the 1970s and early 1980s, which accords with the top-50s' RLS being higher than the RLS of the top 100 and top 200 firms before 1980. This result may reflect relatively stronger union power in the largest manufacturing firms in the 1970s and early 1980s.

[22] Of the 50 largest manufacturing firms in 1980, 18 are related to petroleum & coal products, 2 to computer & other electronic products, and 0 to pharmaceuticals. In 2012, the corresponding numbers are 9, 8, and 5.

the relative labor share of large firms falls—just as in Manufacturing.

Following Elsby et al. (2013), the change in aggregate labor share can be decomposed into a within-sector component and a between-sector component. Formally, we can write

$$\mathrm{LS} = \Sigma_i \omega_i \mathrm{LS}_i \implies \Delta \mathrm{LS} \approx \underbrace{\Sigma_i \mathrm{LS}_i \Delta\omega_i}_{\text{between}} + \underbrace{\Sigma_i \omega_i \Delta \mathrm{LS}_i}_{\text{within}},$$

where i denotes a sector and ω_i represents the share of sector i's value added in the economy. Average values of LS and ω are used to calculate the "between" and "within" components. The baseline measure of sector labor share is the fraction of compensation in value added. As shown in Appendix Table A.15, the within-sector component accounts for most of the decline in labor share; structural change (i.e. the between-sector component) plays a secondary role.[23]. Table 1.4 summarizes the within-sector component. Sectors in which labor share declined significantly are manufacturing, whole trade, retail trade, and transportation. At the same time, the labor share in most finance and services sectors sectors did not decrease.

These results are robust to adjusting labor share for depreciation and proprietor's income. Appendix Table A.16 gives the decomposition results when value added is adjusted for capital depreciation, which includes depreciation not only of traditional capital (equip-

[23]The reported statistics exclude agriculture and government; they also excludes mining, construction, and management of companies & enterprises because these three sectors lack concentration data. The included sectors together accounted, in 2012, for 77% of gross domestic product (GDP) and 89% of private sector GDP. From 1987 to 2013, the labor share in mining decclined by about 20%; that decrease was most pronounced during the years of 2000 and 2003. The labor share in construction declined moderately over this period and primarily during two recessions. Given the relatively small value added of these excluded sectors, their changes in labor share have limited effects on overall labor share. The results in Appendix Table A.15 are in substantial agreement with those of Elsby et al. (2013). The slight difference might be explained by two factors. First, I depart from their approach by focusing on the *nonfarm private* sector (rathern than the corporate sector) and by excluding three subsectors for which there are no concentration data. Second, the data used here reflect IPP revisions whereas Elsby et al. (2013) use pre-revision data.

Table 1.4: Within-sector Component of Declines in LS, 1987-2013

Sector	Aggr.	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ΔLS	-3.62%	2.25	-18.02	-7.36	-6.07	-10.0	-3.79	0.22
$\omega\Delta LS$		0.06	-3.43	-0.56	-0.51	-0.38	-0.23	0.02
	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
ΔLS	-0.99	8.88	2.36	1.72	5.25	3.97	-0.45	10.21
$\omega\Delta LS$	-0.16	0.70	0.08	0.02	0.42	0.05	-0.02	0.33

Note: LS is the share of compensation in value added. (1)-Utilities; (2)-Manufacturing; (3)-Wholesale Trade; (4)-Retail Trade; (5)-Transportation and Warehousing; (6)-Information; (7)-Finance and Insurance; (8)-Real Estate, rental and leasing; (9)-Professional, Scientific and Technical Services; (10)-Administrative and Waste Management Services; (11)-Educational Services; (12)-Health Care and Social Assistance; (13)-Arts, Entertainment, and Recreation; (14)-Accommodation and Food Services; (15)-Other Services. Data Source: NIPA Value-added-by-Industry.

ment and structures) but also of the newly capitalized Intellectual Property and Products (IPP). Proprietor's income does not significantly alter the baseline pattern; the reason is that, in most sectors, the share of proprietors' income in value added (reported in Appendix Table A.17) is relatively stable.[24]

Figure 1.6 presents the relation between change in labor share and change in concentration across 2-digit non-manufacturing sectors.[25]. Labor share is measured as the share of compensation of employees in Value added, and concentration is measured as the 50 largest

[24]Scholars often adjust labor share for proprietors' income by assuming a constant labor share for proprietors and the corporate sector. In this approach, labor share under is defined as compensation/ (value added − proprietor income). The decreasing shares of proprietors' income in value added for the health care sector and the professional, scientific, & technical services sector are partly responsible for the increase of their labor share (reported in Appendix Table A.15).

[25]Labor share can not be measured systematically for disaggregated non-manufacturing sectors because value added data are not available. For services sectors and in 1997, concentration measures are available only for establishments subject to federal income tax. The same measure for these sectors is used in 2012. See Appendix Figure A.4 for the relation between changes in labor share and changes in concentration (measured as revenue share) of the 4 largest firms. In some sectors, the LS fluctuates instead of exhibiting a monotonic trend. Appendix Table A.11 reports the labor shares for all 15 sectors. Based on observing the trend, I tried to make minimum adjustments to the sample dates selected. In particular, I use the 1996 (rather than 1997) labor share for the information sector,; I also use the average LS between 1996 and 1998 for education services instead of the 1997 valued. The pattern that results from these two revisions are presented in Appendix Figure A.5.

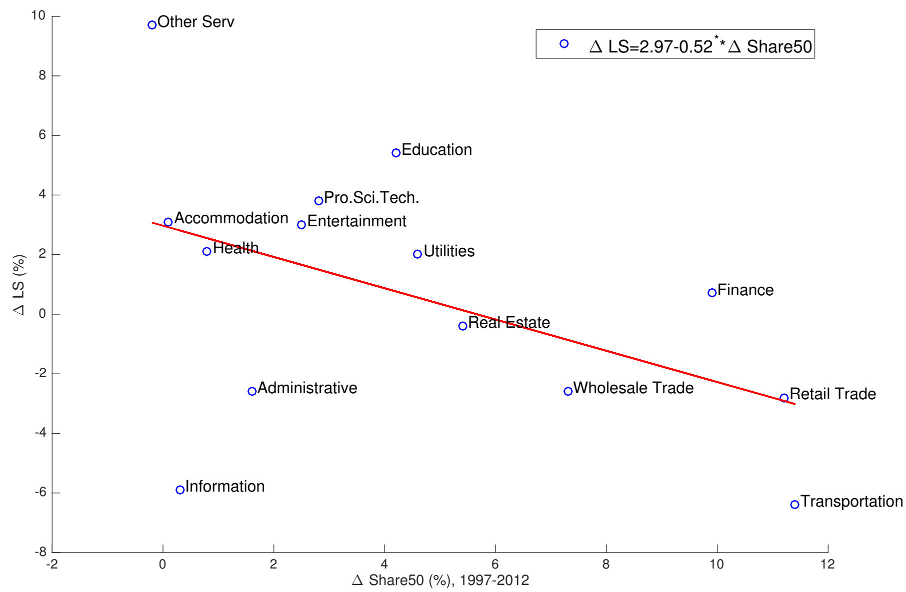


Figure 1.6: Changes in Labor Share versus Change in Concentration, NON-MFG 1997-2012

Note: Labor share (on the vertical axis) is the ratio of employee compensation to value added; concentration (on the horizontal axis) is the revenue share of the 50 largest firms.

firms' share of a sector's revenue[26].

From 1997 to 2012, there is a similar negative association between changes in labor share and changes in concentration. The correlation is significant at the 10% level. Out of Manufacturing, the decline of LS was most evident in transportation, retail trade, and wholesale trade; these are also the subsectors in which concentration increased the largest. The decline of LS in the information sector and in the administrative & waste management sector is probably more reflective of the temporary increase of labor share around 2000 (due, in part, to stock options being realized during the dot-com bubble) than of a long run trend.[27] In

[26] For non-manufacturing sectors, the Economic Census reports concentration in terms of revenue instead of value added. Concentration ratios for the 4, 8, 20 and 50 largest firms in each sector are available. Overall, average concentration ratios for the 4 largest firms at the 6-digit level are comparable to those for 50 largest firms at the 2-digit level.

[27] For the labor share in 2-digit sectors from 1987 to 2015, see Table A.11 and Figure A.7 in the Appendix. Note that the information services sector is a combination of traditional subsectors (e.g. newspaper publish-

most finance and service sectors, labor share actually increased slightly from 1997 to 2012 while increases in concentration were likewise moderate.

For non-manufacturing sectors, relative labor share is approximated by the ratio of large firms' payroll share to their revenue share (instead of value added shares, as in the case of manufacturing sectors). Once again, firms are ranked by revenue.[28] Appendix Table A.13 presents the relative labor share in non-manufacturing sectors at various NAICS digit levels.[29] Just as in manufacturing, the relative labor share for large non-manufacturing firms is less than that for their smaller peers. At the 6-digit level, the average labor share in a sector's 4 largest firms is 18% lower than that for the entire sector whiere a typical firm of the smallest size has a LS that is 20% higher than the sector average.

In wholesale trade, retail trade, and transportation—sectors that saw steep declines in labor share—the relative labor share of large firms also decreased. From 1997 to 2012, the RLS of the 50 largest firms decreased from 39.9% to 26.9% in wholesale trade, from 96.4% to 92.9% in retail trade, and from 103.7% to 80.1% in transportation.

As mentioned previously, the relative labor share can be decomposed into relative labor productivity and relative wage. Table 1.5 presents the relative labor productivity and wage

ing) and othres that are more prone to technological change. In the Broadcasting and Telecommunication subsector (the sector where AT&T belongs to), there is also a declining trend of labor share and rising concentration.

[28] This approximation implicitly assumes that, within a sector, the rank of firms based on value added is the same as that based on sales—especially for the largest firms. It also assumes that the share of value added by large firms is similar to their share of revenue. The first assumption seems reasonable enough; the second can be directly verified from available manufacturing data, which include (for large firm) both share of value added and revenue. Appendix Figure A.6 confirms that there are only small differences between these two measures.

[29] There are several sectors, especially at the 6-digit level, in which the relative labor share exceeds 200% for the "51st largest and smaller" firms. The value added share and payroll share of small firms in these sectors are typically both very small. These sectors are excluded from the analysis.

Table 1.5: Relative Labor Share, Relative Labor Productivity and Relative Wage of Top-50 Firms

	Wholesale Trade			Retail Trade			Transportation		
	R. LS	R. LP	R. Wage	R. LS	R. LP	R. Wage	R. LS	R. LP	R. Wage
1997	39.9%	323.7	129.1	96.4	85.3	82.3	103.7	108.7	112.7
2002	36.4	332.1	120.9	95.2	92.9	88.5	86.6	126.1	109.2
2007	30.8	383.3	117.8	95.4	97.4	93.0	81.9	141.3	115.7
2012	26.9	419.6	112.9	92.9	101.0	93.8	80.1	142.5	114.1
Δ 97-12	-13.0%	95.9	-16.2	-3.5	15.7	11.5	-23.6	32.8	1.6

Note: Relative labor share (resp. labor productivity, wage) in this table is calculated as the ratio of labor share (resp. labor productivity, wage) for the 50 largest firms to the sector average.

for trade and transportation sectors. From 1997 to 2012, the relative labor productivity of large firms in wholesale trade increased by 95.9% while their relative wage *decreased* by 16.2%. In transportation, the 32.8% increase in relative labor productivity is likewise significantly greater than the 1.6% increase in relative wage. This gap is as wide in retail trade as in wholesale trade or transportation, but the increase in relative labor productivity is nonetheless 4.2% higher than relative wage.[30]

In contrast, the relative labor share of large firms in most finance and services sectors does not show a clear trend over the past 15-20 years.[31] These results all continue to hold when, instead of the 50 largest firms in each sector, I use the 20 largest (see Appendix Table A.19) and, for the most part, at more disaggregated levels.[32]

[30]The reason that the relative labor productivity and relative wage of large firms are smaller than 100% in the retail trade sector could be that the top-50 retail firms category is a mix of traditional and new online retailers. In 2012, for example, Walmart was the largest US retailer in terms of revenue and Amazon was ranked 15th. The relative LP for the 4 largest firms in this sector was 96.1% in that year; this percentage was higher (106.6%) for the 5th to 8th largest firms and higher still (108.4%) for the 9th to 20th largest.

[31]See Appendix Table A.18 for details.

[32]See Appendix Table A.20. It is worth emphasizing that results at the 2-digit level are less sensitive to the problem of classifying of multi-establishment firms, changes in industrial code, and reclassification of firms over time. For example, the retail trade sector's increase in concentration is partly caused by the rise of online retailers (e.g. Amazon) that sell goods in various categories. The NAICS 3-digit code for Amazon

1.2.3 The aggregate pattern

The facts presented so far mar be summarized as follows. In manufacturing, trade, and transportation sectors, both the sector labor share and the relative labor share of large firms have been declining. In most services sectors, the labor share has not declined; in these sectors, the relative labor share of large firms does not exhibit an identifiable trend. Finally,concentration has increased across all sectors—but more in manufacturing, trade, and transportation and less in services. Although sector heterogeneity helps to identify the driving forces, this paper seeks to explain the decline of the *aggregate* labor share.[33]

Table 1.6 shows, for the period 1987 to 2012, the economy's labor share as well as the concentration and the relative Labor share for the 50 largest firms.[34] The reported values are averaged across 2-digit sectors, with weight equal to the average value added share of a sector from 1987 to 2012,[35]

is 454 (nonstore retailers) while the code for a typical book store is 451 (sporting goods, hobby, book, and music stores). At the NAICS 2-digit level, both types of stores are grouped in 44-45 (retail trade).

[33]Concentration can increase for reasons that are unrelated to technology. For instance, relaxing the laws that constraint mergers and acquisitions might contribute to the rising concentration in some sectors. This paper does not address the services sector's moderate increase of concentration.

[34]As before, the labor share is not adjusted for proprietor's income—for the sake of consistency with the measure in more dis-aggregated sectors. The choice of 50 is based mainly on data availability. In general, 50 would be too small a number for sectors such as manufacturing (where the 50 largest firms accounted for 26% of 2012 value added) and would be too large a number for sectors such as utilities (where the value added share of the 50 largest firms was 69% in that year). Any choice of a simple cutoff would be forced to make the same trade-off. The relative labor share of top-50 firms in administrative & support and waste management & remediation sector inexplicably rose from 109% in 2007 to 125% in 2012. The value for "RLS-Top50" in 2012 would be 75.86 (rather than 73.37) if the 2007 value were used for this sector.

[35]Recall that the industrial classification standard changed in 1997 from SIC to NAICS. In 1987 and 1992, both concentration and RLS are available for manufacturing, retail trade, and wholesale trade. For 2-digit sectors in transportation, finance, and services, I first check for whether concentration and and relative labor share exhibited any linear trend from 1997 to 2012. When there was a linear trend (significant at the 10% level), I use that trend to obtain values for 1987 and 1992; When there was no significant trend, I use average values from 1997 to 2012 as an approximation for 1987 and 1992.

Table 1.6: Economy-wide Labor Share, Concentration, and Relative Labor Share of Top-50 Firms

Year	1987	1992	1997	2002	2007	2012
LS	51.83%	51.48	50.49	51.21	50.30	49.00
Share50	26.00%	26.57	26.69	29.52	30.51	31.31
RLS-Top50	82.20%	81.23	79.37	78.20	74.86	76.37

Note: Labor share is calculated as the fraction of compensation in value added. Agriculture, government, mining, construction, and management of companies & enterprises are excluded in the statistics.

Table 1.7: Decomposition of Declines in Labor Share, 1987-2012

Δ LS	*Increase in concentration*	*Fall in large*	*Fall in small*
-2.83%	-0.80%	-1.48%	-0.57%
% contributed	28%	52%	20%

Note: See text for details.

Note that change in LS can be decomposed as

$$\text{LS} = \Sigma_{i=l,s}\text{LS}_i\omega_i \implies \Delta\text{LS} \approx \underbrace{\Delta Concentration * (\text{LS}_l - \text{LS}_s)}_{\text{increase in concentration}} + \underbrace{\omega_l * \Delta\text{LS}_l}_{\text{fall in large}} + \underbrace{\omega_s * \Delta\text{LS}_s}_{\text{fall in small}}$$

where l and s denote (respectively) large and small firms. The decline in aggregate LS stems from three sources: decline of labor share in large firms, decline of labor share in small firms, and increasing concentration. From 1987 to 2012, the economy wide labor share declines 2.83%. During the same period, economy-wide concentration increased by 5.31% while the average difference between the LS of large firms and small firms was 15.12%. The first term, "increase in concentration", is responsible for 0.80% of the decline in labor share; the respective contributions of the second and the third components can be calculated similarly. Table 1.7 presents the decomposition results.

The labor share's decline is due mainly to the fall of labor share in large firms (which accounts for 52% of the aggregate decline) and the increasing concentration (which accounts for another 28%); the contribution of "fall in small" firms is only 20%. It is important to bear two facts in mind. First, the labor share in small firms is 55.07% in 1987 and 54.28% in 2012. Its contribution would be smaller if the cutoff for "large" firms were relaxed to include more than 50 firms. Second, the labor share in small firms was 55.85% in 2007. If this value is used (rather than the 2012 value), then the "fall in small" term actually increases the aggregate labor share. As a result, in this paper I focus on the first two components.

Large firms in the the 2010s may little resemble large firms in the 1980s. According to *Compustat* firm-level data, a large portion of LS declines occurs within the same firms in the manufacturing and transportation sectors.[36] In both wholesale and retail trade, the decline seems to come mainly from the creation of new firms (e.g. online retailers) that have lower labor shares and grow large over time. In both cases, large firms in the 2010s had both lower labor shares and larger market shares as compared with their counterparts in the 1980s. Each of these trends is associated with declining aggregate labor share.

1.3 Model

In this section I develop a model that rationalizes the empirical facts. As noted earlier, my model builds on two assumptions: (i) for a given technology (embodied in machines), capital and labor are complementary inputs; and (ii) technological progress is labor saving. Here I begin, in Section 3.1, by using a static model to illustrate the effect of technologies on size (i.e. output) and labor share—namely, more advanced technologies result in larger sizes and

[36] See Appendix Figures A.8 and A.9 for the labor share (measured as the fraction of compensation in revenue) in large manufacturing and transportation firms.

smaller labor shares. Then, in Section 3.2, I add capital accumulation and heterogeneous firms and extend the static model into general equilibrium. It is established there that the introduction of new technologies increases concentration and also reduces the labor shares of large firms. As a result, the aggregate labor share declines.

1.3.1 Static model

There are N vintages of capital, each of which embodies a distinct generation of technology. Denote by $j = 1, 2, ..., N$ the vintage of capital. In the static model, a technology is interpreted as a firm. Technology j combines capital j and labor to produce a single final goods:

$$[(1-\alpha)k_j^\rho + \alpha(\gamma_j \ell)^\rho]^{\frac{1}{\rho}}$$

here γ_j denotes the level of technology embodied in capital j. The assumptions I make are expressed formally as follows.

Assumption 1. [Capital-labor complementarity] $\rho < 0$.

Assumption 2. [Labor-saving technological progress] γ_j *increases with* j.

The elasticity of substitution between capital and labor is given by $\frac{1}{1-\rho}$. The first assumption states that, given a technology (embodied in machines), the elasticity of substitution between capital and labor ranges between 0 and 1; that is, *capital and labor are complementary inputs*. This assumption is consistent with most empirical estimates (e.g. Antras (2004); Klump et al. (2007); Herrendorf et al. (2015)). Recently, Oberfield and Raval (2014)) use plant-level data from the Census of Manufactures to estimate an average (plant-level) elasticity of substitution of about 0.5 for 1987. The estimated aggregate elasticity in manufacturing is 0.71 in 1987 and 0.75 in 2007.

Assumption 2 states that *technological progress is labor saving* in the sense that new technology embodied in new machines requires less labor input per unit of output. Labor-augmenting technological progress is typically assumed in growth models to be consistent with a balanced growth path (Barro and Sala-iMartin (2004). see Acemoglu (2003) and Jones (2005) for theoretical justifications).

Let k_j denote the supply of capital j and let L denote the supply of labor. In order to isolate the effect of technology in the static model, all the k_j are fixed at 1 and the inelastic labor supply is also normalized to 1. Labor moves freely among firms. I study the labor allocation problem and investigate the effects of technologies on firm size and labor share. The marginal productivity of labor in firm j is

$$\mathrm{MPL}_j = \left[(1-\alpha)\left(\frac{k_j}{\ell}\right)^{\rho} + \alpha\gamma_j^{\rho}\right]^{1/\rho-1} \alpha\gamma_j^{\rho}.$$

When the employment in firm j approaches zero, the firm's marginal productivity of labor is

$$\mathrm{MPL}_j(0) = \alpha^{1/\rho}\gamma_j.$$

Since capital is fixed at a positive number, it follows that the marginal productivity of labor at zero employment is not equal to infinity (i.e. the Inada condition does not hold). Hence some firms might not hire any labor in equilibrium. In addition, firms that use more advanced technologies have a higher marginal productivity of labor at zero employment. The employment flow always begins with firm $j = N$ and moves downwards, step by step. Thus the first unit of labor goes to the most productive firm, N. As firm N accumulates labor, its marginal productivity of labor declines. As soon as that level falls to the *second* advanced firm's marginal productivity of labor (at zero employment), that second firm begins to hire labor. This process continues until full employment is reached.

The labor market equilibrium condition is[37]

$$w = \mathrm{MPL}_j = \left[(1-\alpha)\left(\frac{k_j}{\gamma_j \ell_j}\right)^{\rho} + \alpha\right]^{1/\rho-1} \alpha\gamma_j.$$

As a result, technologies that are more advanced (i.e., with a higher γ_j) will cause a decline in the adjusted capital/labor ratio, $k_j/\gamma_j\ell_j$.[38] The labor share in firm j is

$$\mathrm{LS}_j = \frac{w\ell_j}{y_j} = \left(\frac{1-\alpha}{\alpha}\left(\frac{k_j}{\gamma_j \ell_j}\right)^{\rho} + 1\right)^{-1}.$$

Because the labor share is an increasing function of $k_j/\gamma_j\ell_j$, firms that use more advanced technologies have a lower labor share.

Technology also affects firm size, which is (as in the concentration data) equal to value added *divided by* revenue. Combining the formula for firm size and the labor market–clearing condition now yields the following expression:

$$y_j = k_j \left[\frac{1-\alpha(w/\alpha\gamma_j)^{-\rho/(1-\rho)}}{1-\alpha}\right]^{-1/\rho}.$$

According to this formula, output is greater if the value of γ_j is higher. In other words, a more advanced technology increases firm size. Therefore, the single parameter of technology (γ_j) is enough to generate a negative correlation—as observed in the data—between firm size and labor share. Formally, the following proposition holds.

[37]This condition holds only for firms that have positive employment in equilibrium.

[38]I remark that the true capital labor ratio, k_j/ℓ_j, can be either increasing or decreasing in γ_j. The direct effect of technologies with a higher γ is to substitute raw labor. Yet that technology also increases the productivity of capital, which in turn increases labor demand. The equilibrium k_j/ℓ_j depends on which effect dominates (see the appendix for a detailed discussion). In Section 3.2, I show that, if capital is adjustable, then the capital/labor ratio is a positive function of the technology parameter, γ.

Proposition 1. [Effects of technology on firm size and labor share]

- *If* $j > j'$, *then* $\mathrm{LS}(j) < \mathrm{LS}(j')$; *that is, firms that use more advanced technologies have a lower labor share.*

- *If* $j > j'$, *then* $y(j) > y(j')$; *that is, firms that use more advanced technologies produce more output.*

To see the intuition behind this result, recall that technology (γ) both complements capital and increases the productivity of capital, which further increases demand for effective labor ($\gamma\ell$). Hence firms produce more when they use more advanced technologies; however, technology is a substitute for raw labor and so reduces labor's share of income.

1.3.2 Dynamic general equilibrium

It is now possible to incorporate capital accumulation and heterogeneous firms into the model just developed. The goal here is to develop a model that can be used to study how concentration and labor share are affected by the creation of new technologies.

Firms are introduced in the tradition of "span of control" models (cf. Lucas (1978)) but without their career choice component. Production requires three inputs: capital and labor (as before) and also entrepreneurial skill. Hereafter I shall use the terms productivity and entrepreneurial skill interchangeably and without prejudice. There is a continuum of firms $i \in [0, 1]$, each endowed with some productivity. Firm i draws its productivity z_i from the following Pareto distribution[39] (at the beginning of time):

$$z_i \sim f(z) = \begin{cases} \lambda/z^{\lambda+1} & \text{if } z \geq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Firms optimally choose to adopt one among N technologies, which are embodied in different machines. A firm i that adopts technology j thereby accesses the production function

$$y_i(j) = z_i^{1-\eta}[(1-\alpha)k(j)^\rho + \alpha(\gamma(j)\ell)^\rho]^{\eta/\rho},$$

where η ($0 < \eta < 1$) is the span-of-control parameter. Note that the time subscript t has been omitted. To simplify analysis, I assume the following structure of capital. The household supplies and accumulates what I call the *general* capital, which firms purchase at a common interest rate and convert into capital of vintage j at some cost. One unit of the general capital can be converted into $1/q(j)$ units of vintage-j capital. In addition, firms are able to adopt technology j—or, equivalently, to use capital of vintage j—only by first paying a lump-sum fixed cost $\phi(j)$. My last assumption is formalized next.

Assumption 3. *Both $\phi(j)$ and $q(j)$ are increasing in j; that is, more advanced technologies require a larger fixed cost. Also, machines that embody more advanced technologies are more costly to produce.*

Firms optimally choose technology j, and employ capital $k_i(j)$ and labor ℓ_i, while taking wage and the interest rate as given. If none of the N technologies generates a net positive

[39] Axtell (2011) documents that the distribution of firm size is well approximated by a Pareto distribution, as employed by Buera et al. (2011b) and Dinlersoz and Greenwood (2016a).

profit, then firms will be inactive. Firm i's optimal choice problem is written as[40]

$$\Pi_i \equiv \max\Big\{ \max_{j,k_i(j),\ell_i} y_i(j) - r(q(j)k_i(j)) - w\ell_i - \phi(j),\, 0 \Big\}.$$

For future reference, I define an indicator $\sigma_i(j)$ as

$$\sigma_i(j) = \begin{cases} 1 & \text{if firm } i \text{ adopts technology } j, \\ 0 & \text{otherwise.} \end{cases}$$

Note that if firm i chooses to remain inactive and so does not adopt any of the N technologies, then $\sigma_i(j) = 0$ for all j.

There exists a representative household that accumulates the general capital and also inelastically supplies L units of labor to maximize present-value utilities:

$$\sum_{t=0}^{\infty} \beta^t \log C(t);$$

here β is the discount factor and C denotes consumption. The household obtains income from wages, rental income, and profits, and it distributes total income into consumption C and investment I. Its budget constraint is

$$C(t) + I(t) \leq w(t)L + r(t)K(t) + \int \Pi_i(t)\, dF(z_i),$$

where $F(z_i)$ is the cumulative distribution function of z_i. In addition, the household respects

[40]Firms can freely switch technologies from period to period. The model does not incorporate firms' growth because I am focusing instead on how labor share and firm size distribution are affected by technology.

the following law of motion for the general capital:

$$K(t+1) = (1-\delta)K(t) + I(t);$$

here δ denotes the depreciation rate.

The model economy's competitive equilibrium is defined as a sequence of prices $\{r(t)\}_{t=0}^{\infty}$ and $\{w(t)\}_{t=0}^{\infty}$ and a sequence of aggregate quantities $\{C(t)\}_{t=0}^{\infty}$ and $\{K(t)\}_{t=0}^{\infty}$—as well as, for all i, technological adoption decisions $\{\sigma_i(j,t)\}_{t=0}^{\infty}$ and demand for capital $\{k_i(j,t)\}_{t=0}^{\infty}$ and labor $\{\ell_i(t)\}_{t=0}^{\infty}$—such that the following statements hold.

1. Given prices, $\{C(t)\}_{t=0}^{\infty}$ and $\{K(t)\}_{t=0}^{\infty}$ maximize the representative household's utility.

2. Given prices, the technology choices $\sigma_i(j,t)$ and factor demands $k_i(j,t)$ and $\ell_i(t)$ maximize firms' profits for all t.

3. Markets clear:

    - capital market,

    $$\int \Sigma_j k_i(j,t)\sigma_i(j,t)q(j)\, dF(z_i) = K(t) \quad \forall t;$$

    - labor market,

    $$\int \ell_i(t)\, dF(z_i) = L \quad \forall t;$$

    - goods market,

    $$C(t) + K(t+1) - (1-\delta)K(t) + \int \Sigma_j \sigma_i(j,t)\phi(j)\, dF(z_i) = \int y_i(t)\, dF(z_i) \quad \forall t.$$

Technology, labor share, and firm fize The static model showed that more advanced technologies lead to lower labor shares and higher output. Those findings apply also in this extended model. To reduce notation, write the conversion cost $q(j)$ as a function of $\gamma(j)$—thus, q_γ. The production function of a firm i that uses capital of vintage j is

$$y_i = z_i^{1-\eta}\left[(1-\alpha)\left(\frac{k_i}{q_\gamma}\right)^\rho + \alpha(\gamma\ell_i)^\rho\right]^{\eta/\rho};$$

here k_i is the general capital which commands a common interest rate r. From the first-order conditions of firm $i's$ optimization it follows that the capital intensity (K/L)[41] in firm i can be written as

$$\frac{k_i}{\ell_i} = \left(\frac{1-\alpha}{\alpha}\frac{w}{r}\right)^{1/(1-\rho)}(\gamma q_\gamma)^{-\rho/(1-\rho)}.$$

In other words, firms that use more advanced technologies have a higher capital/labor ratio. The labor share in firm i is

$$\mathrm{LS}_i \equiv \frac{w\ell_i}{y_i} = \frac{\eta\alpha}{\alpha + (1-\alpha)\left(\frac{r}{w}\frac{\alpha}{1-\alpha}\right)^{-\rho/(1-\rho)}(\gamma q_\gamma)^{-\rho/(1-\rho)}}.$$

Thus a more advanced technology (i.e., a higher value of γ) results in a lower labor share. All firms face the same wage and so labor productivity, defined as

$$\mathrm{LP}_i \equiv \frac{y_i}{\ell_i} = \frac{w}{\mathrm{LS}_i},$$

is an increasing function of γ. Observe that all these three properties are independent of firm productivity z_i.

[41]The capital used in measuring capital intensity is the general capital. In the data, capital stock's value already contains price information about different machines that reflects quality differences.

Firm $i's$ output y_i is[42]

$$y_i = z_i \underbrace{\left(\frac{\eta(1-\alpha)}{r}\right)^{\eta/1-\eta} q_\gamma^{-\eta/1-\eta}\left[(1-\alpha)+\alpha\left(\frac{r}{w}\frac{\alpha\gamma q_\gamma}{1-\alpha}\right)^{\rho/1-\rho}\right]^{\eta(1-\rho)/\rho(1-\eta)}}_{\equiv g(\gamma,r,w)} \tag{1.1}$$

The effect of technology on firm size, measured as output, has two aspects. On the one hand, more advanced technology increases demand for capital and effective labor and also increases firm size, as in the static case; on the other hand, the conversion cost makes it optimally for firms to cut back on their use of not only labor but also capital.

Firms' technology adoption decision Firm $i's$ profit is

$$\begin{aligned}\Pi_i &= y_i - w\ell_i - rk_i - \phi \\ &= (1-\eta)y_i - \phi \\ &= (1-\eta)z_i g(\gamma, r, w) - \phi\end{aligned}$$

where $g(\gamma, r, w)$ is defined in the firm size formula (1). A nice property is that profit Π_i is a linear and increasing function of productivity z_i. If $g(\gamma, r, w)$ is an increasing function of γ then the profit function, as a function of productivity z_i, has a smaller intercept $(-\phi)$ and a larger slope $(g(\gamma))$ for technologies that are more advanced. In this case, the more productive firms optimally choose more advanced technologies. Formally, we have the following proposition.

Proposition 2. [Firms' optimal technology adoption] *Let $g'(\gamma) > 0$ for $g(\gamma)$ as defined in equation* (1)*. If it is optimal for firm i with productivity z_i to adopt technology j, then firm i' with productivity $z_{i'} > z_i$ adopts technology $j' \geq j$.*

[42] See the Appendix for an expression for capital and labor demand in firm i. Substituting capital and labor demand into the firm's production function gives this result.

Figure 1.7 illustrates the intuition by showing an example case of three technologies, where technology 1 is the least advanced and technology 3 the most advanced. This figure's plot of the profit function associated with technology 1 starts high and increases slowly, whereas the technology 3 function starts low and increases rapidly. Let $\Pi(z_i, \gamma_j)$, $j = 1, 2, 3$, denote the profit function when the firm adopts technology j. There are three intersection points: $\bar{z}_1$, where $\Pi(z_i, \gamma_1)$ intersects the zero-profit line; $\bar{z}_2$, the intersection of $\Pi(z_i, \gamma_2)$ and $\Pi(z_i, \gamma_1)$; and $\bar{z}_3$, the productivity level at which $\Pi(z_i, \gamma_3)$ surpasses $\Pi(z_i, \gamma_2)$. Firms' technology adoption therefore follows a threshold rule:[43]

$$T_i = \begin{cases} \text{stay inactive} & \text{if } z_i < \bar{z}_1, \\ \text{adopt technology 1} & \text{if } z_i \in [\bar{z}_1, \bar{z}_2), \\ \text{adopt technology 2} & \text{if } z_i \in [\bar{z}_2, \bar{z}_3), \\ \text{adopt technology 3} & \text{if } z_i \geq \bar{z}_3. \end{cases}$$

A sufficient condition for $g'(\gamma) > 0$ is that the conversion cost q_γ being a constant. In this case, more advanced technology increases firms' output. The intuition derives from the static case: advanced technology (γ) complements capital and increases employment of effective labor ($\gamma \ell$)[44]. By continuity, , $g'(\gamma) > 0$ holds as long as $q'(\gamma)$ is small enough. In the quantitative analysis, I always choose the conversion cost function such that $g'(\gamma) > 0$. This procedure rules out the uninformative case in which new machines cost so much that *no* firms will adopt them.

Firms that are more productive choose to adopt more advanced technologies. Because

[43]It is possible for (say) ϕ_3 to be slightly smaller than ϕ_2 and hence for the inequality $\bar{z}_3 < \bar{z}_2$ to hold. This situation is equivalent to the case of only two available technologies. I do not consider that possibility because the discussion of N technologies is sufficiently general.

[44]This statement follows also from equation (1) by putting $q_\gamma = c$.

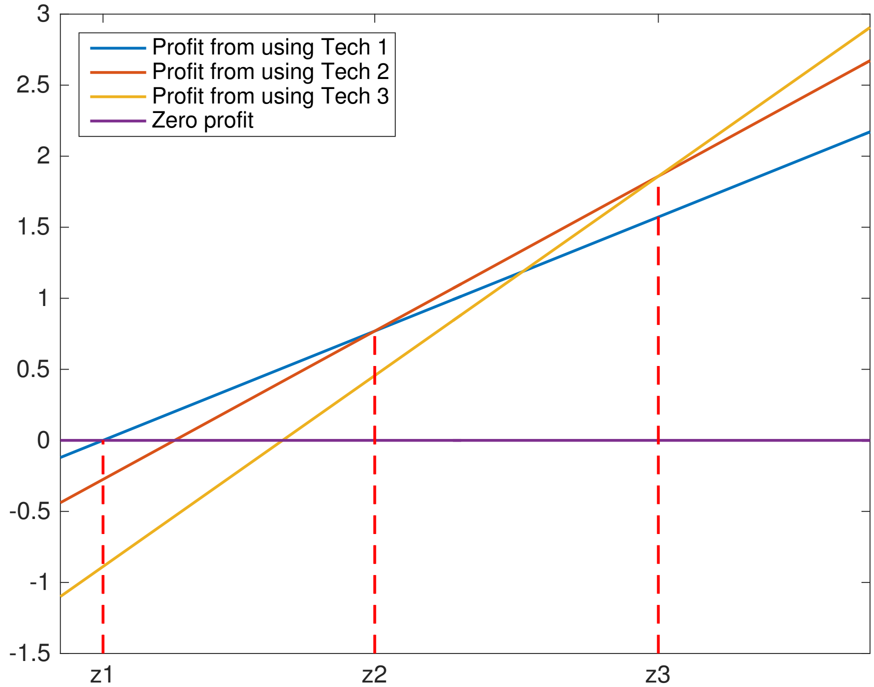


Figure 1.7: Profit Functions and Technology Adoption

Note: This graph shows profit functions in an example case involving three technologies. Technology 1 (resp. 3) is the least (resp. most) advanced. The horizontal axis marks firm productivity.

firm size (as measured by output or value added) is an increasing function of both productivity and technology, firms with higher productivity are also larger. Recall from the static case that more advanced technologies induce higher labor productivity and lower labor share. Hence the extended model predicts a negative relation between firm size and labor share.

Technological change In an economy with N technologies and in which capital and labor are complementary while labor is fixed, the economy will arrive at a steady state. In that steady state, productive firms choose more advanced technologies, are larger, and have a lower labor share. Technological change is modelled as an (exogenous) arrival of the $(N+1)$th technology. This new technology drives the economy to a new steady state. In this subsequent steady state, the most productive firms find it profitable to switch to the cutting-edge technology, and become even larger (which increases concentration) and have lower labor shares than before. As a result, the aggregate labor share declines.[45] To demonstrate this effect more clearly, I next calibrate a version of the model initially with two technologies and study how concentration and labor shares are affected by the introduction of new, third technology.

1.3.3 Quantitative evaluation

This section offers a quantitative evaluation of the mechanism proposed in the paper. The economy initially has two vintages of technology; technological progress is modeled as a third and more advanced technology being exogenously available. I begin with just two vintages of technology since it corresponds to the binary division of firms into large and small ones, for which data is available for calibration; two is also the minimum number needed to demonstrate the mechanism.

[45] There will be some equilibrium effect on wage owing to the availability of new technology and associated capital accumulation. I show in Section 3.3 that this equilibrium effect is small.

The model has five parameters in preference and production functions: the discount rate β, depreciation rate δ, span-of-control parameter η, elasticity of substitution ρ between capital and labor, labor weight α in the production function, and one-tail parameter λ in the Paretro distribution of productivity. The model also has five technology and cost parameters, of which the first four are the levels γ_1 and γ_2 of labor-saving technology and the two fixed costs ϕ_1 and ϕ_2. The conversion cost of capital j, denoted $q(j)$, is assumed to be a power function of $\gamma(j)$: $q(\gamma) = \gamma^{\varepsilon}$. The fifth parameter is ε, the conversion cost parameter. The fixed labor supply is normalized to 1.

The discount rate β and depreciation rate δ are widely used in the macroeconomics literature. I choose $\beta = 0.96$ to match an annual interest rate of 4%, and the discount rate is set at $\delta = 6\%$ per year. In the model, the span-of-control parameter η determines the share of profit (which is usually considered to be part of capital income) in firms' value added. I pick $\eta = 0.75$, which corresponds to 25% of profit share.[46]

As for the elasticity of substitution between capital and labor—$\rho/(1-\rho)$ in my model—most empirical estimates obtain values that are less than 1. Using micro-level Census of Manufactures data and a CES production function, Oberfield and Raval (2014) estimate the average plant-level elasticity of substitution to be 0.5 and the aggregate elasticity of substitution for the manufacturing sector to be 0.71 in 1987 and 0.75 in 2007. I target an elasticity of substitution of 0.5 at the firm/plant level and set $\rho = -1$.

In a 2-digit sector, 50 firms typically account for only a small fraction of the total number

[46]This value for η is slightly smaller than the 0.85 typically used in literature (e.g., Atkeson and Kehoe (2007); Midrigan and Xu (2014a)). A slightly larger profit share is targeted because, in my model, firms must pay fixed costs out of their profit.

of firms in that sector. Large firms typically have multiple establishments, but the model presented here does not distinguish between firms and establishments. I implement the following adjustment procedure. Business Dynamics Statistics (BDS) data classify firms into "bins" of different sizes; size is measured by number of employees, so those bins range from "1 to 4 employees" to "more than 10,000 employees". I use the average number of establishments for firms in the largest-size bin to approximate the number of establishments in the 50 largest firms and then calculate their fraction in the total establishments for each sector. This sector-level fraction is summed up to obtain the economy-wide values. Calculated this way, the 50 largest firms accounted for 1.93% of all establishments in 1987.

According to Table 1.6, top-50 firms account for 26% of revenue share in 1987. Concentration in the mode is a combination of two forces: higher productivity which is governed by the tail parameter λ; and more advanced technology, which is determined by technology γ and conversion costs ε. We do the following to to back out the tail parameter λ: Table 1.6 shows a negative correlation between labor share and concentration over time. In particular, results from a single variable regression is

$$Share50 = 115.6 - 1.72 * LS$$

The labor share is of course affected by technological heterogeneity. In 1987, the relative labor share of top-50 firms is 82.2%. Based on information in Table 1.6, a hypothetical labor share in the case where there is no technological heterogeneity (i.e. RLS of top-50 firms is 100%) can be calculated. We then combine this hypothetical labor share and the relation above to obtain a measure of concentration where there is no technological heterogeneity among firms. The resulted ratio is 20.9%, i.e. the 50 largest firms (or 1.93% of all firms) account for 20.9% of value added. These yield the tail parameter of the Pareto distribution,

$\lambda = 1.66$.

There are still six parameters to be determined. labor weight in the production function, α, technology parameter γ_1 and $_2$, fixed costs ϕ_1 and ϕ_2, and conversion cost parameter, ε. For technology 1, the level of labor saving technology is normalized to $\gamma_1 = 1$. Since technology affects both labor share and size, the relative labor share and concentration data in 1987 are chosen as two moments to target in the calibration. The labor weight α affects labor share across all firms and thus the aggregate labor share, which is used as a third moment. A 5% exit rate for firms or establishments is used as the fourth moment since exit is a function of fixed costs[47]. A higher conversion cost reduces demand for the general capital; hence the rate of net labor-saving technological progress, which reduces equilibrium employment, is amplified by the conversion cost parameter ε.[48] The last moment used is concentration of employment in top-50 firms in 1987[49]. The five parameters are jointly calibrated to match these five moments.[50] Table 1.8 summarizes the calibration results.

Under these parameter values, the aggregate labor share in the model economy is 51.72% and top-50 firms accounts for 26.07% of total revenue and 25.09% of total employment. In

[47]The establishment exit rate, calculated using BDS data, was 11.9% in 1987. Because the model does not have firm entry and exit, I choose a more conservative target of 5% inactive firms.

[48]To see this, note that the production function in terms of the general capital is

$$y_i = z_i^{1-\eta}\left[(1-\alpha)\left(\frac{k}{q_\gamma}\right)^{\rho} + \alpha(\gamma\ell)^{\rho}\right]^{\eta/\rho}.$$

given that $q_\gamma = \gamma^\varepsilon$, the rate of labor-saving technological progress is $(1+\varepsilon)\dot{\gamma}/\gamma$.

[49]The employment share of top-50 firms is an weighted average of this share across 2-digit sectors. At the two digit level, the share is interpolated for values in 1987 and 1992 in some sectors. The same interpolation method as in constructing the revenue share is employed. This share for the aggregate economy is 20.00% in 1987, 19.85% in 1992, 19.63% in 1997, 21.55% in 2002, 20.66% in 2007, and 20.90% in 2012.

[50]An equal weights on all five moments would give a relatively large value for the conversion cost parameter ε; Due to the convexity of the cost function, a large ε would greatly dampen technology adoption of firms when a new technology becomes available. Generally, the value ε needs to be high in order to match the employment share and the relative labor share of large firms. To avoid the unwanted case where new technology are so expensive that no firms would like to adopt them, I choose to assign a lower weight for these two moments in the calibration.

Table 1.8: Summary of Calibration Results

Para.	Meaning	Values	Target/sources
β	discount rate	0.96	4% interest rate
δ	depreciation rate	0.06	6% capital depreciation
η	span-of-control	0.75	25% profit share
ρ	elasticity of substitution, K and L	-1	Oberfield and Raval (2014)
λ	shape of Pareto distribution	1.66	20.9% in tail
γ_1	technology para. in Tech. 1	1	normalization
jointly calibrated parameters			
α	labor weight in prod. fun.	0.44	labor share in 1987
γ_2	technology para. in Tech.	1.5	relative LS-*top-50* in 1987
ϕ_1	fixed cost of Tech. 1	0.16	5% exit rate
ϕ_2	fixed cost of Tech. 2	0.45	revenue share-*top-50* in 1987
ε	power in conversion cost fun.	0.9	emp. share -*top-50* in 1987

equilibrium, 5.2% of firms/establishments stays inactive. The equilibrium wage is 1.11, and the two productivity cutoffs are 1.03 and 2.57. Equivalently: 4.8% of firms exit, 74.3% of (small) firms adopt technology 1 and have a labor share of 55.77%; and the remaining 20.9% of (large) firms adopt technology 2 and have a labor share of 49.77%.

As mentioned earlier, technological progress is modeled as the exogenous availability of a more advanced technology that is more labor saving. That is, this new and more advanced technology has a higher value of γ_3. Furthermore, The fixed cost ϕ_3 associated with this new technology is also larger. We choose the technology parameter γ_3 and the fixed cost ϕ_3 such that in equilibrium, exactly top-50 firms choose to adopt this new technology, and the increase in their concentration ratio matches what is observed in data.

When this more advanced technology becomes available, the most productive firms (*top-50* firms in this case) optimally switch to the new technology, which reduces the labor share in these firms to 45.67%. The concentration ratio, measured as the value-added share of top-50 firms, increases from 26.07% to 31.36%; at the same time, the aggregate labor share declines from 51.72% to 50.35%, or 48.4% of the decline as observed in the data.[51] Note that there is an equilibrium effect when a more advanced technology arrives: capital accumulates and so the wage's absolute level might rise.[52] Under the calibrated parameter values, the wage increases, and this slight increase accounts for the reduced output and also for the increase in small firms' labor shares. The increase in equilibrium wage also induces more firms to exit. In the new steady state, 9.1% of firms choose to exit—as compared with 4.8% before introduction of the new technology.

Large firms in the model grow by adopting labor saving technologies. One implication of the model is that the increase in employment concentration will be slower than the increase in revenue concentration. In the data, the employment share of top-50 firms has increased 0.90% from 1987 to 2012, while their revenue share increases 5.31% for the same period.[53] While the revenue increase is a target in the model, the implied increase in the employment share in the model increases 3.43%, much smaller than the increase in revenue share.

1.4 Discussions and Further Evidence

This section addresses three related issues. First, I calculate the return to capital using NIPA data and while explicitly accounting for changes in the relative prices of investment goods. Second, I present the evolution of sector labor productivity and show that the (labor-

[51]The facts presented in the empirical section show that technological progress, particularly labor saving technological progress, is of first order importance in driving up concentration. It is possible that concentration might increase for non-technological reasons. I tried a more conservative calibration where 50% of increase in concentration observed in data from 1987 to 2012 is caused by technological change. In that case, the technological channel accounts for 21.2% of the decline of labor share for the same period.

[52]The new technology is also a substitute for labor and therefore reduces wages. How the equilibrium wage changes depends on which of these effects dominates.

[53]Even though not focusing on concentration, Kehrig and Vincent (2017) use census data and reports that hyper productive manufacturing firms increases their revenue while maintain their employment relatively unchanged.

saving) technological progress that lowers the labor share also raises labor productivity. Third, because my arguments rely on the assumption of technological heterogeneity across firms, I document (using Compustat data) the heterogeneity of capital intensity at the firm level.

1.4.1 The return to capital

With a production function where capital and labor are complementary, a declining rate of return to capital would encourage capital accumulation and increase the labor share. It has been well documented that the relative price of investment goods (i.e., the price deflator of investment goods *divided by* the price deflator of non-durable consumption goods and services) has declined since 1950s. Yet changes in the price of investment goods reflect only the capital gain component of investment, so the return to capital is equal to the capital gain *plus* the real return. Gomme et al. (2011) measures the return to capital as

$$R_t = \frac{\text{After-tax capital income at } t}{\text{Capital stock at } t} + \frac{\text{Relative price of investment goods at } t}{\text{Relative price of investment goods at } t-1} - 1;$$

here capital stock excludes housing, and capital income (calculated using NIPA data) equals after-tax nonlabor income. The second term captures the capital gain of investment—that is, changes in the relative prices of investment goods over time. Figure 1.8 presents the (annualized) results from Gomme et al.. The return to capital fluctuates around a value of 5.16% and shows no declining trend since the 1980s.

1.4.2 Labor productivity

According to the model, new (labor-saving) technologies that decrease the labor share simultaneously increase labor productivity. Labor share declines occurred mostly in the sectors of

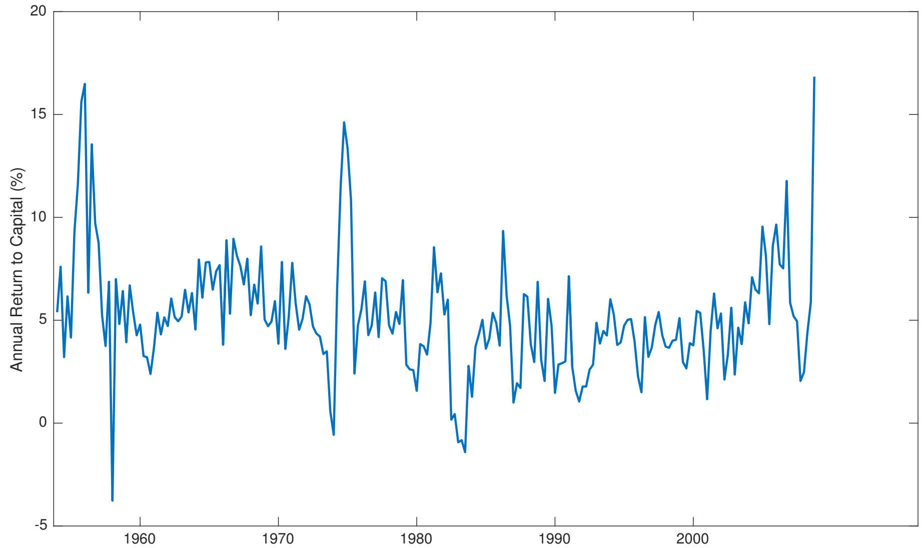


Figure 1.8: Return to Capital, 1954-2008

Note: The return to capital is equal to the real return plus capital gains.
Data Source: Table 2 in Gomme et al. (2011).

manufacturing, trade, and transportation. A greater increase in labor productivity should also be observed in these sectors. Note that there need not be any ex ante relation between labor share (WL/PY) and labor productivity (Y/L). For example, in a Cobb-Douglas production function, irrespective of changes in labor productivity, the labor share is a constant. On the empirical side, labor productivity has been increasing for hundreds of years, while the decline of labor share is a fairly recent phenomenon.

The Labor Productivity and Costs program of Bureau of Labor Statistics provides data on labor productivity for different sectors. Labor productivity is the ratio of real output (net of price change) to hours of labor input.[54] Figure 1.9 plots labor productivity for the

[54]BLS uses different output concepts to measure labor productivity. For the non-farm business sector, real output is measured net of inter-industry transactions and is equivalent to value added. For manufacturing, retail trade and wholesale trade, what is employed is sector output which is total output minus intra-industry transactions. intermediates goods are not included value added, but included in sector output. We

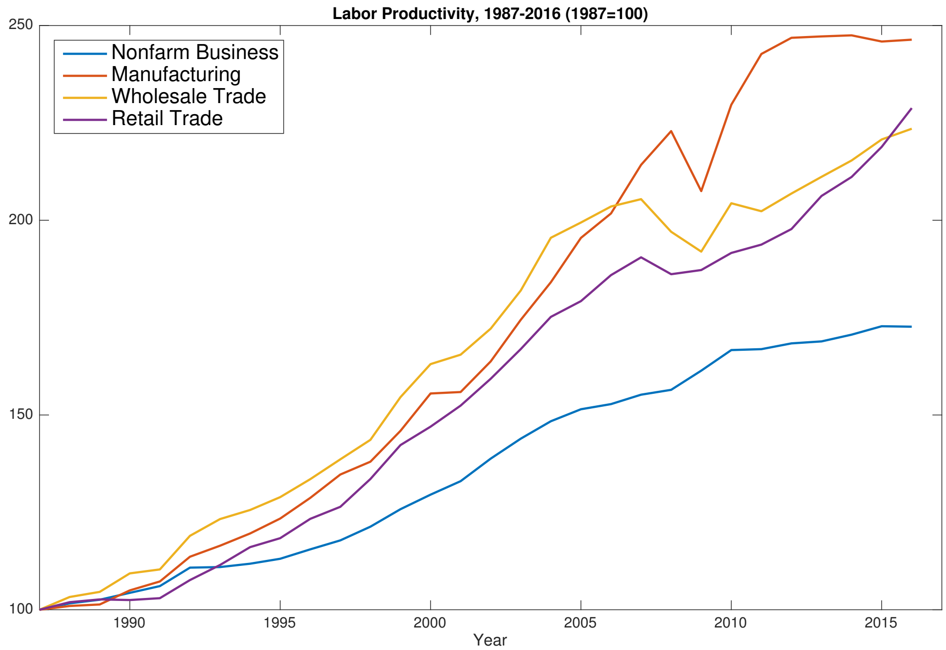


Figure 1.9: Labor Productivity, 1987–2016 (1987=100)

Note: Labor productivity is the ratio of real value added to total working hours.

Data Source: Bureau of Labor Statistics, Labor Productivity and Costs program.

overall economy (nonfarm business sector) and also for the manufacturing, wholesale trade, and retail trade sectors from 1987 to 2016 (the 1987 values are normalized to 100).[55]

Economy-wide labor productivity increased from 100 in 1987 to 172.7 2016—a 72.7% rise. Increases for the same period in the manufacturing, wholesale trade, and retail trade

can calculate the share of out-of-sector intermediate goods in a sector's output from input-output tables in order to evaluate the potential bias. For manufacturing, this share was 26.30% in 1997 and 29.50% in 2016; for retail trade, this ratio was 30.95% in 1997 and 36.16% in 2016; the same share increases slightly from 27.56% to 27.82 in the wholesale trade sector from 1997 to 2016. These relatively small changes suggest that the difference in output measures is not likely to be the driver for the much larger divergence in labor productivity.

[55]Prior to 1987, labor productivity data were not available for different industries. The BLS Labor Productivity and Cost by Industry tables do not include labor productivity data for the transportation & warehousing (NAICS 48–49) sector. The message from subsectors is mixed: labor productivity in air transportation (NAICS 481) increased 134% from 1987 to 2016; the increase in line-haul railroads (NAICS 482111) was 196% for the same period. From 1987 to 2016, labor productivity increased 16% in postal service (NAICS 491) but decreased by 43% in couriers and messengers (NAICS 492).

sectors were (respectively) 146.4%, 123.5%, and 128.8%. It is clear that, over the last three decades, labor productivity in manufacturing and trade increased much faster than the rest of the economy.[56] This piece of evidence also supports the technological channel, rather than monopoly power, as the likely explanation for the negative correlation between changes in concentration and changes in labor share. Since labor productivity is measured as real output (net of price changes) per hour, it follows that increasing monopoly power drives up the price but does *not* increase labor productivity.[57]

Appendix Figure A.11 plots the manufacturing sector's labor productivity and labor share from 1987 to 2014. From 1987 to 1997, labor productivity increased 34.7% while the labor share declined 10.3%. Since the late 1990s, both of these trends have accelerated. From 1997 to 2007, labor productivity rose 59.0% and the labor share fell 20.2%.

Labor productivity in more disaggregated manufacturing sectors can be measured us-

[56]One example of technological progress in retail trade is adoption of information technology by online retailers, such as Amazon. From 1987 to 2016, labor productivity in retail trade increased 128.8%; the corresponding increase in nonstore retailers was 860% and in electronic shopping & mail-order houses was 1486%. In 2012, the share of payroll in total revenue for the retail trade sector (NAICS 44–45) was 8.74%, for nonstore retailers (NAICS 454) was 7.03%, and for the electronic shopping & mail-order houses (NAICS 4541) was 6.12%.

[57]Additional evidence against the monopoly power account can be found in the finance sector. In finance and insurance (NAICS 52), the concentration ratio—measured by revenue share of the 50 largest firms—increased from 38.6% in 1997 to 46% in 2007 and to 48.5% in 2012. (A similar pattern is observed for other concentration measures; for example, revenue share of the 20 largest firms increased from 22.6% in 1997 to 28.5% in 2007 and to 31.6% in 2012, and the average revenue share of 4 largest firms across NAICS 6-digit finance sectors increased from 26.0% in 1997 to 36.1% in 2007 and to 35.4% in 2012.) In finance, the sector labor share (measured as the fraction of compensation in value added) actually increased slightly from 23.2% in 1997 to 25.4% in 2007 but declined during the financial crisis. The gist of these observations is that concentration alone cannot fully explain the behavior of labor share. The Compustat data provide information on labor shares (measured as the fraction of compensation in revenue) in large financial firms; see Appendix Figure A.10. Over the period in question, the labor share actually increased in most large financial firms and by a nontrivial amount in many of them. These features of the data reveal the value of combining concentration and changes of labor shares within firms, especially large ones, for understanding the behavior of aggregate labor share. That large financial firms continue to have a larger market share and larger labor share seems not to be a technologically driven phenomenon and so is beyond the scope of this paper.

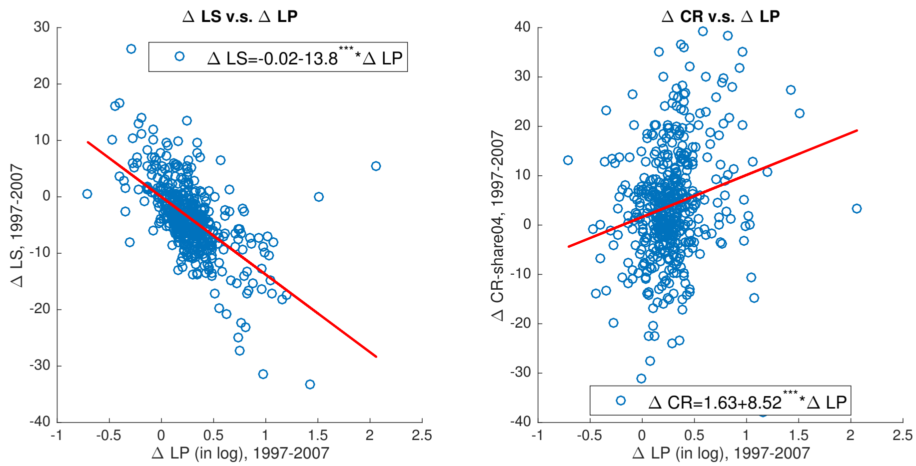


Figure 1.10: Labor Productivity, Labor Share, and Concentration, MFG 1997–2007
2mm *Note:* Labor share is the fraction of payroll in value added; labor productivity LP is the ratio of real value added to total working hours; concentration CR is the value-added share of the 4 largest firms. Each circle represents an NAICS 6-digit sector.

ing Census of Manufactures data. The census data provide—for each 6-digit manufacturing sector—value added, employment, number of production workers, hours of production workers, and a deflator for the value of shipments (1997 = 1). Total hours are constructed by assuming that the average working hours of nonproduction workers are the same as those of production workers. Labor productivity is calculated as the ratio of value added to total working hours. The deflator for value of shipments is used for each sector to obtain the real value added. Figure 1.10 plots the change in labor share and concentration against labor productivity for 6-digit manufacturing sectors. From 1997 to 2007, the following dynamic prevailed: larger increases in sector labor productivity lead to greater declines in the labor share and also to an increase in concentration.

Congress (1995) offers a detailed description of how progress in information technology since the 1980s has transformed the transportation and trade sectors. The use of sophisticated information systems and automation (e.g., bar-coding) has contributed to the rise of carriers such as UPS (United Parcel Service) and Federal Express. Information technology has changed the wholesale trade sector from a system of stocked warehouses to one of fewer but larger-scale distribution centers. Prominent technologies include electronic data interchange (a.k.a. computer-to-computer information interchange), which facilitates the communication of inventory and demand information; bar-coding, which has improved logistics and inventory control while raising the percentage of accurate deliveries; and automation of distribution facilities (e.g., a conveyor system). These new technologies replace certain tasks previously performed by labor, so they increase both productivity and concentration—since facilities must be large enough to support dedicated automated equipment and achieve economies of scale.

At the firm level, one implication of my model is that large firms expand by adopting more advanced technologies that improve their labor productivity. Data from the Economic Census can be used to identify differences between large and small firms. I therefore calculated, for different firm groups, the percentage increase in employment if revenue were to increase by 1%. The benchmark period is from 2007 to 2012. "Large" firms are defined as the top 4 firms (in terms of revenue) in a 6-digit non-manufacturing sector; all other firms in that subsector are then "small " firms. Table 1.9 reports the results. A 1% increase in revenue is associated with an 0.84% increase in employment by small firms as compared with an 0.62% increase by large ones. This difference is both statistically significant and economically large. In addition, the pattern is robust to redefining “large” firms via different cutoffs and to using various digit levels and different years. These findings indicate that growth in small firms relies heavily on more hiring whereas most of the growth in large firms is due to improved labor productivity. Further confirmation is provided by this table's reported R^2 values: the value for small firms (0.69, in column (2)) is much higher than for large firms (0.29, in column (1)).

Table 1.9: Employment Changes associated with 1% Increase in Revenue, 2007–2012

	(1) Large	(2) Small	(3) All	(4)-WLS All
RC_REV	0.62***	0.84***	0.84***	0.79***
	(0.04)	(0.03)	(0.05)	(0.05)
Large			0.02	0.04**
			(0.02)	(0.02)
RC_REV*Large			-0.22***	-0.26***
			(0.06)	(0.06)
R^2	0.29	0.69	0.38	0.35
NO.	489	488	977	977

Note: The dependent variable is the percentage change of employment from 2007 to 2012; the independent variable, "*RC_REV*" the percentage change of revenue for the same period. 'Large' is a dummy and denotes 4 largest firms in a NAICS 6-digit sector. Column (4) uses sector employment in 2007 as weights.

1.4.3 Firm size, concentration and capital intensity

My model implies that large firms adopt advanced technologies which are more capital intensive (higher K/L).[58] Firm-level capital intensity can be measured using data from *Compustat*. The data that I use cover the period from 1980 to 2016; in total, there are 421,501 firm $\times$ year observations. Firm size is measured by assets (and also, for robustness checks, by employment and sales). Capital is defined as the sum of two items: *PPEGT* (property, plant and equipment-total (gross)) and *INTAN* (intangible assets-total). Capital intensity is defined as capital *divided by* number of employees (and is log-transformed).

Table 1.10 gives the results of ordinary least-squares (OLS) regressions in which the dependent variable is (log) capital intensity and the only independent variable is firm size. Dummies for each year and each SIC 4-digit sector are included as controls. I find that

[58] Abow et al. (1999) shows that firms paying higher wages in French are more productive and also more capital intensive.

Table 1.10: Results from Regressing Capital Intensity on Firm Size

	(1)	(2)	(3)	(4)	(5)	(6)
log-assets	0.18***			0.21***		
	(0.001)			(0.001)		
log-sales		0.10***			0.14***	
		(0.001)			(0.001)	
log-emp.			0.03***			0.07***
			(0.001)			(0.002)
Year D.	Yes	Yes	Yes	Yes	Yes	Yes
Sector D.	Yes	Yes	Yes	Yes	Yes	Yes
Sample	1980-	1980-	1980-	2000-	2000-	2000-
R^2	0.62	0.62	0.57	0.58	0.58	0.50
Obs.	225,798	212,927	225,811	107,875	98,103	107,885

Note: Capital intensity is measured as capital *divided by* number of employees (and is log-transformed). Sector dummies are assigned for each SIC 4-digit sector.
Data Source: Compustat, 1980-2016.

capital intensity is positively and significantly correlated with firm size.

The Census of Manufactures also provides values of capital stock for manufacturing subsectors. The capital stock contains equipment and structures both of which are measured in real terms.[59] Capital intensity is defined as the ratio of real capital to number of employees. Figure 1.11 plots capital intensity against the concentration ratio (measured as the value-added share of the 4 largest firms) in 6-digit manufacturing sectors. Sectors that are more capital intensive tend to be more concentrated. Over time, sectors in which concentration increases also become more capital intensive.

[59] The recently capitalized Intellectual Properties and Products (IPP) are not included in the Census of Manufactures data.

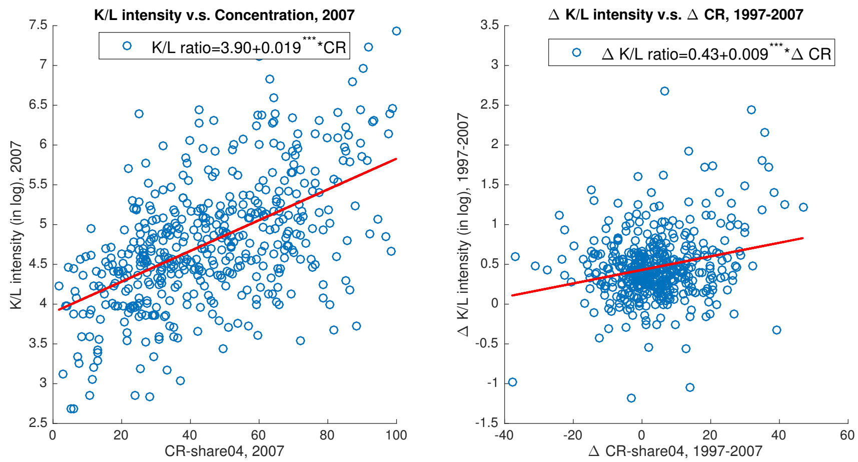


Figure 1.11: Concentration and Capital Intensity in the Manufacturing Sector

Note: Capital intensity, on the vertical axis, is the ratio of capital (equipment + structure) to number of employees; concentration, on the horizontal axis, is the value-added share of the sector's 4 largest firms. Each circle represents an NAICS 6-digit sector.

1.5 Conclusion

This paper documents firm heterogeneity with regard to labor share; in particular, large firms tend to have a lower labor share. . It also shows that a declining aggregate labor share is due to falling labor share in large firms combined with the rising market share of those firms (higher concentration). The sectors of manufacturing, trade, and transportation exhibit the most decline in labor share; these are also the sectors in which concentration has increased the most and the relative labor share of large firms has declined the most. The increases in labor productivity of large firms in these sectors far exceed the increases in wage.

I provide a rationale for these empirical facts by assuming that capital and labor are complementary inputs and that technological progress is labor saving. Under these assumptions, my model predicts a negative correlation between firm size and labor share. Given the complementarity of capital and labor, (labor-saving) technology increases the productivity of capital and the demand for effective labor, thereby increasing output; however, technology substitutes raw labor and so reduces the latter's share of income. Furthermore, the adoption of new technologies diminishes the labor shares in large firms and increases their market share. Hence the aggregate labor share declines. This technological channel is consistent with the evolution of labor productivity across sectors during the last 30 years. From 1987 to 2016, economy wide labor productivity increased by 72.7%. For the same period, the labor productivity increases in manufacturing, wholesale trade, and retail trade were (respectively) 146.4%, 123.5%, and 128.8%.

My paper focuses on post-1980 period. Possibly at a slower pace, there is of course labor saving technological change before 1980s. Two facts suggest the mechanism proposed in the paper also work before 1980. First, as reported in Table A.3, the negative correlation between change in concentration and change in labor share holds in 1960s and 1970s; even the coefficient has a smaller absolute value, all coefficients are negative and significant. Second, the labor share calculated using post-2013 revision NIPA data shows a slower and declining trend before 1980 (See Koh et al. (2016)). The stationary concentration ratio in manufacturing in late 1960 and 1970s might be due to counteracting forces (e.g. union power) and I leave that for future investigation.

The model in paper does not address sector heterogeneity. A promising extension would be to embed that heterogeneity into the model for quantitative analysis. Also, in some sectors, new technologies are adopted by small firms that grow large over time (e.g. Amazon in retail trade). Another extension would be to capture a richer firm dynamics by incorporating the creation and death of incumbents firms into the model.

Chapter 2

Social-Economic Change and its Impact on Violence: Homicide History of Qing China

Zhiwu Chen Kaixiang Peng Lijun Zhu[1]

2.1 Introduction

Historians often divide China's Qing Dynasty (1644-1911) into an early period of prosperity and a late period of stagnation or decline (Rowe, 2011). Starting from the Kangxi reign in 1661 and ending after the Qianlong reign (about 1813), the prosperous period was distinguished by high income growth and social tranquility (hence this period is often referred to

[1]Chen: School of Management, Yale University, 165 Whitney Avenue, New Haven, CT 06511 (e-mail:*zhiwu.chen@yale.edu*); Peng: Department of Economics, Henan University, Kaifeng, Henan, China (e-mail: *kaixiangp@qq.com*); Zhu: Department of Economics, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130 (email: *lijun.zhu@wustl.edu*). This chapter of dissertation has been published as Chen et al. (2017), Social-economic change and its impact on violence: Homicide history of Qing China. Explorations in Economic History, 63, 8-25. https://doi.org/10.1016/j.eeh.2016.12.001.

as the 'Kangxi-Qianlong Prosperity'); in contrast, the rest of the 19th century was marked by economic stagnation, or decline, and war[2]. This conventional view about the Qing is supported by standard economic measures such as GDP (Maddison, 2007), living standards Allen et al. (2011), and population gain. According to Maddison, China's population grew from 138 million in 1700 to 381 million in 1820 and to 437 million by 1913–a cumulative growth of 176% from 1700 to 1820 but only 15% for the century that followed. Thus, economic growth was far greater in the 18th than in the 19th century. China also experienced major humiliating wars in the latter period: the First Opium War (1839-1842), the Second Opium War (1856-1860), the Sino-Japanese Naval War (1894-1895), and the war against the Eight-Nation Alliance (1900).

Yet standard economic measures are only partial indicators of a society's development progress and cannot reflect the full picture. In this paper, we construct China's homicide rate history and investigate the socio-economic drivers of changes in violence during the period from 1661 to 1898. Our goals are to shed new light on Qing China's economic history and to improve our understanding of the interacting dynamics between economic growth and social change. In particular, we focus on ordinary interpersonal violence by excluding war and other organized intergroup violent acts. Because the lack of suitable data makes it difficult to estimate general 'ordinary' violence, we rely instead on the homicide rate as a proxy. In so doing, we assume that ordinary violence and the homicide rate are highly correlated. Although the homicide rate is not a performance measure in the vein of income growth, population change, or other economic measures, it does capture an important dimension of ordinary people's well-being and living standards. A society in which the homicide rate rises is one characterized by reduced well-being of its members and increasingly insecure property

[2]There is some disagreement over when the Kangxi-Qianlong Prosperity started and when it ended. Gao (1993) argues that it started in 1681 and ended in 1795, whereas Li (1999) gives 1684 and 1813 as the respective starting and ending years.

rights. According to North et al. (2009), the use of violence is restrained by political and economic institutions that give individuals control over resources, which in turn shapes the incentives faced by those in a position to commit violent acts. It follows that the level of interpersonal violence is a good indicator of progress in institutional and economic development. The 'civilizing process' theory of Elias (2000) holds that, at the individual level, humans have developed a higher level of self-control by way of literacy, education, and cultural consumption (e.g., reading and group learning). Because the homicide rate is driven by social, economic, and institutional factors, it is also an intertemporally and internationally consistent measure of interpersonal violence and associated insecurity (Baten et al. (????)). For these reasons, we seek to assess the different periods of Qing China by using the homicide rate's level and trajectory.

Using sources kept at the First National Historical Archives of China, we offer the first estimate of interpersonal homicide rates for the period 1661-1898. Our main finding is that the national homicide rate ranged between 0.35 and 1.47 homicides per 100,000 population annually, which was much lower than in Western Europe at the time[3]. More specifically, China's homicide rate rose steadily from about 0.6 (per 100,000 population) in 1661 to about 1.47 in 1821–an increase of 145% over the 140-year period! Thus, underlying this increasing homicide rate was a significant increase in ordinary interpersonal violence during the Kangxi-Qianlong Prosperity. The opposite occurred from 1821 onward, when the national homicide rate was indecline.[4]

[3]According to Eisner (2003), Western European communities during the 17th-19th centuries had a homicide rate that ranged between 0.6 and12. The European rates did not approach the low levels in China until the 19th century. Why did China have much less violence among ordinary people than did its Western counterparts? One could follow North et al. (2009) and develop a complete explanation of the contrast in homicide rates between China and the West, but that undertaking is beyond the scope of this paper.

[4]The death rate due to war was probably falling (and lower) during the Kangxi-Qianlong period than during the post-1820 Qing period, which saw such deadly conflicts as the Taiping Rebellion. We follow Eisner (2003) and Elias (2000) in focusing on ordinary interpersonal violence.

In order to explain the intertemporal variation in China's homicide rate, we propose–and use cross-provincial panel data to test–several hypotheses. Our population pressure hypothesis states that significant population growth and large-scale migration put considerable stress on society and cause more conflicts to occur until new norms are firmly established. This hypothesis is consistent with Buoye (2000) finding that, when large numbers of migrants enter a region, the effect of ambiguous property rights on disputes may be exacerbated until new norms emerge. Our survival distress hypothesis assumes that, when grain prices rise (because of crop failures or other risk events), the ability of ordinary people to survive is challenged and forces some individuals to seek violence. The link between crop failure and violence is well established in the literature (e.g., Anderson et al. (2013); Bai and sing Kung (2011); Jia, 2013).

According to Elias (2000) and Eisner (2003), state formation represents both a civilizing and a pacifying process because social order is likely to improve once the state monopolizes the legal use of violence, imposes rules, and enforces them. Miller (2013) and Wakeman (1998) document that state power was on the rise–and civilian self-governance was in decline–during the Kangxi-Qianlong era, although these trends reversed starting early in the 19th century. We hypothesize that the level of state power must have effects on violence, though the net impact may be difficult to determine. There are at least three channels through which state power affects the level of violence.

First, state power might make government agencies more efficient and improve the overall society's law and order, leading to lower violence rates. We refer to this as the state capacity channel, which in our empirical implementation is captured by a region's 'Chong' rating (applied to key administrative zones) by the Qing government; when a region was rated

Chong, the government would likely send a more capable official to govern that region and in that way increase state capacity there (or at least signal such an increase). Second, the rise of state power might weaken local self-governance institutions and thus reduce the role of the gentry (Miller (2013); Wakeman (1998)), leading to greater social disorder and more violence at the local level; we refer to this as the gentry channel. Third, newly gained state power might be directed at setting up regional border barriers to prevent grains and other goods from flowing between provinces or other administrative zones. Thus, for example, grain markets actually became less integrated across regions from the early 18th century to the early 19th century; as a consequence, ordinary people became less able to cope with crop failure (and other income shocks), which in turn led to more violence. We refer to this as the market integration channel.

Our empirical exercise uses Chinese cross-provincial homicide data to show that, during the 18th and 19th centuries, provinces with higher population density and higher grain prices (reflecting both population pressure and food supply conditions) experienced higher homicide rates–especially if these conditions were accompanied by less integration of grain markets, lower state capacity (as proxied by a sub-Chong rating at the provincial level), and fewer gentry in the province. These findings are largely consistent with our population pressure and survival distress hypotheses. At the same time, the channels of state capacity, local governance, and cross-regional market integration could all serve to reduce ordinary violence.

The cross-provincial regression results allow us to offer a partial explanation for the pre-1821 upward trend and the post-1821 downward trend of China's national homicide rate. We demonstrate that, prior to 1821, China experienced fast population growth, rising grain prices, and increasing disintegration of the grain market; all these factors contributed to the observed continuous rise in the national homicide rate. After 1821, however, the opposite scenario obtained: population growth slowed down, grain prices stabilized or declined, grain markets became more integrated across regions, state power weakened, and local self-governance strengthened. As a result, the national homicide rate declined for most of the 19th century's remaining decades. Thus Chinese societal pressures due to rapid economic growth and institutional changes led to a rise in ordinary violence and property insecurity during the era of Kangxi-Qianlong Prosperity; but as population and economic growth pressures lessened and state power retreated in the 19th century, so did interpersonal violence and property insecurity.

Our paper introduces historical China data-based insights to the literature on the economics of criminal behavior (Becker (1968); Grossman (1991)). In particular, we examine the effect of grain price shocks on violence in the context of the 18th-19th century China while exploring how various factors influenced this effect. Higher grain prices could be a response to crop failure due to natural disasters, which reduced the returns to land-based labor and increased the gains from violent behavior. In addition, higher grain prices increased land values and so created more incentives for land-related disputes (Buoye (2000)). Given these effects, homicide rates tended to be higher when grain prices increased; yet the strength of these effects depended on the institutional environment. As our empirical work shows, a higher level of state capacity or more local gentry governance reduced these negative effects by increasing the cost of violence while reducing its benefits and also by increased sharing of risks through state and civilian granary networks. Trading networks that were better geographically integrated would provide a market alternative and improve the capacity of ordinary people to cope with food shortages, thus reducing the need for violence and (by extension) acts of homicide.

Besides expanding our knowledge of Qing China's socio-economic history, our work contributes to the literature on the history of violence. Much research has addressed the history of interpersonal violence–in particular, the homicide rate–in Western Europe since the late Middle Ages. For example, Elias (2000) and Pinker (2011) document a remarkable long-term decline in interpersonal violence due to the 'civilizing process' and to institutional, cultural, and market development. Gurr (1981) collects estimates of homicide cases for 30 English localities and finds that their annual homicide rates fell from about 20 per 100,000 inhabitants in the high and late Middle Ages to 10 by 1600 and to a mere 0.1 by the end of the 20th century. However, we are not aware of any published efforts at estimating China's rates of homicide and other violence during different historical periods. It has therefore been difficult either to evaluate China's process of civilization quantitatively or to compare China with other countries. Our paper fills thisgap.

2.2 Data sources and description

In Qing China, local governors were required to report all homicide and other death-penalty cases to the central government, using a standardized template known as the Tiben, or 'case memorials' or simply 'memorials' in English. For each important case (and certainly for homicides), the local governor would submit a memorial (i.e., a General Report or *Tongben*) to the Grand Secretariat, where it was copied and transmitted to the Ministry of Justice; the latter would then return a memorial (i.e., the Ministry Report or *Buben*), along with its opinion on the case, to the Grand Secretariat. Thus for each case there were two memorial reports, the General Report and the Ministry Report. These reports were originally kept by the Red Book Archives (Hongben Ku) of the Grand Secretariat, and most of them ended up in the First National Historical Archives.

By a 1745 order of Emperor Qianlong, the Ministry of Justice began collecting statistics on key cases (including homicides) based on memorials submitted from the previous year into booklets that we refer to as *Case Summary Books*[5]. The extant Case Summary Books are available for 39 years from 1744 to 1898. No summary statistics on homicide cases (and other criminal acts) are available before 1743. We therefore rely on the estimates of total Red Books (1661 being the first year with data available) based on information recorded by Fang (1934) for 1661-1743. Due to serveal institutional changes after 1860s, we used the extant Tiben case memorials, which are publicly available at the First National Historical Archives to estimate homicide rate from 1860 to 1898. Accessible are memorials on two types of homicide: (i)land- and debt-related homicide, and (ii)marriage- and adultery-related homicide. We refer to these two types of memorilas as *Land & marriage memorials*. They represent about half of all surviving homicide case memorials.

Table 2.1 presents summary statistics for these three data sources, homicide cases from the *Case Summary Books*, annual estimates of Red Books, and number of Land marriage memorials, as well as the annual number of death penalty executions as reported by the *Qing Chronicles*. The time series pattens are plotted in Figure 2.1.

[5]These books are also known 'Yellow Books' (*Huang Ce*). Because this term was used with reference to many of the imperial court's documents, we use the terminology Case Summary Books to avoid confusion

Table 2.1: Summary statistics for annual data, 1661-1898

Variables	First year	Last year	Obs.	Mean	S.D.	Min.	Max.
Homicide cases	1744	1896	39	2,422	842	972	4,459
Red Books	1661	1888	96	6,122	2,654	1,170	10,578
Land & marriage memorials	1736	1898	163	1,477	517	178	2,573
Death penalty	1738	1849	82	813	271	296	1,662

Note: 'Homicide cases' is the total number of homicide cases concluded by local governors and summarized in the *Case Summary Books*. 'Red Books' is the annual total of Red Books (in the Red Books Archives). 'Land & marriage memorials' is the annual total of extant case reports concerning land/debt or marriage/adultery homicide. 'Death penalty' is the annual total of death penalties carried out–through *Qiushen* and *Chaoshen* deliberations (see the Appendix) but excluding executions that proceeded without due process.

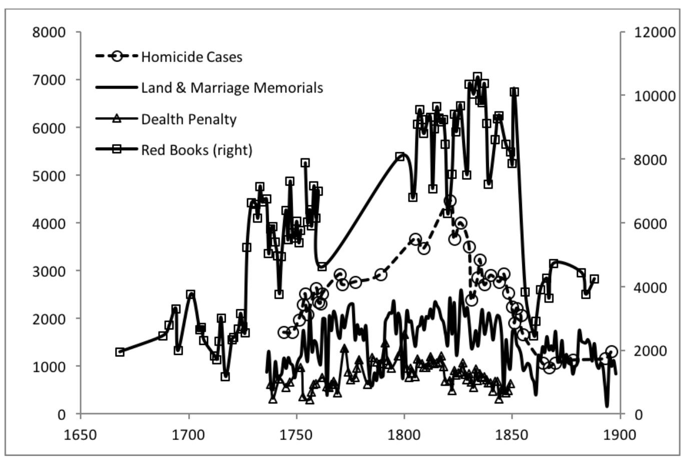


Figure 2.1: Distribution and trend of homicide-related series, 1661-1898

Note: See note to Table 2.1 for descriptions of each variable.

Figure 2.1 shows that the national homicide total, the Red Book total, the number of land/marriage memorials (combining both land/debt and marriage/adultery cases), and the national death-penalty total all track each other closely from 1744 to 1860; this is the period for which data availability and quality are good for each of the four series. In particular, all these reported crimes increased during the 18th century and peaked after the start of the 19th century–suggesting a general rise in violence followed by a decline. Collecting data from different sources allows us to cross-check the accuracy of these data, which is crucial for the discussion to follow. We will rely on this cross-validation when employing either the Red Books series or the land/marriage memorials series to approximate missing national homicide totals for earlier and later periods. More institutional background and comparisons are provided in the Appendix.

2.3 National homicide rate trend

Our homicide statistics for 1661-1898 are estimated separately for three subperiods: 1661-1743, 1744-1860, and 1861-1898. As explained in Section II, annual homicide case counts for the 1744-1860 period are taken directly from the *Case Summary Books.*

For the 1661-1743 subperiod (i.e., prior to the existence of *Case Summary Books*), we use the annual Red Book estimates[6] to arrive at our approximation. In particular, we divide the annual Red Books (conditional on the number of Red Books being available for the year in question) by the average Red-Books-to-Homicide ratio rom 1744 to 1850, when Red Book counts and the national homicide totals (from *Case Summary Books*) are *both* available for 16 year to obtain an estimate for that year's homicide case total . Altogether, we have homicide estimates for 34 years for the 1661-1743 subperiod. To limit the effect of fluctuations in the Red Books series,[7] we use the average annual homicide total over consecutive five-year periods.[8] We thus end up with 11 yearly estimates for the 1661-1743 period.

[6]See appendix for details on our estimates of annual red books. Note that many non-memorial administrative files were also recorded in the Red Books Archives. So we can't use that directly as homicide estimates.

[7]We highlight the long-term trend by averaging out short-term fluctuations.

[8]For example, the average value for 1661-1665 is used as the national homicide total for 1663.

Several changes were made after 1860 in response to the Taiping Rebellion and other conflicts. First, the emperors began issuing orders to pardon certain categories of alleged criminals, such as some homicide cases awaiting the Ministry of Justice's review; these cases were excluded from the Case Summary Books homicide statistics even though fatal violence had occurred.[9] Hence the associated Case Summary Books statistics under-represent interpersonal violence for such years. Nonetheless, local governors should still have filed Tiben memorials with the Grand Secretariat. Figure 2.1 shows no dramatic change around 1860 in the number of land/marriage case memorials, which suggests that changes in the pardon policy did not undermine the Tiben reporting practice despite affecting the Case Summary Books counts. It follows that, after 1860, the number of Tiben case memorials should more accurately reflect the extent of homicide occurrence–and thus also of ordinary violence. A second post-1860 change was that, thereafter, local governors were required to follow the Tiben format only when reporting homicide cases to the Grand Secretariat. Subsequently, the Tiben template was no longer required for other categories of legal and bureaucratic matters or for less important criminal cases. This change in practice altered the meaning of aggregated Red Book counts after that year.

Given these two significant changes, we did not rely either on the Red Books or *Case Summary Books* when estimating homicide statistics for the post-1860 Qing Dynasty. Instead, we approximated homicide rates by using the total number of land/marriage case memorials for each year, employing the same approach as the pre-1743 period[10]. There are 38 annual observations estimated from 1861 to 1898, and averaging this series for each five-

[9]The *Case Summary Books* often stated outright that the reported statistics exclude cases in which the defendant was pardoned or had his sentence reduced.

[10]i.e. multiplying land/marriage memorials for 1860-1898 by the average ratio of Homicide totals to land/marriage memorials between 1744 and 1850.

year interval generates eight homicide estimates for the post-1860 period.

In total, we have 53 data points concerning national homicide totals for the entire 1661-1898 period: 11 estimates for 1661-1743, 33 for 1744-1860, and 8 for 1861-1898. The homicide series is graphed in Figure 2.2, where the plot clearly exhibits an inverted U-shape. From 1661 to 1821, the number of homicides rose from fewer than 1,000 annually to more than 4,000, a threefold increase; however, homicides gradually declined in number after 1821 and for the rest of the 19th century.

Figure 2.2: Total number of homicides in China, 1661-1898

Note: Each solid squares represents the homicide total for that year as estimated from data in the *Case Summary Books*. The open diamonds represent five-year averages of homicide cases estimated from the Red Books, and the open triangles represent five-year averages of homicides estimated from land/marriage case memorials.

China's population fluctuated considerably during the Qing Dynasty. Cao (2001) provides detailed population estimates for six selected years during the Qing Dynasty: 1678, 1776, 1820, 1851, 1880, and 1910. We obtain the country's population for other years via interpolation (while assuming a constant growth rate between any two data points). We then divided the homicide counts reported in Figure 2.3 by the estimated population for the corresponding year, thereby obtaining the (unadjusted) homicide rate series. Yet we must bear in mind that infanticide, as well as killing the killing of a wife or concubine by her husband, were not treated as homicide in Qing China, so it is not included either in the Case Summary Books or in the extant homicide case archives. We inflated the unadjusted homicide rate series for China (as just derived) by 25% (see the Appendix for details).

The adjusted values are presented in Figure 2.3. This homicide rate–like the homicide totals plotted in Figure 2.2–exhibits an inverted U-shape. From 1661 to 1821, China's homicide rate increased from about 0.6 during the 1661–1665 period to 1.47 in 1821 (a two-fold increase). Our estimates indicate that, in China, the homicide rate at the end of the 19th century was comparable to that of the 1660s.

Figure 2.3: Annual homicide rate (per 100,000 population) in Qing China

Note: A solid circle represents statistics from the Case Summary Books, and the open circles signify approximations.

Estimates by Eisner (2003) indicate that the average European homicide rate started to decline around 1500 and continued to do so until the 20th century.[11] Western European

[11]One difference between our homicide series and those reported in Eisner's work is that our estimation is at the national level and is based on statistics collected by the Qing Ministry of Justice, whereas Eisner

cities had average homicide rates of about 6 per 100,000 population during the late 17th century, 3 to 4 in the late 18th century, and 2 to 3 before 1900. During the same periods, China's corresponding homicide rates were much lower: respectively 0.6, 1.5, and 0.6. European homicide rates did not approach China's low level until late in the 19th century.

There are several factors that could account for this difference between China and the West. First, state formation in China began in the Qin Dynasty from 221 bc onward–long before state formation in Western Europe. According to Elias (2000) and Eisner (2003), state formation is both a civilizing and a pacifying process because the state monopolizes the legal use of violence while imposing and enforcing law and order. Given China's much longer history of centralized governance, these civilizing and pacifying processes likely explain the lower ordinary violence there than at the same time in Europe. Second, Confucianism might have contributed to lower levels of violence (Miller (2013)). Confucianism emphasizes community governance by local elites or gentry as well as on ancestor worshipping within each clan, so there was no ambiguity about whom was vested with authority. Clear authority often goes hand in hand with order, which may be why Confucian societies in general (even today) have less violence.[12]

The late introduction of guns to Qing China may also have played a role in China's low homicide rate. Guns were invented and widely available throughout Western Europe in the 16th century.[13] Yet even though Ming Dynasty soldiers encountered Western matchlock guns

(2003) pre-modern estimates are all local (mainly at the county level).

[12]According to Baten et al. (????), the homicide rate for 2000-2010 is 0.7 (per 100,000 population) in France, 1.4 in U. K.; 2.5 in Italy, and 6.1 in U. S.; it is 1.6 in China and 0.5 in Japan. According to the 'intentional homicide' data compiled by the World Bank (http://data.worldbank.org/indicator/VC.IHR.PSRC.P5), homicide rates for South Korea and Hong Kong are (respectively) 1.0 and0.0. Thus homicide rates in today's societies that were influenced by Confucianism (China, Hong Kong, Japan, and Korea) are generally lower than elsewhere

[13]The first recorded use of a firearm, in 1364, was in Europe. Handguns were present across Europe by 1380, and the matchlock gun was invented in the 1400s; rifles were popular in Europe by the mid-16th

in 1521 (when fighting the Portuguese in Canton), handguns were not widely available in China until the late 19th century.[14] Figure 2.4, which is based on the 49,627 Case Summary Books cases we examined, plots the fraction of cases (by category) in which weapons were used. Prior to 1850, only 0.29% of criminal cases involved the use of a weapon (e.g., knife or gun). After 1850, however, the use of weapons rose significantly; this increase resulted mainly from the introduction of rifles to China in the late 19th century. Even so, China's homicide rate did not rise significantly during that time.

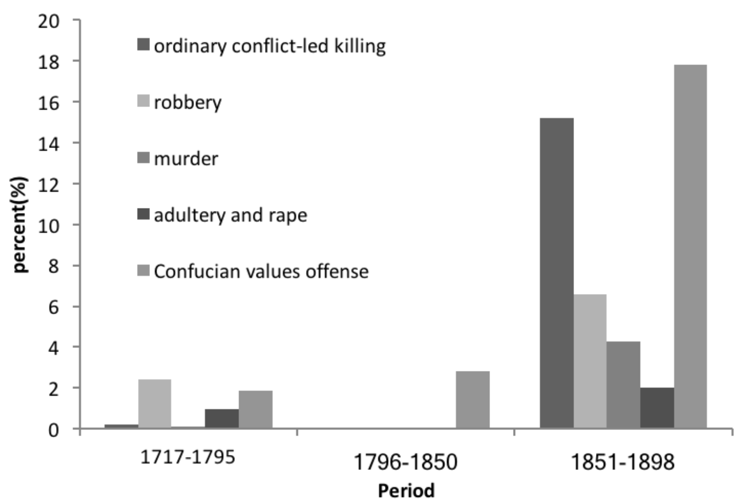


Figure 2.4: Fraction of homicide cases involving use of a weapon

Note: A solid circle represents statistics from the Case Summary Books, and the open circles signify approximations.

2.3.1 Verification and cross-checking

It could be that the decline of homicide during the 19th century is due to under-reporting and deterioration of local government efficiency (as opposed to real changes in homicide rate). The Qing state increased its power from the reign of Emperor Kangxi to that of Qianlong

century (http://www.pbs.org/opb/historydetectives/technique/gun-timeline/).

[14]Interpretative caution is advised, however, since these country estimates are from different types of data sources: those for China (resp., Western Europe) are based on national (resp., local) data. Because the reporting of homicides and the administrative dealing with cases is fairly homogenous within a given country, interregional comparisons within a country are more robust than those across countries.

(i.e., from 1661 to 1795), so there are reasonable grounds for believing that local officials during the so-called Kangxi-Qianlong Prosperity would not dare to under-report. Yet after Emperor Jiaqing began his reign in 1796, the state's grip on power gradually loosened and compliance with reporting requirements became less reliable–a trend that continued well into the 19th century (Miller (2013); Sng (2014); Wakeman (1998)). Thus many homicide cases may not have been reported to the central government in the 19th century as local governors shirked their responsibilities, creating a false impression of reduced violence during that time. That being said, the Qing Code explicitly threatened local officials with punishment for non-performance. Local governors failing to report homicide cases would be dismissed, and those who knowingly misreported homicide cases would be demoted and perhaps charged with acrime.

For the 49,627 cases drawn from the Case Summary Books, we calculated the time lag between the date violence occurred and the date of Tiben reporting to the Grand Secretariat; our aim was to see whether there had been structural changes in reporting practices over time. The average lag, displayed in Figure 2.5, was 15.19 months for the 1717-1795 period and 15.43 months from 1795 to 1850. There is no clear evidence of any decline in administrative efficiency or state capacity (as reflected in reporting lag) from the 18th to the 19th century. However, the average time lag did increase to 21.32 months after 1850. There are a number of reasons for this increase, which include the Taiping Rebellion's impact as well as changes to (and the eventual abolishment of) the Tiben memorial system–since those changes led local officials to be less compliant with reporting standards.

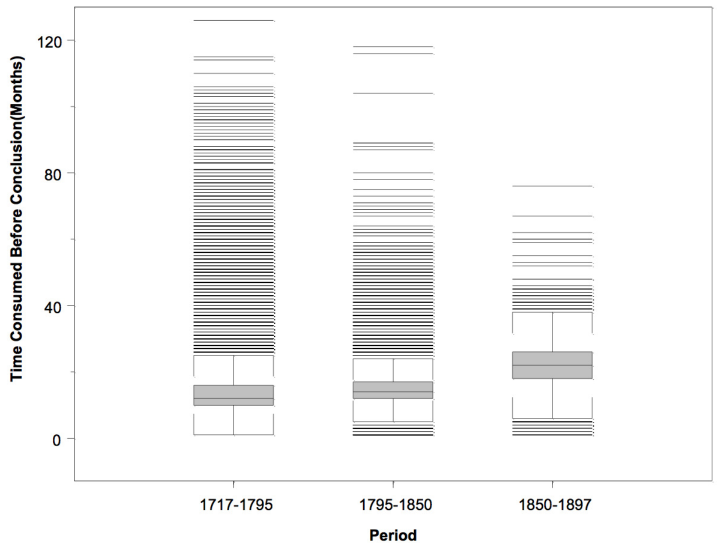


Figure 2.5: Occurrence-reporting time lag of *Case Summary Books* cases

Note: The occurrence-reporting time lag is defined as the number of months from the day a crime occurred to the day of its being reported, via *Tiben*, to the Grand Secretariat. The dark horizontal line inside each shaded box (which represents half of the respective subperiods' observations) indicates the average timelag.

We can cross-check to see whether our collected homicide statistics are consistent with Qing China's social conditions as perceived by officials at the time. The Qing government created a system for rating each prefecture along four dimensions[15]. In particular, a prefecture was labeled Fan (corresponding to a value of 1 for our Fan indicator variable) if its administrative burden was heavy and cumbersome. In Figure 2.6 we present both the average annual homicide total from 1744 to 1860 and the average Fan measure across prefectures for each of 17 provinces[16]. The chart demonstrates a significant and positive correlation between the two. Note that the Chong, Fan, Pi, and Nan ratings were made and recorded independently of the *Tiben* case reporting system. This simple correlation exercise suggests

[15]These dimensions are: Chong (if the prefecture was geographically and/or strategically important); Fan (if it was administratively burdensome); Pi (if tax compliance/collection was difficult); and Nan (if social order and local institutions presented challenges).

[16]Sichuan is not included in this figure. The average annual number of homicide cases for Sichuan was 355 yet the average annual number for the other 18 provinces was 125; hence Sichuan was clearly an outlier. It had both a large number of homicide cases and a high prevalence of Fan. We plot provincial-level data owing to the sparsity of county- and prefectural-level cases and to the nonavailability of county population data.

that our homicide data series is consistent with the Qing government's original ratings of each province's governability.

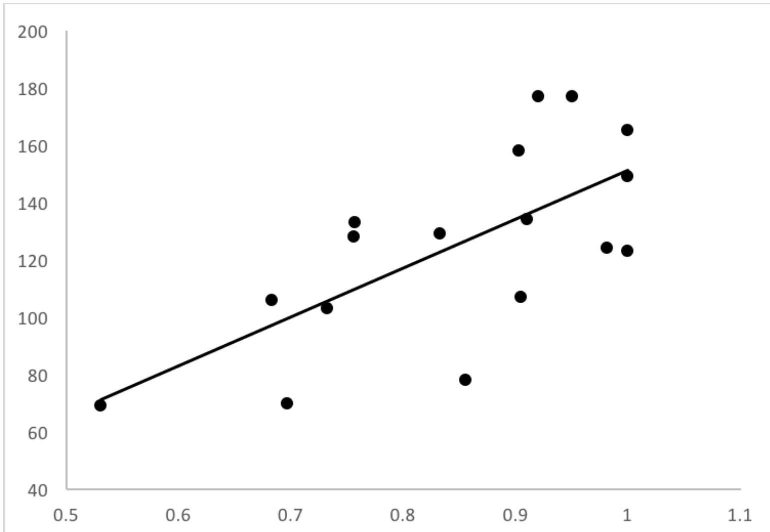


Figure 2.6: Average annual homicide cases versus *Fan* (horizontal axis) for 17 provinces

Note: The vertical axis indicates the average annual homicide case total for each province from 1744 to 1860; the horizontal axis indicates the population-weighted average of the Fan dummy across all prefectures in the given province.

2.4 Social and economic drivers of the homicide rate

Violence occurs in a social, economic, and institutional context. Before attempting to explain the homicide rate trends during Qing China, we use cross-provincial data to identify the social and economic drivers of differences across the provinces: population density, grain prices, market integration, state capacity, and local gentry governance.

2.4.1 Homicide rates across Qing provinces

Table A.1 in the Appendix presents statistics on homicide cases and homicide rates for 18 provinces in China Proper from 1774 to 1849, all based on the *Case Summary Books*.[17] There

[17] Since many *Case Summary Books* were damaged or lost over the past three centuries, our search of available sources at the First National Historical Archives yielded the required details for only some years.

are about 30 annual observations for each province. To avoid the bias problems discussed previously, we exclude data after1850.

Sichuan had the highest average number of homicide cases per year: 357, or almost 3 times the average for other provinces. Its homicide rate was also the highest (1.67), followed by Guizhou; of all the provinces, these two absorbed the most migrants before and during the 18th century. Peripheral provinces–such as Guangxi, Gansu, and Yunnan–experienced fewer homicides (owing to their relatively sparse populations), but their homicide rates were in the middle of the distribution. The most developed provinces (in the Yangtze River delta) had the lowest homicide rates. For example, Jiangsu's rate was 0.39, or about one quarter of that for Sichuan; Anhui and Zhejiang enjoyed similarly low violence rates.

We can also examine the time trends of provincial homicide rates and make comparisons across provinces. Figure 2.7 presents the homicide rate history of four provinces: Guangdong, Sichuan, Jiangsu, and Shandong. For 1860-1895, we estimate homicides based on land/marriage memorials data in the same way as for the post-1860 national homicide rates. The *Tiben* memorial counts are averaged for each 10-year span from 1856 to 1895, resulting in homicide rate averages for the decades of 1860, 1870, 1880, and1890.[18]

For other years, there are no (or only partial) data available for homicide and non-homicide cases. For the years 1755, 1761, 1823, 1835, and 1848 (and, in some provinces, also for 1748 and 1777), the total numbers of *Case Summary Books* cases–that is, including both homicide and non-homicide cases–are available; however, we are unable to distinguish between these case types because the data are not sufficiently detailed. For those years, we estimate the homicide total for a province in two steps: (i)calculate the average ratio of the *Case Summary Books* case total to actual homicide total for that province in the years for which both types of case counts are available; (ii)divide the Case Summary Books case total by this average ratio to obtain that year's homicide estimate for the province. The highest standard deviation for this ratio is 0.14 for Zhili (with a mean of 1.2), and the lowest is 0.03 for Jiangxi (with a mean of1.1).

[18]Provincial homicide rates are not extended to the pre-1744 years because Red Books statistics cannot be disaggregated into provinces.

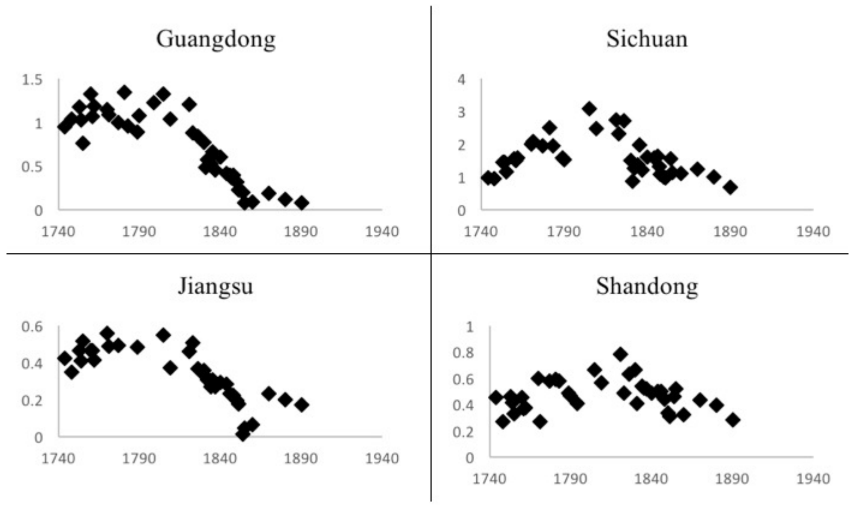


Figure 2.7: Homicide rates in four provinces, 1744-1895

Note: Each diamond represents an estimated annual homicide rate value. The pre-1860 rates are based on the *Case Summary Books*; post-1860 rates are based on land/marriage memorials.

All four provinces experienced rising homicide rates from 1744 until about 1821 and a decline thereafter, although the slopes of the rise and fall differed among them. Sichuan stood out in this regard also, as its rise and fall were the sharpest among these provinces. The implication is that the dramatic demographic changes due to migration indeed shook up Sichuan and led to more violence there (Buoye (2000)). By the 19th century, the early migrants to Sichuan and other regions had settled in to their new homes, after which the pressures inciting violence probably decreased relative to the early Qing decades. Thus the costs and consequences of violence increased as its impetus declined, and these developments led to a reduction in homicide from the 1820s onward. Population change likely has a strong effect on homicides, a hypothesis that we test next. Yet because the other provinces also exhibited a 'rise then fall' pattern in homicide trends, it is likely that other drivers of violence were also atplay.

2.4.2 Explanatory variables of provincial homicide rates

To explain differences across provinces, we need to identify the likely drivers of interpersonal violence. Our *population pressure hypothesis* (and first explanatory variable) concern population change. As an agrarian society, China had limited ability to cope with rapid population growth. A rapidly increasing population density could result in degraded living conditions and hence in Malthusian stress for ordinary people, causing conflict to arise. There was substantial variation in population growth across the provinces. During the 17th-18th centuries, the largest migration wave was the movement of 'filling Sichuan with people from Hunan and Hubei' (*Huguang tian Sichuan*). Sichuan's population was destroyed in the civil war of 1630-1640 and so, when the Qing Dynasty was founded, the government encouraged millions of peasants from nearby provinces (e.g., Hubei, Hunan, Jiangxi) to migrate there.[19] From the mid-18th century onward, hilly and mountainous areas were the main destinations for migrant peasants as the population pressure in the Sichuan plains intensified (Buoye (2000)).

In agrarian societies, interpersonal trust and behavioral norms are generally established through repeated exchanges and interactions; that dynamic makes cost-benefit calculations work, howsoever gradually, against violence. For this reason, large-scale migration will not only disrupt established bonds and norms in the communities left behind; it will also create frictions both among newcomers from different regions and between migrants and established residents. In a new environment, there are fewer costs to a perpetrator of violence against strangers. Over time, however, new norms will emerge that curb incentives for violence. Using land homicide *Tiben* memorials, Buoye (2000) shows that–as the primary destination of migrants during the 17th-18th centuries–Sichuan indeed experienced many homicide cases.

[19]Sichuan's population was decimated by warfare prior to the Ming Dynasty's collapse in1644. In1776, migrants and their descendants accounted for some 60% of Sichuan's 10 million inhabitants (Buoye (2000)). By 1851, Sichuan's population was nearly 30million.

Cao (2001) provides provincial population estimates for five years during Qing China: 1776, 1820, 1851, 1880, and 1910.[20] Since our annual cross-sectional regressions focus on the 1744-1849 period, there are only two independent population-growth data points for each province (1776-1820 and 1820-1851). So for the years during, say, 1776-1820, the annual population growth rate does not vary; hence the annual provincial population growth is too similar (for estimation purposes) to the province fixed effect. We therefore use each province's annual population *density* as a proxy for population pressure. Province populations for each year are obtained by interpolating between Cao's estimates for the two years closest to the focal year, and population density–denoted *PopDense* (in log value)–is calculated as population per square kilometer ofland.

Our *survival distress hypothesis* concerns grain prices. We use each province's price of grains (averaged across its prefectures), which are probably the most important goods in agrarian societies, to approximate the overall distress level of that province's inhabitants. The effect of grain prices on overall homicide operates through two main channels. First, a short-term increase in grain prices may reflect crop failure due to drought, flood, or other natural disasters; in that case, high grain prices proxy for food-supply stress that can induce otherwise law-abiding citizens to steal, rob, or commit even worse crimes. High grain prices can also result from wars that cut off normal supply chains of grains and other materials. The connection between natural disasters and violence is well documented in the literature (e.g., Anderson et al. (2013); Bai and sing Kung (2011); Jia (2013); Edward Miguel and Sergenti (2004)). The second channel is that higher grain prices make land worth more and thereby lead to more land disputes and related violence–especially when land property rights are ambiguous (Buoye (2000)). In addition, grain price changes may also be a response to

[20] Maddison (2007) provides *national* population estimates for only three non-Qing years, whereas Cao (2001) reports estimates of *provincial* populations.

population pressure, thus capturing a different type of survival distress.

During the Qing Dynasty, grain prices were reported by local officials on a monthly basis; these price reports are now kept in the Grain Price Database for Qing Dynasty at Academia Sinica's Institute of Modern History in Taiwan. We use the average grain price across a province's prefectures as its provincial grain price, denoted by *GPrice* (in log value).

We advance three hypotheses related to the three main channels though which state power affects violence. First is the 'state capacity channel' whereby, according to Elias (2000) and Eisner (2003), the law and order supported by state power both reduces the benefits of violence and increases its costs. For example, Buoye (2000) demonstrates that whether the government created and maintained an unambiguous property rights system made a significant difference in the occurrence of homicide. Buoye uses this public good to explain the homicide rate's inverted U-shaped pattern in 18th-century Guangdong.[21] According to Buoye, Sichuan's increase in property rights related homicide continued well into the 19th century because this province did not provide similar public goods.

A direct measure of each province's state capacity is not possible, so we use the *Chong* 'governability' rating (averaged across the prefectures within each province) as a proxy.[22] Our assumption is that a higher *Chong* rating implies greater geographic and strategic im-

[21] Buoye (2000) focuses on the ratio between the occurrences of property rights-related homicide and all land/debt homicide; in Guandong, that ratio rose steadily from the early to the mid-18th century and then began to decline. He reports that land rights were ambiguous in rural Guangdong even before rapid population growth increased the population-to-land ratio and hence the value of land. The combination of ambiguous property rights and increasing land values created a context for conflicts to rise in the first half of the 18th century. At the same time, higher land values also incentivized communities and officials to establish previously absent boundaries and rights ofland. After the Guangdong governor and local leaders did just that in the mid-18th century, the number and severity of land disputes declined.

[22] *Chong* represents that the prefecture was geographically and/or strategically important); See footnote 15 and texts there for detailed explanations of the various ratings.

portance of that province to the national government; this should increase the likelihood of an imperial court assigning a more capable governor to that province, from which should follow better state capacity. Our data for the *Chong* variable is from Liu (1994).

Second is the 'gentry channel', reflecting the governance of local communities by local gentry together with clan leaders. As explained by Miller (2013) and Wakeman (1998), the rise of state power often forced a retreat of local self-governance–notwithstanding the possibility, at least in theory, that high state capacity could co-exist with meaningful local self-governance by the gentry. Our hypothesis is that more local self-governance is associated with greater social order and less violence. We use the number of local gentry figures in Qing dynasty, denoted by *Gentry*, to capture the extent of local self-governance in each province. Data from Zhang (1991) is used to calculate the *Gentry* variable for each province.

Third is the 'market integration channel', through which market development reduces violence as interregional and interpersonal exchanges improve households' ability to handle distress and reduce the impetus for resorting to violence.[23] When population pressure was rising from the mid-17th century onward–or in the wake of natural disasters and crop failures–not all regions felt the same impact. Well-developed and unconstrained interregional markets should help households absorb negative shocks. The literature documents that commercial networks were indeed expanding beyond local areas in the 17th-19th centuries, although these developments characterized only some of the provinces.

Our analysis proceeds by approximating, for each year, the degree of market disintegration within a province by the coefficient of variation in grain price across its prefectures

[23] Sen (1982) argues that famine occurs not only from the lack of food but also from poor mechanisms for food distribution. For instance, the Bengal famine of 1770 was due to an urban boom that raised food prices, after which millions of rural workers starved because of lagging wages.

(*PriceCV*): the more integrated the regional grain markets, the lower the *PriceCV*. Although the rules and order imposed via state power likely reduce violence, such power may actually impede the cross-regional integration of grain markets if state agencies hinder or block the movement of goods (Anderson et al. (2013)). In this sense, then, increased state power could spur more violence and hence a higher homicide rate–especially during times of distress.

In our regressions, we use the *PriceCV* averaged over the most recent five-year period to measure the degree of market disintegration for each province in a given year. We also construct a market integration dummy, *Mkt*, which is set to1 if *PriceCV* is in the lowest quartile (and set to0 otherwise).

Our *War* variable represents the portion of a province's counties at war in each given year; the data for this calculation are from *The Chronological Timetable of Wars for Qing China*. Our regressions include, as additional controls, each province's 'governability' rating as assigned by the central government. Three 'emperor' dummies are used to control for heterogeneity (in strength of law enforcement, bureaucratic efficiency, etc.) across the periods during which different emperors ruled: *LateQianlong* for 1766-1795, *Jiaqing* for 1796-1820, and *Daoguang* for 1821-1849. Table B.2 in the Appendix reports summary statistics for all of our variables.

2.4.3 Cross-provincial regressions

The panel data used for our regressions cover 15 provinces[24] for about 30 nonconsecutive years during the period 1744-1849 (depending on data availability). As already mentioned, all post-1850 observations are omitted to preclude any bias due to the Taiping Rebellion.

[24]We follow Yan and Liu (2011) in excluding Sichuan, Yunnan, and Gansu because grain price data for these three provinces are incomplete and of low quality.

Our regression results are summarized in Table 3.4.[25] The baseline regression in column (1), which excludes provincial fixed effects (FEs), shows a positive correlation between *GPrice*, *PriceCV*, and the homicide rate; in contrast, the effect of *PopDense* on provincial homicide rates is statistically insignificant. In column (2) of the table, where we control for both province and emperor fixed effects, the coefficients for *PopDense*, *GPrice*, and *PriceCV* are all positive and statistically significant. In other words: an increase in population density, grain prices, or grain market disintegration is associated with an increase in homicide rate. This finding is consistent with our hypotheses that (a) high grain prices proxy for food-supply stress that leads some individuals to commit crimes and (b) high population density creates distress and increases violence among residents.

[25]Regression results with *robust* (i.e., heteroscedasticity-consistent) standard errors are similar to the results (with normal standard errors) reported in Table 2.

Table 2.2: Cross-provincial panel regression results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
PopDense	-0.07	1.15***	1.15***	1.13***	1.21***	1.19***	3.47***
	(0.06)	(0.25)	(0.25)	(0.25)	(0.25)	(0.25)	(0.72)
GPrice	0.33***	0.15$^+$	1.85***	0.54**	0.15$^+$	2.39***	1.22***
	(0.10)	(0.098)	(0.42)	(0.24)	(0.098)	(0.49)	(0.34)
PriceCV	0.89*	0.76*	0.82*	0.78*			
	(0.50)	(0.44)	(0.44)	(0.43)			
Mkt×*GPrice*					-0.014*	-0.017**	-0.34**
					(0.008)	(0.008)	(0.17)
Chong×*GPrice*			-2.26***			-2.41***	-0.41
			(0.55)			(0.55)	(0.52)
Gentry×*GPrice*				-0.09*		-0.11*	-0.19***
				(0.05)		(0.05)	(0.02)
War						-0.06	2.77
						(0.05)	(2.46)
Other controls	Yes	No	No	No	No	No	No
Emperor FEs	No	Yes	Yes	Yes	Yes	Yes	Yes
Province FEs	No	Yes	Yes	Yes	Yes	Yes	No
R^2	0.53	0.36	0.19	0.15	0.36	0.13	0.09
N	394	394	394	394	394	394	361

Note: Standard errors are reported in parentheses. 'Other controls' include *Longitude* and *Latitude* (for provincial capital cities), three governability ratings (*Chong*, *Pi*, and *Nan*; see Section A.2 in the Appendix), a dummy variable for southeastern provinces, and year. $+p < 0.15, *p < 0.1, **p < 0.05, ***p < 0.01$.

To see how different institutions can help mitigate the impact of food shocks to violence, columns (3)-(6) investigate how the interaction of grain prices with institutional variables affects the impact of food distress. The regression results reported in column (3) include a term for the interaction between *Chong* and *GPrice*. The coefficient for this interaction term is both negative (-2.26) and statistically significant. Recall that *Chong* represents state capacity in that this rating reflects the imperial court (presumably) assigning a more capable official to govern that region. Thus, even as higher grain prices lead to higher homicide rates, that relation is weaker when state capacity is stronger. This result supports our hypothesis that greater state capacity effectively lowers conflict and violence.

The Gentry×GPrice interaction reported in column (4) of the table yields a similar result. The negative (and statistically significant) coefficient for this term implies that more local self-governance (as proxied by more gentry members) improves social order and diminishes the effect of food distress on violence at the local level.

In column (5) we add the interaction term between *Mkt* and *GPrice*.[26] This term's coefficient is also negative and statistically significant, confirming our hypothesis that a grain market that is better integrated across regions reduces the impact of food shocks on violence. The notion of a well-connected trading network reducing violence is supported by several recent studies. For example, Burgess and Donaldson (2010) use data from 1875 to 1919 and conclude that 'the arrival of railroads in Indian districts dramatically constrained the ability of rainfall shocks to cause famine in colonial India.' Cao and Chen (2016) treats the 1826 abandonment of China's Grand Canal as a natural experiment; these authors find that the abandonment significantly increased the frequency of rebellions in counties bordering the canal–a result of the subsequent collapse of the interregional trade network.

The regression whose results are reported in column (6) of Table 2.2 includes all interaction terms in addition to our *War* variable. In this regression, the coefficients for the

[26] Because *Mkt* is generated from *PriceCV*, these two variables are highly correlated. We therefore exclude *PriceCV* from column (5) because the main purpose of that regression is to derive the coefficient for the $Mkt \times GPrice interaction$.

interactions of *GPrice* with *Mkt*, *Chong*, and *Gentry* are robust and similar to those seen in columns (3)-(5). The coefficient for *War* is negative but statistically insignificant–perhaps because our period of study here ends in 1849 and so does not include the Taiping Rebellion. Because there are only 15 (of 420) provincial observations for which the *War* dummy is nonzero, one can hardly expect that variable to have a first-order effect on the homicide rate during this period (i.e., from 1744 to 1849).

For the regression in column (7), we use the first-difference terms for *homicide rate*, *Popdense*, *GPrice* and *War* in order to address the issue due to the potential presence of unit roots.[27] The basic conclusions from the previous regressions remain robust, except that the interaction term of *GPrice* with *Chong* now becomes statistically insignificant.

2.5 Explanatory narrative of the national homicide trend

In the previous section we relied on provincial panel data to investigate drivers of differences in homicide rate among the provinces. In this section we use those cross-sectional findings to shed light on the upward and downward trends in the national homicide rate of Qing China.

First of all, the Chinese population suffered heavy losses during the civil wars that raged from the mid-16th to the mid-17th century; this warfare led to the Ming Dynasty's demise and to the Qing Dynasty's founding in 1644. According to Cao (2001) and as shown in Table 2.3, China's population was160 million in 1678 and thereafter grew at an annual rate of 0.664%, reaching 306.6million by1776. Population growth continued until the Taiping Rebel-

[27]The Fisher-type unit root test for our dependent variable which we conducted can not reject the null hypothesis that there exists a unit root. The homicide rate series, however, is difference-stationary, that is, its first difference does not have a unit root. The results in column (7) are obtained under standard robust errors. We have also tried to add *year* and *year squared* into the regression in column (6), but the results are qualitatively robust to the inclusion of these extra terms.

lion during the 1850s-1860s, when large-scale casualties caused the population to decline by 17% between 1851 and 1880. Population growth resumed after 1880 and recovered all of the civil war losses by 1910. Table 2.3 shows that China's population increased by 35% between 1678 and 1820 and then declined by 5.8% between 1820 and 1880. It follows from our results in Section IV that the pre-1821 rise in national homicide was likely due, at least in part, to the rapid rise in population pressure during that time,[28] whereas the post-1821 decline in homicide was due to slower growth (or no growth) in population for several decades.

Table 2.3: Population of China (millions), 17th-19th centuries

Year	1678	1776	1820	1851	1880	1910
Population	160.0.	306.6	377.1	428.2	355.0	421.6
Annual growth rate (%)	0.664	0.471	0.410	-0.646	0.573	–

Note: Reported figures are based on data in Cao (2001).

Our second insight is illustrated in Figure 2.8, which plots two time series for the 1736-1895 period: the annual homicide rate and the national grain price (each averaged over 10-year intervals). Here the national grain price for a given year is defined as grain prices averaged across all prefectures. The grain price exhibits a rising trend with considerable volatility. Yet close examination reveals co-movement in the two measures *except* during the civil war years. Based on estimation results in Table 2, when grain prices increase by 1%, the homicide rate goes up by 0.15%.[29] The national grain prices rose by 48.6% from 1744 to 1821, which translates into an increase of 0.07 cases in homicide rate, accounting for about

[28]Our estimation results in Table 2 show that a 1% increase in population density is associated with an increase of 1.19% in homicide rate. By multiplying the population-density coefficient with the actual rate of change in population density, we estimate that from 1721 to 1821, the population change increased the homicide rate by 0.56 cases per 100,000 population. Since the actual homicide rate went from 0.36 in 1720 to 1.47 in 1820, the population-pressure effect probably accounted for 50.4% of the period's net increase in homicide rate.

[29]Note that changes in grain prices affect the homicide rate both directly and indirectly through the interaction terms. In our estimation exercise here on the effect of the grain price changes, we only use the direct effect for simplicity.

11% of the actual homicide-rate increase for the period. Similarly estimated, the decline in grain prices during the 1821-1850 period contributed 3% to the period's net decline in homicide rate.

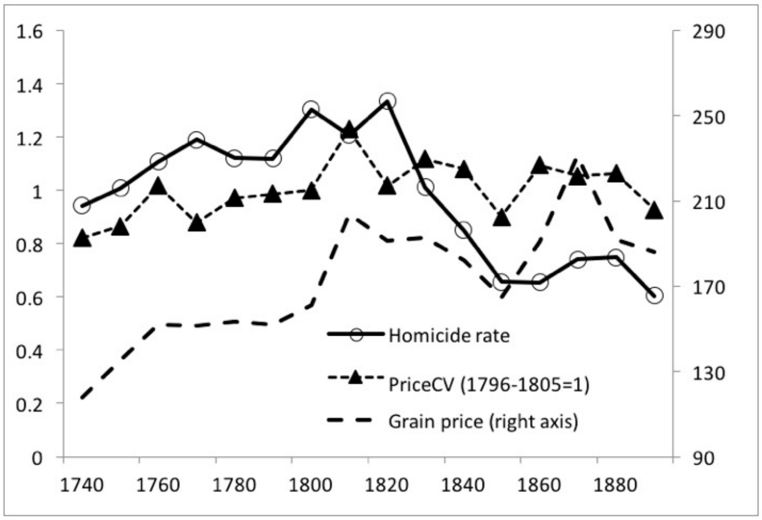


Figure 2.8: Grain price, coefficient of variation, and national homicide rate

Note: 'Homicide rate' represents the annual rates in Figure 2.3 averaged over 10-year intervals. 'Grain price' is the decadal average grain price of all prefectures (units: silver tael per shi). '*PriceCV*' is the decadal average coefficient of variation for grain prices across prefectures.

A positive correlation between homicide and the overall price level is also found for England, where the homicide rate declined in periods with stable prices but increased in periods of grain price instability (**?**, p.309). As grain and other prices increased, wages remained 'sticky' and did not increase as fast or by as much; the result was a decline in real purchasing power for ordinary people, which in turn led to more violence and homicide. In this sense, the long-term historical experience was similar in England and Qing China.

Third, changes in Qing China's interregional market disintegration may also have contributed to homicide rate variations. Figure 2.8 shows the decadal average national *PriceCV* of grain prices across all Chinese prefectures (for which data are available). Two patterns stand out. First, *PriceCV* and the national homicide rate exhibited significant co-movement during the 18th-19th centuries. Second, market disintegration became increasingly more severe in the 18th century until about 1821–in sharp contrast to the literature's consensus view that China's market development was improving rapidly prior to the 19th century.[30] As explained in Section IV, grain market disintegration makes it more difficult for cross-regional arbitrageurs to transport grains to areas hit by crop failure or other natural disasters; the result is an increase in interpersonal violence. Economically, the estimation results in Table 3.4 suggest that the change in market disintegration accounted for 17% of the increase in homicide rate from the 1740s to the 1810s and for 10% of the decrease from the 1820s to 1849.[31]

To understand why regional markets were becoming more disintegrated in 18th-century China, we refer to work by Bernhofena et al. (2015). These authors also find that China experienced a prolonged process of market disintegration from 1740 to 1821. This disintegration trend is robust and holds even after removing, from the prefectural price series, the effects of common exogenous shocks within regional and agro-climatic boundaries. Despite the Qing state's innovations meant to liberalize markets by establishing both government and civilian granaries to stabilize the grain supply, government power was also employed to interfere with markets in a number of ways: direct control of supply and marketing, policing of supplies, forced sales at reduced prices, and disaster relief (**?**). In particular, Bernhofena et al. (2015) conclude that physical barriers–such as setting up checkpoints along borders (between provinces or across prefectures) to prevent grain transportation by speculators–were among the most significant drivers of disintegration.[32]The Qing government outlawed

[30]According to **?**, p. 70), '18th-century China' came closer to resembling the neoclassical ideal of a market economy than did western Europe' (cf. **?**). That description is clearly at odds with our findings.

[31]Here, the exact cutoff year used is 1814 as $PriceCV$ rose until the 1810s but dropped in the 1820s. From Table 3.4, we take 0.8 as the coefficient for $PriceCV$.

[32]Cheung (2008) reports many instances, during the second half of the 18th century, where the emperor warned provincial governors to permit cross-regional grain exports lest they be punished for imposing export bans. These instances indicate that barriers to the movement of cross-regional goods must have been

grain hoarding and speculation because such arbitrage activities were viewed as being harmful to society in the longrun.

Of course, Qing officials could restrict market flows only if the state had enough power and control. The balance of power between the state and society (as represented by the rural gentry) went through fluctuations that mirrored the rise-then-fall pattern in both homicide and market disintegration during Qing China. Miller (2013) reports that, from 1572 until the collapse of the Ming Dynasty in 1644, Ming emperors repeatedly but unsuccessfully tried to consolidate power in the state by weakening the governance roles of the gentry and other non-official players. Emperors of the new Qing Dynasty inherited the Ming Dynasty's governance structure, under which the state had little control over local affairs. Hence the Qing emperors soon faced the same struggle with local gentry and the larger civil society. Starting in 1661, the emperors Kangxi (1661-1722), Yongzheng (1722-1735), and Qianlong (1735-1796) launched a series of successful efforts to concentrate power in the state. By the end of the 18th century, China's state power was at its peak and civil society had shrunk considerably (Wakeman, 1998). In this process of power consolidation by the state, which spanned nearly 150 years, not much room was given to develop bottom-up, self-governing institutions and rules. It follows that societies dominated by a powerful state may be poorly equipped to deal with distress events, which makes it more likely that the society's members will employ violence as a means of securing their ends (Anderson et al. (2013); Bai and sing Kung (2011); Jia (2013)).

So even though the rise in state power from 1661 until the end of the 18th century should have led to a declining homicide rate via the 'state capacity' channel, the negative effect of state power on local self-governance and on grain market integration probably made it

widespread, since otherwise the emperor would not have felt impelled to speak against them so frequently.

harder for Chinese localities to cope with food distress. Thus the overall effect of increased state power may actually have contributed to the homicide rate's rise from 1661 until the early 19th century.

Near the start of the 19th century, state power in China began to retreat and so the balance of power shifted once again in favor of market and self-governing institutions (**?**; Wakeman (1998)). Figure 2.8 reveals that this is about when market disintegration, grain prices, and the national homicide rate all began to decline. The 19th-century experience is therefore largely consistent with our explanation that more room for market and self-governance institutions permits both market and social solutions to offer relief, reducing the impetus for citizens resorting to violence. Market development is typically an adaptive process characterized by many trial-and-error steps. That is, market institutions and rules that are more egalitarian do not appear automatically; rather, they are innovative responses to conflicts that arise because of their absence ('conflict then order').[33] Hence it is not surprising that the increase in homicides from the mid-17th to the early 19th century was followed by a decline in homicides.

2.6 Conclusion

In this paper we have constructed Qing China's homicide history, examined its trends, and used cross-provincial data to investigate possible drivers of its evolution. This exercise is largely supportive of five hypotheses concerning how population change, food distress, state capacity, gentry governance, and cross-regional grain market integration affect incentives for violence and homicide. Although rapid population growth and rising grain prices probably

[33] North et al. (2009) theorize that violence ultimately leads to political and economic institutional development that creates a more stable social order that has less violence but could be either better or worse for economic development. Adaptive institutional development can require multiple trial-and-error rounds before violence is significantly reduced

contributed to the increased violence in 19th-century China, the steady rise of state power during that period may have stifled both local self-governance and cross-regional market integration, thereby undermining the ability of the local community and the market to mitigate the impact of risk events on ordinary people's lives–thus leading more people, especially those who experienced marginalization, to resort to violence for survival. By the same token, weakening state power in the 19th century may have resulted in more room both for the gentry and for market forces to play their civilizing and pacifying roles, reducing the extent of violence during that period. Our work has thus shed new light on Qing China's history: the rapid economic and population growth throughout the Kangxi-Qianlong Prosperity was at the cost of increased property insecurity and rates of violence. These findings have also enriched our understanding of the socioeconomic drivers of violence.

Our paper contributes to the literature not only by establishing China's homicide rate history from 1661 to 1898 but also in other ways. We establish that, at least from the mid-17th century to the late 19th century, China enjoyed a lower homicide rate than Western Europe–with the latter not approaching the former until late in the 19th century. This quantitative finding has implications for researchers in the field of comparative civilizations and also for the intellectual debate on the divergence between East and West. Local gentry-based governance structure for communities below the county level may have contributed to China's lower violence rates.

As China's First National Historical Archives and other archival sources have become more available in digital formats, new research efforts are clearly feasible. These sources not only offer new opportunities for research on various economic aspects of Qing China (e.g., financial contracting, interest rates, marriage patterns, family structures) but also allow us to re-examine some conventional beliefs about life in China during that era. So even though the Kangxi-Qianlong Prosperity may have been the best 'boom' period ever experienced by the Chinese economy, it was characterized also by increasing homicide and general violence (although the homicide rate during this period was not especially high). More efforts are required to establish causal linkages and to gain a fuller understanding of China's long-term history of interpersonal violence and economic development.

Chapter 3

Labor Unions and the Labor Wedge: A Macroeconomic Perspective

Lijun Zhu

3.1 Introduction

General equilibrium based macroeconomic models build on two pillars: household's present value utility maximization, and firm's profit maximization. Taken prices as given, the household and firm optimally make consumption/production and work/hiring decisions. In equilibrium, demand equates supply, i.e. the regular marginal condition in each market holds, and markets clear. The deviation from equilibrium conditions, i.e. the discrepancy between the marginal condition between households and firms, wedges as labeled in the business cycle literature (Chari et al. (2007)), provides a natural metric to investigate market (in-)efficiency.

In this paper, we focus on labor wedge, the wedge between the marginal rate of substitution of consumption for leisure ($MRS_{c,n}$), and marginal productivity of labor (MP_n).

Without distortions, labor market efficiency requires that

$$MRS_{c,n} = MP_n$$

The labor wedge measures the violation of this condition, which could be a result of inefficiency in either the labor market or other markets. Our paper investigates the effect of a specific form of labor market institution, labor unions, on the labor wedge. The main intuition lies in the following formula

$$MRS_{c,n} = W^c < W^u$$

with W^c and W^u denoting wages in the competitive market and wages controlled by labor unions. From a macro perspective, the marginal rate of substitution, or people's willingness to work, equals to the wage level in the competitive market. Labor unions, however, demand a wage premium and creates a wedge between unionized wages and the marginal rate of substitution[1].

Figure 3.1 presents the long run trend of the labor wedge in the US from the third quarter of 1947, the earliest date relevant data is available, to the third quarter of 2007. The series is truncated at 2007 to avoid the effect of the Great Recession.[2] The labor wedge is relatively stable from 1940s to 70s, and has declined continuously since around 1980. The Wald test for structural break with unknown break point shows that $1977-Q3$ is the break point[3]. On the

[1]The labor wedge $\frac{MP}{MRS} = \frac{MP}{W}\frac{W}{MRS}$. In this paper, there is no wedge between MP and W. In Cobb-Douglas production function, Marginal productivity is proportional to the average productivity, AP, and the trend of $\frac{AP}{W}$ reflects the behavior of the labor share. While the latter only declines at a small magnitude after 1980s, the relatively large decline in the labor wedge comes mainly from the household side wedge.

[2]See section II for details of measurement. It is well known that the labor wedge has a countercyclical pattern, and rises in recessions. The measured labor wedge increases from 2007 to 2010 (see Appendix for the overall pattern.

[3]See appendix for details. The time series trend for labor wedge is sightly increasing before 1977-Q3 and decreasing after that.

other hand, the union density in U.S., measured as the percentage of union members in total employment, follows a similar trend: relatively stable until around 1980, and declines after that[4]. The correlation coefficient between the two series, measured in a yearly frequency, from 1947 to 2007 is 0.75 and statistically significant.

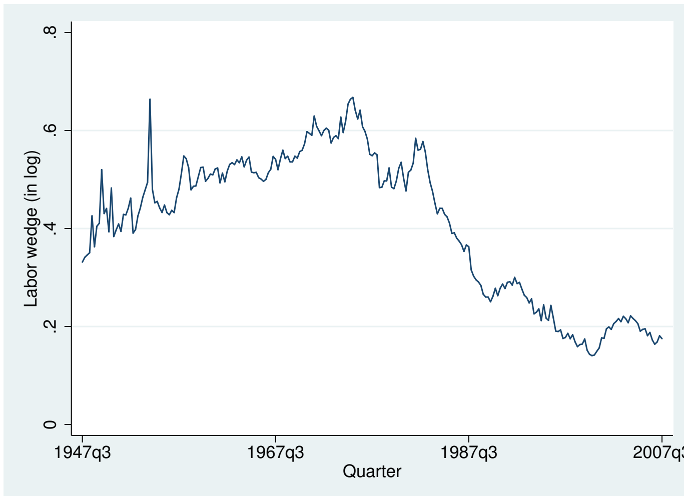


Figure 3.1: The Labor Wedge, 1947Q3-2007Q3

The labor wedge has been a focus of research especially in one branch of the business cycle literature. Rotemberg and Woodford (1991) defines markup as the wedge between marginal product of labor and wage, which is the first component of labor wedge, with the second the gap between wage and marginal rate of substitution, and documents that the markup has a countercyclical pattern. Hall (1997) decomposes the labor market equilibrium condition, i.e. marginal product of labor equals marginal rate of substitution in order to investigate the sources of fluctuation of hours. The residual in their decomposition equation[5] is essentially labor wedge. Gali et al. (2007) uses the gap between marginal product of labor and the marginal rate of substitution[6] as a measure of economic efficiency, and uses it to calculate the efficiency cost of business fluctuations. In an influential paper, Chari et al.

[4]See appendix Figure C.1 for the time series pattern of union density

[5]Equation 3.2 in paper

[6]The inefficiency gap defined this way equals to the negative labor wedge

(2007) builds four reduced-form wedges, corresponding to productivity, labor, investment, and government expenditure, into the stochastic Neoclassical growth model, and finds that efficiency and labor wedges account for most of the fluctuations over business cycles.

Chari et al. (2007) has generated a series of research that has the labor wedge as the central focus. Several papers, in particular, examined the effect of labor market frictions on the labor wedge. Shimer (2009) reviews the labor wedge literature and suggests that search friction combined with real wage rigidities are promising explanations for endogenous cyclical labor wedge. Pescatori and Tasci (2011) augments the benchmark RBC model with a labor market featuring search frictions, in which employed workers and firms bargain over both wage and working hours. According to their results, the search friction itself doesn't not cause variation in labor wedge over the business cycle since the effect of searching friction is completely absorbed by wage instead of working hours. Cheremukhin and Restrepo-Echavarria (2014) decomposes the labor wedge, associated with search frictions, and finds that fluctuations in matching efficiency account for 90% of variations in the labor wedge. We share the same focus of labor wedge as this line of literature. Our paper, however, concerns the long run trend, instead of cyclical patterns, of the labor wedge.

This paper is also closely related to a small but growing set of papers that incorporate union into macroeconomic general equilibrium models. Observing the inverted-U-shape of union density and U-shape of inequality in U.S. over the 20th century, Dinlersoz and Greenwood (2016b) develops a model of union in a general equilibrium framework. In their model, A continuum of firms with varied productivities hire skilled and unskilled labor in a competitive labor market. Only unskilled labor can be unionized. Bearing an organizing cost, unions target firms with relatively high productivities. In that framework, skilled-biased technological change (or unskilled-biased in early 20th century) generates simultaneously the inverted-U-shape of Union density and U-shape of inequality in U.S. in the 20th century. Rudanko and Krusell (2015) models a monopoly union in an economy with search frictions. Rather than determined by bargaining between firms and workers, the wage is set unilaterally by a universal-coverage union which values the welfare of both employed and unemployed workers. While making wage proposals, the Union takes their effect on job creation into account. Efficiency is achieved in all but the initial periods in the case the union fully commits to proposed wages. Without full commitment, employment is lower than efficient levels in both the short and long run. Taschereau-Dumouchel (2015) investigates the effect of Unions and the threat of unionization on wage distribution. Non-unionized firms respond to the Union threat by hiring more anti-union high skilled workers, and less low-skill ones who supports unionization. That strategy endogenously compresses the wage distribution in non-unionized firms under the assumption of decreasing return to scale for each level of skills.

Perhaps the closest paper to ours is Cole and Ohanian (2004). That paper is motivated by the fact that, during the Great Depression, consumption is significantly below trend in 1939, comparing to its 1929 level, while leisure time (non-working time) and wage are much above trend. These combined, don't satisfy the marginal condition of labor supply, $\frac{c}{\ell} = w$ in the case of log utilities[7]. Cole and Ohanian (2004) argues that the increasing influence of labor unions in 1930s is responsible for this divergence. In their framework, there are two sectors, one competitive and the other unionized. Insiders in the unionized sector determine the size of union and the wage premium each period. Outside workers has to wait to be rationed a position in order to enter the unionized sector. The rationing creates a wedge between wage and household's marginal rate of substitution, with the gap reflecting the value of waiting. Different from Cole and Ohanian (2004), in our paper where unions control both

[7] Cole and Ohanian (2004) doesn't use the term 'labor wedge' explicitly in their paper. However, the gap between the left and right hand side of Household's marginal condition is the first component of labor wedge

wage and employment, in our model, Unions decides wages, and unionized firms optimally post vacancies, taking wages as given.

The rest of paper is organized as following: section 2 provides a detailed description of the measurement of labor wedge; sector level evidence that supports the connection between union power and labor wedge is presented in section 3. Section 4 lays out the full model. The quantitative results of model are explored in section 5. Section 6 concludes.

3.2 Measurement of the Labor Wedge

The section details on the measurement of the labor wedge. The procedure here follows Shimer (2009) which itself adopts the approach commonly used in the labor wedge literature. The economy features a representative household and firm. Time is discrete and infinite. The representative household's problem is to maximize lifetime utility given by

$$\Sigma_{t=1}^{\infty}\beta^t(\log c_t - \frac{\gamma\epsilon}{1+\epsilon}n_t^{\frac{1+\epsilon}{\epsilon}})$$

where β is the discount rate, and c_t and n_t denotes consumption and working hours respectively[8]. γ measures the disutility of working, while $\epsilon > 0$ is the Frisch elasticity of labor supply. The household respects its period budget constraint

$$c_t + k_{t+1} - (1-\delta)k_t \leq r_t k_t + w_t n_t$$

Denote λ_t the Lagrange multiplier for the budget constraint. First order conditions for

[8]It is assumed that the (dis)utilities from consumption and working are separable, with the former in the form of log, and the latter CRRA. This specific functional form is to ensure the existence of a balanced growth path

consumption and labor supply are given by

$$\frac{1}{c_t} = \lambda_t$$
$$\gamma {n_t}^{\frac{1}{\epsilon}} = \lambda_t w_t$$

A combination of the two first order conditions above leads to

$$w_t = \gamma c_t {n_t}^{\frac{1}{\epsilon}} \tag{3.1}$$

Assume a Cobb-Douglas production technology. The representative firm's problem is standard, rent capital and hire labor in spot markets and maximizes period profit

$$\max_{k_t, n_t} \quad A_t {k_t}^{\alpha} {n_t}^{1-\alpha} - r_t k_t - w_t n_t$$

The firm's optimal condition corresponding to its choice of labor reads

$$w_t = (1-\alpha)\frac{y_t}{n_t} \tag{3.2}$$

where $y_t = A_t k_t^{\alpha} n_t^{1-\alpha}$ denotes total products.

Combining the labor supply and demand, i.e. (1) and (2), yields the standard labor market equilibrium condition. Define the labor wedge as $\tau_t \equiv \log(\frac{MP_n}{MRS_{c,n}})$[9], and it satisfies

$$\tau_t \equiv \log(\frac{MP_n}{MRS_{c,n}}) = \log\frac{1-\alpha}{\gamma} + \log\frac{y_t}{c_t} - (1+\frac{1}{\epsilon})\log n_t \tag{3.3}$$

Data from the following sources are utilized to measure the labor wedge

[9]Note that the labor wedge can also be measured as $\tau_t = 1 - \frac{MRS_{c,n}}{MP_n}$. The two measures give very similar long run trends. We choose the measurement above since it is in logarithm and unit free.

- n_t, hours time employment-population ratio, both taken from Cociuba et al. (2012)[10]

- c_t, nominal personal consumption expenditure, and $y(t)$, nominal GDP, from NIPA; used to produce consumption-income ratio on the right hand side of equation (3).

- The value of $\frac{\gamma}{1-\alpha}$, acting as a shift coefficient, does not affect the trend over time, which is the focus of the current paper. This value is chosen such that the average labor wedge equals 0.4.

For the baseline case, we pick 1 as the value for the Frisch elasticity of labor supply. The measured labor wedge is presented in Figure 3.1. Note that from the definition here, a deviation from the Cobb-Douglas technology, e.g. a decreasing-return-to-scale production function, does not change the trend of the labor wedge.

The representative household's marginal condition reads

$$\gamma c n^{\epsilon} = w$$

or equivalently,

$$\gamma \frac{c}{y} n^{\epsilon+1} = \frac{wn}{y}$$

The right hand side is the labor share, which declines but at a relatively small magnitude. In this paper, we take it as constant, and focus on the behavior of the left hand side. Figures C.4 and C.5 in appendix provide the trend of working hours and Consumption-Income ratios in U.S. from 1947 onwards. Both demonstrate an increasing trend from 1970s to 2000s. Table 3.1 lists their values for 1970s and 2000s.

[10] We have tried to use instead average hours time employment-labour force ratio and average hours only. The decrease from 1970s onwards is smaller in both case. In the paper, however, we follow the literature (Shimer (2009) etc.) and use average hours times employment-population ratio.

Table 3.1: Working Hours and C-Y ratio

	1970-1979	2000-2007
Hours-1	24.56	28.04
Hours-2	39.83	42.29
C-Y ratio-1	60.46	67.08
C-Y ratio-2	77.32	81.89

Both increases in hours and the consumption income ratio increases the measured MRS, i.e. $\gamma\frac{c}{y}n^{\epsilon+1}$. Table 3.2 shows the relative change in MRS from 1970s to 2000s, for different values of ϵ, and under different measures. We use 0.17 as the benchmark.

Table 3.2: Changes in MRS

	Measure-I	Measure-II
$\epsilon = 1$	0.34	0.17
$\epsilon = 0.5$	0.44	0.24
$\epsilon = 2$	0.29	0.15

3.3 Sector Level Evidence

This section provides sectoral level evidence for the union-wedge connection. We have constructed a database which covers 75 manufacturing sectors from 2005 to 2014[11]. Data on union density comes from the Current Population Survey (i.e. CPS). Union density is measured as the fraction of union members in total wage and salary earners in each sector. The industrial classification in CPS is based on 2000 Census Industry Code. One industry in 2000 Census Code might correspond to one or several 3, 4, 5 or 6 digit NAICS sectors. In total, there are 75 Manufacturing sectors according to 2000 Census Code. Table 3.3 presents the relevant summary statistics.

[11] We choose this interval since Annual Survey of Manufactures only provide publicly accessible data from 2005 onwards.

Table 3.3: Summary statistics: Union density

Year	Mean	Std. Dev.	Min.	Max.	Obs.
2005	13.4%	7.8	1.3	35.2	75
2009	11.3	8.1	0	32.5	75
2014	10	8.2	0	42.3	75

Data Source: CPS and Annual Survey of Manufactures.

The union density is in decline in the U.S.. As can be seen from the table above, over the last decade and across manufacturing sectors, union density decreases from 13.4% to 10%. On the other hand, there are still relatively big variations in union density across sectors. The standard deviation of union density is stable at around 8, a relatively big number. Even in 2014, union members in the highest manufacturing sector accounts for over 40% of total employment. The lower bound of union density reaches its lowest possible value, 0%, in 2014. The relatively large variation in union density provides us the opportunity to investigate its effect on labor wedge.

To calculate sectoral labor wedge, we extracts data on value added and hours from Annual Survey of Manufactures (i.e. ASM). ASM uses NAICS codes to define sectors (3, 4, 5, and 6 digit). We merge the two data sets using the industry crosswalk tables from the census website. To incorporate multiple sectors, we adopt the following utility function.

$$U(c_t, n_{i,t}) = \log(c_t) - \gamma \frac{\epsilon}{1+\epsilon} \Sigma_{i=1}^{N} n_{i,t}^{\frac{1+\epsilon}{\epsilon}}$$

where c_t denotes aggregate consumption[12]. Consumption for disaggregated sectors would be very difficult, if not impossible, to obtain. The advantage of using the utility function

[12] See appendix for the derivation of the utility function specified here.

above is that, only aggregate consumption is needed to calculate sectoral level labor wedge. Labor wedge for sector i in year t is then measured as

$$\begin{aligned}\tau_{i,t} \equiv \log \frac{MP_{n_{i,t}}}{MRS_{c,n_{i,t}}} &= \log \alpha \frac{y_{i,t}}{n_{i,t}} - \log \frac{\gamma n_{i,t}^{1/\epsilon}}{1/c_t} \\ &= \log y_{i,t} - (1 + 1/\epsilon) \log n_{i,t} - \log(c_t) + constant\end{aligned}$$

$y_{i,t}$ and $n_{i,t}$ denote sectoral level value added (deflated by GDP deflator) and employee hours respectively[13], both from Annual Survey of Manufactures, 2005-2014. c_t denotes personal consumption expenditure in 2009 dollar.

Merging the two datasets produces a panel dataset for 75 sectors and over 10 years[14]. To test whether a higher union density leads to a larger labor wedge, we run the following regression

$$\tau_{i,t} = \beta_0 + \beta_1 * union_{i,t} + \beta_2 * X_i + \beta_3 * D_t + \beta_4 * Z_{i,t} + \epsilon_{i,t}$$

- D_t: year dummies, 2005-2014;

- X_i: sectoral controls, including mean and standard deviation of annual growth rates (in value added and employment) from 2005 to 2014 for each sector;

- $Z_{i,t}$: ratio of production to non-production employees.

The baseline value of the Frisch elasticity of labor supply is set to be 1. Table 3.4 presents the baseline regression results.

[13] calculated as $\frac{total\ employment}{production\ workers} \times production\ worker\ hours$

[14] For 2012, 2013, and 2014, one new sector, 1190, is added into the census code. That is why in later regressions, there are 753 instead of 750 observations

Table 3.4: Dependent var.: Labor wedge ($\epsilon = 1$, in log)

	(1)	(2)	(3)
union density	0.011**	0.027***	0.030***
	(0.006)	(0.006)	(0.006)
$\frac{prod.emp.}{non-prod.emp.}$		-0.109***	-0.111***
		(0.017)	(0.030)
$\Delta y - mean$		-0.067***	-0.068***
		(0.009)	(0.01)
$\Delta y - st.dev.$		0.025***	0.026***
		(0.003)	(0.004)
Year Dummy	No	No	Yes
R^2	0.006	0.08	0.10
Obs.	753	753	753

Note: $*p < 0.1; **p < 0.05; ***p < 0.01$. Data Source: CPS and Annual Survey of Manufactures, 2005-2014.

Column (1) shows the univariate regression result, with sectoral labor wedge and union density as the dependent and independent variables, respectively[15]. It can be seen that the effect of union density on labor wedge is positive and significant at 5% confidence level. The coefficient values at 0.011, which means that a 1% increase in union density leads to about 1% increase in labor wedge.

In the second column, several variables are added to the regression to control for sectoral level heterogeneities. In particular, we add *the ratio of production to non-production*

[15]Note that labor wedge is measured in log. We don't use log for union density since it is already in percentage and unit-free.

workers, *average growth rate of sectoral value added*, and *standard deviation of growth rate of value added*. The fraction of production workers is correlated with union density if production workers are more likely to be unionized. It might also affect the aggregate efficiency as represented by the marginal conditions if the two set of workers are subject to different compensation rules in reality. Whether a sector is high-growth or low-growth, and whether the growth rates are relatively stable over time, are also taken as controls.

The positive and significant effect of the union density is robust to additional controls, and the coefficient increases to 0.027. This effect is stronger than that in the single variable regression. In addition, labor wedge tends to be lower in sectors with higher fraction of production workers, higher average growth rate, and where growth rates are relatively stable over time. It is well known that labor wedge has a strong cyclical pattern. To control for cyclical fluctuations, we add year dummies into the regression. The results are presented in the third regression. All coefficients from column (2) are robust[16].

The value of the Frisch elasticity of labor supply, ϵ, is critical to determine the response of labor supply to the change in wage rate. In our baseline regression, we set that $\epsilon = 1$. In the RBC literature, a value as large as 4 has been employed to match the relatively large effect of wage changes on aggregate labor supply. Micro evidence, however, usually supports a value of ϵ smaller than 1. To verify the robustness of the baseline results, we vary the value of ϵ, and present regression results for $\epsilon = 0.5$ and $\epsilon = 2$ in Table 3.5. It shows very similar results as baseline cases (i.e. $\epsilon = 1$), with the only exception that for $\epsilon = 0.5$ and single variable regression, the effect of union density becomes insignificant.

[16]We have tried to measure the union density as percentage of wage earners whose wage contract are covered by unions' collective bargaining. The results presented above are robust to this alternative measure. We have also tried to use unemployment rate, instead of year dummies, to control for cyclic fluctuations, and found similar patterns. The qualitative results remain if mean and standard deviation of growth rates in employment (instead of valued added as in the baseline case) are used as sectoral level controls

Table 3.5: Dependent var.: Labor wedge (in log)

	$\epsilon = 0.5$			$\epsilon = 2$		
	I-(1)	I-(2)	I-(3)	II-(1)	II-(2)	II-(3)
union density	0.005	0.035***	0.040***	0.015***	0.022***	0.025***
	(0.01)	(0.01)	(0.010)	(0.004)	(0.004)	(0.004)
$\frac{prod.emp.}{non-prod.emp.}$		-0.032	-0.034		-0.148***	-0.149***
		(0.010)	(0.053)		(0.017)	(0.020)
$\Delta y - mean$		-0.15***	-0.15***		-0.026***	-0.028***
		(0.009)	(0.018)		(0.006)	(0.007)
$\Delta y - st.dev.$		0.056***	0.057***		0.01***	0.01***
		(0.003)	(0.006)		(0.002)	(0.002)
Year Dummy	No	No	Yes	No	No	Yes
R^2	0.0004	0.10	0.12	0.02	0.09	0.12
Obs.	753	753	753	753	753	753

Note: $*p < 0.1; **p < 0.05; ***p < 0.01$. Data Source: CPS and Annual Survey of Manufactures, 2005-2014.

3.4 The Model

This section presents the model, which incorporates union into an otherwise standard Neoclassical growth model. The economy features two sectors, one competitive and the other unionized. In the competitive sector, wage and employment are determined by the usual supple and demand. Wages in the unionized sector are controlled by a monopoly union. The union values both the wage premium and membership size. The higher wage in the unionized sector invites an application queue, and jobs are rationed by the labor union. The labor rationing process and associated waiting value create a wedge between wage and worker's willing to work.

Time is discrete and infinite. Within the representative household, some household members work in the competitive sector, and some search and work for unionized firms. As standard, household owns capital and makes investment decision. The objective of the representative household is to maximize lifetime utility, i.e.

$$\max \quad \Sigma_{t=0}^{\infty}\beta^t(u(c_t) - \nu(n_t)) \tag{3.4}$$

c_t and n_t are consumption and non-leisure time, respectively. There are a continuum of firms, with index $i \in [0, 1]$ in the economy. Firms locate in $[0, \phi], 0 < \phi < 1$, are unionized, and the $(\phi, 1]$ range behave competitively. Denotes $n_t(i)$ the employment of firm i. Since there would be wage premiums for unionized works, outsiders have to queue, and wait to be rationed a union position. The rationing process models the practice of membership restrictions and organization costs (such as certification elections) of labor unions in the real world. Denote $q_t(i)$, which is determined in equilibrium, the probability of getting a union job, total working hours is given by

$$n_t = \int_0^{\phi} \frac{n_t(i)}{q_t(i)} di + \int_{\phi}^{1} n_t(i) di \tag{3.5}$$

Household income includes wage income of workers in both competitive and unionized sectors, capital rental income, and firms' profits in both sectors. The budget constraint for the household is

$$c_t + i_t = \int_0^{\phi} (w_t^u n_t(i) + \Pi_t(i))di + \int_{\phi}^{1} (w_t^c n_t(i) + \Pi_t(i))di + r_t \int_0^{1} k_t(i)di, \tag{3.6}$$

where w_t^u and w_t^c are wages in the unionized and competitive sector, and $\Pi_t(i)$ profits. Denote

$k_t = \int_0^1 k_t(i)di$ the aggregate capital stock, and i_t investment, the following law of motion for capital holds

$$k_{t+1} = (1 - \delta)k_t + i_t \tag{3.7}$$

The goal of the representative household is then to maximize (4), subject to constraints (5), (6), and (7).

The final goods is produced by combining the composite goods of the competitive and unionized sectors, y_t^c and y_t^u, according to

$$Y_t = (y_t^{u\rho} + y_t^{c\rho})^{\frac{1}{\rho}}$$

ρ governs the elasticity of substitution between sectoral goods, which itself is an aggregation of intermediated goods

$$y_t^u = (\int_0^\phi y_t(i)^\zeta di)^{\frac{1}{\zeta}}; \qquad y_t^c = (\int_\phi^1 y_t(i)^\zeta di)^{\frac{1}{\zeta}}.$$

ζ determines elasticities of substitution among firms within each sector. We use the final goods as numeraire and normalize its price to 1. Denote $p_t(i)$ the price of the intermediate goods produced by firm i, the final goods producers' problem is

$$\max \quad Y_t - \int_0^\phi p_t(i)y_t(i)di - \int_\phi^1 p_t(i)y_t(i)di \tag{3.8}$$

In both sectors, intermediate goods producing firms take factor prices as given, and optimally make hiring decisions. Firms in both sectors solve the following standard optimization problem

$$\Pi_t(i) \equiv \max_{k_t(i), n_t(i)} \quad p_t(i)F^i(k_t(i), n_t(i)) - r_t k_t(i) - w_t n_t(i) \tag{3.9}$$

Note that $w(t) = w^u(t)$ for $0 < i < \phi$, and $w(t) = w^c(t)$ if $\phi < i < 1$. The differences between the competitive and unionized sectors lie in how wages are determined. In the competitive sector, wages are determined by marginal conditions. Unionized wage is controlled by a monopoly union, whose objective is to

$$\max_{w^u{}_{t=0}^{\infty}} \Sigma_{t=0}^{\infty} \beta^t G(w_t^u - w_t^c, \int_0^{\phi} n_t(i) di) \tag{3.10}$$

That is, the union values both the wage premium, $w_t^u - w_t^c$, and the size of union members. The union takes into account the hiring behavior of firms while making wage proposals. Denote $D(w_t^u, [black, fill = black](0,0)circle(.3ex);)$ the labor demand function of unionized firms, and the union respects the following constraint

$$n_t^i = D(w_t^u, [black, fill = black](0,0)circle(.3ex);), \quad for\ 0 < i < \phi \tag{3.11}$$

In addition, firms in the unionized sector should maintain a nonnegative profit

$$\Pi_t(i, w_t^u, [black, fill = black](0,0)circle(.3ex);) \geq 0, \quad for\ 0 < i < \phi$$

The dynamic competitive equilibrium of the economy consists of a sequence of wages in both sectors, $\{w_t^c\}_{t=0}^{\infty}$ and $\{w_t^u\}_{t=0}^{\infty}$, interest rates, $\{r_t\}_{t=0}^{\infty}$, prices of intermediate goods, $\{p_t(i)\}_{t=0}^{\infty}$, employment in both sectors, $\{n_t^c\}_{t=0}^{\infty}$ and $\{n_t^u\}_{t=0}^{\infty}$, job rationing probability, $\{p_t\}_{t=0}^{\infty}$, capital employed in both sectors, $\{k_t^c\}_{t=0}^{\infty}$ and $\{k_t^u\}_{t=0}^{\infty}$, intermediate and final goods, $\{y_t(i)\}_{t=0}^{\infty}$ and $\{Y_t\}_{t=0}^{\infty}$, consumption $\{c_t\}_{t=0}^{\infty}$, investment $\{i_t\}_{t=0}^{\infty}$, and capital stock $\{k_t\}_{t=0}^{\infty}$, such that

1. Given wages and interest rates, the representative households maximizes lifetime utilities, (4), subject to (5) – (7), final goods producers maximize profits in (8), and intermediate goods producing firms in both sectors maximize profits and solve (9);

2. The monopoly union maximizes its objective, (9), subject to (10) and (11);

3. Markets Clear[17]

   - Capital Market
$$\int_0^{\phi} k_t(i)di + \int_{\phi}^{1} k_t(i) = k_t, \quad \forall t.$$

   - Goods Market
$$c_t + i_t = Y_t, \quad \forall t$$

3.4.1 A Static Case

In this subsection, we present a static version of the full model and focus on the symmetric equilibrium, which illustrates the main mechanism at work. All notations remain the same as the previous section. The representative household's problem is

$$\begin{aligned} \max_{n^c, n^u} \quad & u(c) - \nu(n) \\ s.t. \quad & c = \phi w^u n^u + (1-\phi) w^c n^c + \phi \Pi^u + (1-\phi)\Pi^c \\ & n = \phi \frac{n^u}{q} + (1-\phi) n^c \end{aligned}$$

Firms in both the competitive and unionized sectors take wage as given and optimally make hiring decisions

$$\Pi^i = \max_{n^i} \; p(i)F(n^i) - w^i n^i$$

Assume the union' objective function is of Cobb-Douglas[18]. The union proposes wages while taking account for the fact that firms in the unionized sector hire workers optimally. Denote $D(w^u)$ firms' demand for unionized workers at the wage rate w^u. The problem of

[17]Note that the labor market clearing condition is implicitly assumed by employing the same notation for labor demand and supply.

[18]Dinlersoz and Greenwood (2016b) share the identical objective function of labor unions.

unions is

$$\max_{w^u} \quad (w^u - w^c)^\eta n^{u1-\eta}$$

$$s.t. \quad n^u = D(w^u, [black, fill = black](0,0)circle(.3ex);)$$

To see the effect of unions on the labor wedge, note that the two marginal conditions, w.r.t. working time in competitive (n^c) and unionized (n^u) sectors, are

$$u' * w^c = \nu'$$

$$u' * w^u = \nu' \frac{1}{q}$$

It is frictionless in the competitive sector. There exists a labor wedge, given by $\frac{1}{q}$, in the unionized sector. Note that the labor wedge is created by the wage premium and its associated labor rationing process in the unionized sector. Denote $\tilde{\phi} \equiv \frac{\phi n^u}{\phi n^u + (1-\phi)n^c}$ the fraction of employees that work in the unionized sector, and $1 - \tilde{\phi}$ in the competitive sector. We have the following relation

$$W = \tilde{\phi} w^u + (1 - \tilde{\phi}) w^c = [\tilde{\phi} \frac{1}{q} + (1 - \tilde{\phi})] * MRS$$

The size of the wedge is determined by the exogenous parameter ϕ, endogenous variables n^u, n^c, and q.

We parameterize the economy by choosing the following functional forms

$$u(c) = \log(c); \quad \nu(n) = \gamma \frac{\epsilon}{1+\epsilon} n^{\frac{1+\epsilon}{\epsilon}}; \quad F(n) = n^\alpha$$

The labor demand function for unionized firms, $D(n^u)$, is

$$D(n^u) = (\frac{w^u}{\alpha p})^{\frac{1}{\alpha - 1}}$$

The union's problem becomes

$$\max_{w^u} \quad (w^u - w^c)^\eta n^{u1-\eta}$$
$$s.t. \quad n^u = (\frac{w^u}{\alpha p})^{\frac{1}{\alpha-1}}$$

Note that the union takes wage in the competitive sector as given. It is straightforward to solve the optimization above and obtain

$$w^u = \frac{1-\eta}{1-\eta+(\alpha-1)\eta} w^c$$

We restrict attention to the case $1-\eta+(\alpha-1)\eta > 0$. Denote $\Delta \equiv \frac{1-\eta}{1-\eta+(\alpha-1)\eta}$. Since $\alpha - 1 < 0$, it follows that $\Delta > 1$, and $w^u > w^c$. Combining this relation with the marginal conditions on the household side leads to

$$\frac{1}{q} = \frac{1-\eta}{1-\eta+(\alpha-1)\eta}$$

The intuition for this equation is, if the wage premium of unions is higher, then the queue is longer outsize of the labor union, which implies a lower matching probability.

The union density, i.e. the fraction of union members in total employment, is

$$\tilde{\phi} = \frac{\phi n^u}{\phi n^u + (1-\phi) n^c} = \frac{\phi}{\phi + (1-\phi)\Delta^{\frac{1}{1-\alpha}}},$$

and the wedge between the average wage, $W \equiv \tilde{\phi} w^u + (1-\tilde{\phi}) w^c$, and the marginal rate of substitution, is $\tilde{\phi}\Delta + (1-\tilde{\phi})$.

Note that the way we calculate the marginal product of labor is

$$\begin{aligned} MPL = \alpha\frac{Y}{N} &\equiv \alpha\frac{\phi p^u y^u + (1-\phi)p^c y^c}{\phi n^u + (1-\phi)n^c} \\ &= \tilde{\phi}\alpha\frac{p^u y^u}{n^u} + (1-\tilde{\phi})\alpha\frac{p^c y^c}{n^c} \\ &= \tilde{\phi}w^u + (1-\tilde{)}\phi w^c \equiv W \end{aligned}$$

The size of the labor wedge, defined as $\log\frac{\text{MP}}{\text{MRS}}$ is therefore proportional to $\tilde{\phi}\Delta + (1-\tilde{\phi})$.

3.5 Quantitative Analysis

This section implements quantitative analyses. We choose the following functional forms

$$u(c) = \log(c); \quad \nu(n) = \gamma\frac{\epsilon}{1+\epsilon}n^{\frac{1+\epsilon}{\epsilon}}; \quad F^j(k,n) = A^j(k^{1-\alpha}n^{\alpha})^{\chi}, j = u, c.$$

In principle, productivities in the unionized and competitive sectors can be different from each other[19].

The objective function of the union is chosen to be[20]

$$G(w_t^u - w_t^n, n_t(i)) = (w_t^{u\theta_1} - w_t^{c\theta_2})^{\eta}(\int_0^{\phi} n_t(i)di)^{1-\eta}$$

The union values the wage premium, $w_t^u - w^{ct}$, but might have different elasticities towards increases in w_t^u and decreases in w_t^c.

[19] Dinlersoz and Greenwood (2016b) documents that unions generally target more productive firms.

[20] A more general utility function is used since the wage premium is constant under a Cobb-Douglas objective function.

Several parameters are standard especially in the business cycle literature, and we choose widely used values for these parameters. One period in the model corresponds to one year in the data. We choose the annual depreciation rate δ to be 10%. The value of the discount rate is set as $\beta = 0.96$ such that the annual net return of investment equals 4%. In the baseline calibration, we set the Frisch elasticity of labor supply as $\epsilon = 1$. The value of γ, the weight on working disutility in household's preference, is chosen such that the household spend one third of hours on working[21]. On the production side, we follow Buera et al. (2011a), and choose the span of control parameter in production function $\chi = 0.79$[22]. The value of α is calibrated to be 0.81 to match a labor share of 64%, as in Kydland and Prescott (1982).

Without loss of generality, we normalize productivity in the competitive sector, $A^i, \phi < i < 1$ (or simply A^c), to be 1. Due to the existence of decreasing return to scale, a small difference in wages translates into a relatively large gap in employment. That is, employment in firms in the the competitive sector is significantly higher than unionized employment even if the union wage premium is moderate[23]. Table 3.6 lists the distribution of firm size groups among union members and regular workers. The median union worker works in a firms with more than 1000 persons in both 1992 and 2007, while the total number of persons in the firm the median non-union worker works in is between 100 and 500.

[21] This is for the year 1970. Total working hours are affected by the size of the unionized sector, as measured by ϕ. The value of ϕ varies over years, so as the working hours.

[22] There are different calibrations for this span of control parameter. For example, the parameter is set as 0.85 in Midrigan and Xu (2014b).

[23] See Appendix for a detailed derivation. It can be seen there that the employment ratio of competitive to unionized firms is a function of $\frac{A^u}{A^c}$ and the wage ratio $\frac{w^u}{w^c}$.

Table 3.6: Distribution of workers among firm size groups

	< 10	[10,50)	[50,100)	[100,500)	[500,1000)	≥ 1000
1992						
Non-Union	15.28%	11.16	16.50	16.44	5.77	34.85
Union Member	3.47	3.94	9.73	17.21	7.86	57.78
Union Coverage	3.99	4.23	10.03	17.56	8.15	56.04
2007						
Non-Union	15.89%	12.25	15.15	14.11	5.98	36.62
Union Member	4.92	5.69	8.73	16.11	8.11	56.44
Union Coverage	5.00	6.12	9.53	16.00	7.98	55.36

Data Source: CPS. Firm size indicates the total number of persons who work in the firm. The universe is workers who work for wage and salaries in the private sector.

We increase the productivity of the unionized firms, $A^i, 0 < i < \phi$ (or A^u), to be consistent with the fact that union members, in average, command a wage premium and as well work in larger firms than non-union workers. In the baseline calibration, we set $A^u = 1.23$ such that, under a 20% of union wage premium, the average employment size of unionized firms is 20% higher than that in the competitive sector. A 26% of union density, i.e. fraction of union members in total employment, requires 22.6% of firms to be unionized[24].

The parameter ζ determines the elasticity of substitution among firms within each sector. This parameter is widely used in the New Keynesian business cycle literature. We pick the value of $\zeta = 0.83$, the benchmark value used in Christiano et al. (2005). For the parameter ρ, which governs the elasticity of substitution between sectoral goods, note that the following

[24] Results under different values of A^u are presented in appendix.

relation holds under the aggregate production function, $Y_t = (y_t^{u\rho} + y_t^{c\rho})^{\frac{1}{\rho}}$,

$$\log \frac{p_t^u y_t^u}{p_t^c y_t^c} = \frac{\rho}{\rho - 1} \log \frac{p_t^u}{p_t^c}.$$

The relation between relative expenditure and relative price between the competitive and unionized sectors from data can therefore be used to estimate the parameter ρ[25]. Table C.1 in appendix lists the average union density, ranked from low to high, for 2 digit sectors in NAICS. Sectors that have a union density higher than 15% are labeled as unionized sectors, and the rest competitive sectors[26]. We then run a single variable regression with relative expenditure share and relative price[27] of unionized sectors, comparing to competitive sectors, as dependent and independent variables, respectively. Each year is one observation, and there are 34 years (1983-2017). The regression results yields a value of $\rho = 0.5$ for all economy[28].

We normalize $\theta_2 = 1$. The Union's optimization problem implies a relation between wages in competitive and unionized sectors, $\varphi(w^u, w^c, \theta, [black, fill = black](0,0)circle(.3ex);)$. We use this relation from data to estimate θ_1, and obtain the value $\theta_1 = 1.26$. We jointly calibrate parameters γ, ϕ, and η, to target a working hour of about 1/3, a wage premium of 20%[29], and a union density of 21%. All moments refer to their 1977 values. Table 3.7 lists the results.

[25]Similar methods has been employed in Cole and Ohanian (2004), and Acemoglu (2003), to estimate the parameter governing the elasticity of substitution between sectors.
[26]Defined this way, the value added share of unionized sectors is 43.1%, and 33.6% in the private economy in 1987.
[27]we calculate price index for each sector as the weighted average of prices of industries within that sector, with value added used as weights
[28]As robustness check, we have tried to use $\rho = 0.1$ and found robust the main results regarding changes of the labor wedge
[29]See Table C.2 in appendix.

Table 3.7: Calibration moments

Moments	Model	Data
working hour	33.63%	1/3
wage premium	20.00%	20%
union density	21.00%	21%

Table 3.8 summarizes the calibration results.

Table 3.8: Summary of Calibration

Para.	Description	Value	Target/Source
β	Discount rate	0.96	Annual 4% net return of inv.
δ	Depreciation rate of K	0.1	Annual 10% depreciation
γ	Disutility of work	4.04	$\frac{1}{3}$ working time
ϵ	Elasticity of labor supply	1	Standard
A^c	Productivity in competitive sector	1	Normalization
A^u	Productivity in unionized sector	1.23	Relative employment size
χ	Span-of-control in prod. fun.	0.79	Buera et al. (2011)
α	Labor's share	0.81	Labor share in NIPA
ζ	Elasticity of sub. within sectors	0.83	CEE(2005)
ρ	Elasticity of sub. btw. sectors	0.5	Expenditure-price elasticity
θ_1	Union's elasticity w.r.t. w^c	1	Normalization
θ_2	Union's elasticity w.r.t. w^u	1.26	See text
η	Union's weight on wages	0.38	Union wage premium
ϕ	Fraction of unionized firms	0.18	Union density in 1977

Note: see text for details.

To see the implications on the labor wedge, note that, the labor wedge is originally from the gap between average wage and the marginal rate of substitution, due to the wage premium commanded by the union and the associated job rationing. The fraction of employment in the unionized sector is measured as $\tilde{\phi} \equiv \frac{\phi n^u}{\phi n^u + (1-\phi) n^c}$, and the economy wide wage is given by

$$W = \tilde{\phi} w^u + (1-\tilde{\phi}) w^c$$

Denote $MPL \equiv \alpha\chi\frac{Y}{N}$ the economy wide marginal productivity of labor. As in the static case, it follows that

$$MPL = W = lw * MRS$$

where

$$lw = \tilde{\phi}\frac{1}{q} + (1-\tilde{\phi})$$

q is the probability of a successful job rationing in the unionized sector, which is determined endogenously. We treat each year as a steady state, and vary the value of ϕ across years, by targeting union densities in data. Table 3.9 lists results for several selected years.

Table 3.9: Model moments

Period	ϕ	w^u	w^c	n^u	n^c	Union D.	LW (in log)
1970-79	0.21	1.05	0.86	0.38	0.32	23.95%	5.13%
1980-89	0.16	1.04	0.85	0.40	0.33	18.22%	4.01%
1990-99	0.13	1.03	0.84	0.41	0.35	14.49%	3.23%
2000-07	0.11	1.02	0.83	0.43	0.37	12.26%	2.76%

Based on previous calculations, the labor wedge in data has declined 0.17 log points from 1970s to 2000s. For the same period, our model, by matching the magnitude of union density, implies a decrease of labor wedge by 0.024 log points, which accounts for 14% of the overall decline in data.

Union premiums means a higher wage in the model. In reality, however, a large proportion of union premium does not come in the form of wages, but in things such as a flexible working hours, and larger retirement benefits. To capture these non-wage benefits, we have tried an alternative calibration to target a wage premium of 30%, instead of 20%. Table 3.10 lists main results under these alternative calibration.

Table 3.10: Model moments

Period	ϕ	Union D.	LW (in log)
1970-79	0.21	23.95%	7.99%
1980-89	0.16	18.22%	6.22%
1990-99	0.13	14.49%	3.23%
2000-07	0.11	12.26%	4.26%

Under this alternative calibration, the model generates 3.73% log points of decrease in the labor wedge, which accounts for 22% of the decline observed in data.

3.6 Conclusion

Labor wedge, the difference between marginal product of labor and marginal rate of substitution, is a reduced form representation of deviations from competitive market and allocation efficiency. The measured labor wedge in U.S. since world war II has been relatively stable until early to middle 1970s, and steadily declined for the recent 3-4 decades. In this paper, we propose that the existence of labor unions, and their power to influence wages and employment in the labor market affect the behavior of labor wedge. Overall, union density follows a similar stable-then-decline trend. We further assemble a panel dataset of 75 manufacturing sectors from 2005 to 2014, and provide empirical support for the union-wedge connection.

We developed a dynamic general equilibrium model to quantify the effect of union power on labor wedge. In our model economy, there are two sectors, competitive and unionized. Wages in the competitive market are determined competitively, while the unionized wages are controlled by a monopolistic union. The union commands a wage premium, which invites job application queues and job rationing in the unionized sector. This job rationing process creates a wedge between wages and households' willing to work. According to our results, approximately 20% of the decline in labor wedge from 1970s to 2000s can be accounted for by the decreases in union density.

The long run trend of labor wedge in general, and its decline since 1970s in particular provide a summary of market efficiency and its overall change. The decline of labor wedge, or the improvement of market efficiency, might also be a result of changes beyond declining power of labor unions. One example is the decrease of tax rate, which also create a wedge between marginal conditions. We leave explorations along these directions for future research.

Bibliography

ABOW, J. M., F. KRAMARZ, AND D. N. MARGOLIS (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67, 251–333.

ACEMOGLU, D. (2003): "Labor and Capital Augmenting Technical Change," *Journal of the European Economic Association*, 1, 1–37.

ALLEN, R. C., J.-P. BASSINO, D. MA, C. MOLL-MURATA, AND J. L. V. ZANDEN (2011): "Wages, Prices, and Living Standards in China, 1738-1925: in Comparison with Europe, Japan, and India," *The Economic History Review*, 46, 8–38.

ANDERSON, R., N. JOHNSON, AND M. KOYAMA (2013): "Jewish Persecutions and Weather Shocks: 1100-1800," *Working paper, University of Michigan.*

ANTRAS, P. (2004): "Is the U.S. Aggregate Production Function Cobb-Douglass? New Estimates of the Elasticity of Substitution." *Contributions in Macroeconomics*, 4.

ATKESON, A. AND P. J. KEHOE (2007): "Modelling the Transition to a New Economy: Lessons from Two Technological Revolution," *The American Economic Review*, 97, 64–88.

AUTOR, D., D. DORN, L. F. KATZS, C. PATTERSON, AND J. V. REENEN (2017): "The Fall of the Labor Share and the rise of Superstar Firms," *Working Paper.*

AWUKU-BUDU, C. AND C. A. ROBBINS (2014): "The Role of Industry Classification in the

Estimation of Research and Development Expenditures," *Discussion Papers, U.S. Census Bureau, Center for Economic Studies.*

Axtell, R. L. (2011): "Zipf Distribution of U.S. Firm Size," *Science*, 293, 1818–1820.

Bai, Y. and J. K. sing Kung (2011): "Climate shocks and Sino-nomadic conflict," *The Review of Economics and Statistics*, 93, 970–981.

Barkai, S. (2016): "Declining Labor and Capital Shares," *Working paper.*

Barro, R. J. and X. Sala-iMartin (2004): *Economic Growth*, Cambridge, Massachusetts: The MIT Press, 2 ed.

Baten, J., W. Bierman, J. L. van Zanden, and P. Foldvari (????): .

Becker, G. (1968): "Crime and Punishment: An Economic Approach," *The Journal of Political Economy*, 76, 169–217.

Bernhofena, D., M. Eberhardt, J. Li, and S. Morgan (2015): "Assessing market (dis)integration in pre-modern China," *Working paper, American University.*

Brown, M. and J. S. DeCani (1963): "Technological Change and the Distribution of Income," *International Economic Review*, 4, 289–309.

Buera, F., J. Kaboski, and Y. Shin (2011a): "Finance and Development: A Tale of Two Sectors," *American Economic Review*, 101, 1964–2002.

Buera, F. J., J. P. Kaboski, and Y. Shin (2011b): "Finance and Development: A Tale of Two Sectors," *The American Economic Review*, 101, 1964–2002.

Buoye, T. M. (2000): *Manslaughter, Markets, and Moral Economy: Violent Disputes over Property Rights in 18th-Century China*, London: Cambridge University Press.

Burgess, R. and D. Donaldson (2010): "Can Openness Mitigate the Effects of Weather Shocks? Evidence from India's Famine Era," *American Economic Review*, 100, 449–453.

Cao, S. (2001): *Zhongguo Renkou Shi (Qing Shiqi) [The Population History of China: Qing]*, Shanghai: Fudan University Press.

Cao, Y. and S. Chen (2016): "Robin Hood on the Grand Canal: Economic Shock and Rebellions in Qing China, 1650-1911," *Working paper.*

Chari, V., P. Kehoe, and E. Magrattan (2007): "Business Cycle Accounting," *Econometrica*, 75, 781–836.

Chen, Z., K. Peng, and L. Zhu (2017): "Social-Economic Change and its Impact on Violence: Homicide History of Qing China," *Explorations in Economic History*, 63, 8–25.

Cheremukhin, A. and P. Restrepo-Echavarria (2014): "The Labor Wedge as a Matching Function," *European Economic Review*, 68, 71–92.

Cheung, S.-W. (2008): *The Price of Rice: Market Integration in 18th-Century China*, Bellingham: Western Washington University Press.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113, 1–45.

Cociuba, S., E. Prescott, and A. Ueberfeldt (2012): "U.S. Hours and Productivity Behavior: Using CPS Hours Worked Data: 1947-III to 2011-IV," *Working paper.*

Cole, H. and L. Ohanian (2004): "New Deal Policies and the Persistence of the Great Depression," *Journal of Political Economy*, 112, 779–816.

Congress, U. (1995): "The Technological Reshaping of Metropolitan America," *U.S. Congress, Office of Technology Assessment*.

David, P. A. and T. van de Klundert (1965): "Biased Efficiency Growth and Capital-Labor Substitution in the U.S., 1899-1960," *The American Economic Review*, 55, 357–394.

Dinlersoz, E. and J. Greenwood (2016a): "The Rise and Fall of Unions in the United States," *Journal of Monetary Economics*, 83, 129–146.

——— (2016b): "The Rise and Fall of Unions in the U.S.: A Macroeconomic Perspective," *Journal of Monetary Economics*, 83, 129–146.

Edward Miguel, S. S. and E. Sergenti (2004): "Economic Shocks and Civil Conflict: An Instrumental Variables Approach," *Journal of Political Economy*, 112, 725–753.

Eisner, M. (2003): "Long-Term Historical Trends in Violent Crime," *Crime and Justice*, 83–142.

Elias, N. (2000): *The Civilizing Process: Sociogenetic and Psychogenetic Investigations*, Oxford: Blackwell Publishing; Revised edition.

Elsby, M. W. L., B. Hobijn, and A. Sahin (2013): "The Decline of the U.S. Labor Share," *Brookings Papers on Economic Activity*, 1–52.

Feenstra, R. C. and G. Hansen (1996): "Globalization, Outsourcing, and Wage Inequality," *The American Economic Review Papers and Proceedings*, 86, 240–245.

——— (1998): "The Impact of Outsourcing, and High-Technology Capital on Wages: Estimates for the United States, 1979-1990," *The Quarterly Journal of Economics*, 114, 907–940.

GALI, J., M. GERTLER, AND D. LOPES-SALIDO (2007): "Markups, Gaps, and the Welfare Costs of Business Fluctuations," *The Review of Economics and Statistics*, 89, 44–59.

GAO, X. (1993): "Kangqian Shengshi Qianyi [A Discussion of Kangxi and Qianlong Prosperity]," *Qingshi Yanjiu [The Qing History Journal]*, 1.

GOMME, P., B. RAVIKUMAR, AND P. RUPERT (2011): "The Return to Capital and the Business Cycle," *Review of Economic Dynamics*, 14, 262–278.

GROSSMAN, H. (1991): "A General Equilibrium Model of Insurrections," *The American Economic Review*, 81, 912–921.

GURR, T. R. (1981): "Historical Trends in Violent Crime: A Critical Review of the Evidence," *Crime and Justice*, 3, 295–353.

HERRENDORF, B., C. HERRINGTON, AND A. VALENTINYI (2015): "Sectoral Technology and Structural Transformation," *American Economic Journal: Macroeconomics*, 7, 104–133.

JIA, R. (2013): "Weather Shocks, Sweet Potatoes and Peasant Revolts in Historical China," *The Economic Journal*, 124, 92–118.

JONES, C. I. (2005): "The Shape of Production Function and the Direction of Technical Change," *The Quarterly Journal of Economics*, 120, 517–549.

KALDOR, N. (1957): "A Model of Economic Growth," *The Economic Journal*, 67, 591–624.

KARABARBOUNIS, L. AND B. NEIMAN (2014a): "Capital Depreciation and Labor Shares Around the World: Measurement and Implications," *Working paper.*

——— (2014b): "The Global Decline of the Labor Share," *The Quarterly Journal of Economics*, 61–103.

KEHRIG, M. AND N. VINCENT (2017): "Growing Productivity Without Growing Wages: The Micro-Level Anatomy of the Aggregate Labor Share Decline," *Economic Research Initiatives at Duke (ERID) Working Paper No. 244.*

KLUMP, R., P. MCADAM, AND A. WILLMAN (2007): "Factor Substitution and Factor-Augmenting Technical Progress in the United State: A Normalized Supply-Side System Approach," *Review of Economics and Statistics*, 89, 183–92.

KOH, D., R. SANTAEULALIA-LIOPIS, AND Y. ZHENG (2016): "Labor Share Decline and Intellectual Property Products Capital," *Working paper.*

KRUSELL, P., L. OHANIAN, V. RIOS-RULL, AND V. GIOVANNI (2000): "Capital Skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica*, 68, 1029–1053.

KYDLAND, F. AND E. PRESCOTT (1982): "Time to Build and Aggregate Fluctuations," *Econometrica*, 50, 1345–1370.

LAWRENCE, R. Z. (2015): "Recent Declines in Labor's Share in US Income: A Preliminary Neoclassical Account," *Working paper.*

LI, S. (1999): "Guanyu Kangqian Shengshi de Lishi Fenqi Wenti [Periodization of Kangxi and Qianlong Prosperity]," *Zhongguo Shehui Kexueyuan Yanjiushengyuan Xueban [Journal of Graduate School of Chinese Academy of Social Sciences].*

LIU, Z. (1994): "Qin Shigao Dili Zhi Fu Zhou Xian Ting Zhiguan Quefen Fanjian Dingwu [Correction of the Official Ratings of Prefectures and Counties in the Geography in Draft History of Qing Dynasty]," *Zhongyang Yanjiuyuan Lishi Yuyan Yanjiusuo Jikan [Collected Papers of Institute of History and Philology of Academia Sinica].*

Loecker, J. D. and E. Jan (2017): "The Rise of Market Power and the Macroeconomic implications," *Working paper.*

Lucas, R. E. (1978): "On the Size Distribution of Business Firms," *The Bell Journal of Economics*, 9, 508–523.

Maddison, A. (2007): *Chinese Economic Performance in the Long Run: 960-2030 A.D.*, Paris: Development Centre of OECD.

Midrigan, V. and D. Y. Xu (2014a): "Finance and Misallocation: Evidence from Plant-Level Data," *The American Economic Review*, 104, 422–458.

——— (2014b): "Finance and Misallocation: Evidence from Plant-Level Data," *American Economic Review*, 104, 422–458.

Miller, H. (2013): *State versus Gentry in Early Qing Dynasty China, 1644-1699*, Basingstoke, U.K.: Palgrave Macmillan.

Moylan, C. E. (2008): "Employee Stock Options and the National Economic Accounts," *Bureau of Economic Analysis.*

North, D. C., J. J. Wallis, and B. R. Weingast (2009): *Violence and Social Order*, Cambridge, U.K.: Cambridge University Press.

Oberfield, E. and D. Raval (2014): "Micro Data and Macro Technology," *Working paper.*

Parker, R. P. (2012): "Restoring the Enterprise Statistics Program (ESP) for the 2012 Economic Census," *Reports of the Census Bureau.*

Pescatori, A. and M. Tasci (2011): "Search Frictions and the Labor Wedge," *Working paper.*

PIKETTY, T. AND G. ZUCMAN (2014): "Capital is Back: Wealth-Income Ratios in Rich Countries 1700-2010," *The Quarterly Journal of Economics*, 1255–1310.

PINKER, S. (2011): *The Better Angels of Our Nature: Why Violence Has Decline*, New York: The Penguin Group.

RUDANKO, L. AND P. KRUSELL (2015): "Unions in a Frictional Labor Market," *Journal of Monetary Economics*, 80, 35–50.

SEN, A. (1982): *Poverty and Famines: An Essay on Entitlement and Deprivation*, Oxford New York: Clarendon Press Oxford University Press.

SHIMER, R. (2009): "Convergence in Macroeconomics: The Labor Wedge," *American Economic Journal: Macroeconomics*, 1, 280–297.

SNG, T. H. (2014): "Size and Dynastic Decline: The Principal-Agent Problem in Late Imperial China 1700-1850," *Explorations in Economic History*, 54, 107–127.

SOLOW, R. (1958): "A Skeptical Note on the Constancy of Relative Shares," *The American Economic Review*, 48, 618–631.

TASCHEREAU-DUMOUCHEL, M. (2015): "The Union Threat," *Working paper*.

WAKEMAN, F. J. (1998): "Boundaries of the Public Sphere in Ming and Qing China," *Explorations in Economic History*, 127, 167–189.

WESTERN, B. AND J. ROSENFELD (2011): "Unions, Norms, and the Rise in U.S. Wage Inequality," *American Sociological Review*, 76, 513–537.

WILKINSON, M. (1968): "Factor Supply and the Direction of Technological Change," *The American Economic Review*, 58, 120–128.

Yan, S. and C. Liu (2011): "18 Shiji Zhongguo Nan Bei Fang Shichang Zhenghe Chengdu de Bijiao: Liyong Qingdai Liangjia Shuju de Yanjiu. [Market Integration in Sourthern and Northern China in the 18th Century: A Study Based on Grain Price Data in Qing Dynasty]," *Jingji Yanjiu [Economic Research Journal]*, 12, 124–137.

Zhang, Z. (1991): *Zhongguo Shenshi: Guanyu qizai Shijiu Shiji Zhongguo Shehui zhong Zuoyong de Yanjiu [Local Gentlemen: its social role in 19th century's China]*, Shanghai: Social Science Press.

Appendix A

Supplementary Material to Chapter 1

K/L *intensity under different technologies in Static Case.* The equilibrium condition in the labor market is

$$w = \text{MPL}_j = \left[(1-\alpha)\left(\frac{k_j}{\ell_j}\right)^{\rho} + \alpha\gamma_j^{\rho}\right]^{1/\rho-1} \alpha\gamma_j^{\rho}.$$

Note that this expression contains two γ_j. The last γ_j^{ρ} captures a direct effect: technology with a higher γ_j increases labor productivity and requires less labor to produce. The first γ_j^{ρ} captures the opposite effect—namely, that this technology allows firms to operate at a larger scale, which increases labor demand.

Rewrite the labor market marginal condition as

$$(1-\alpha)\left(\frac{k}{\ell}\right)^{\rho} = \left(\frac{w}{\alpha\gamma^{\rho}}\right)^{\rho/(1-\rho)} - \alpha\gamma^{\rho} \equiv f(\gamma).$$

A sufficient condition for a positive correlation between γ and k/ℓ is $f'(\gamma) < 0$. Note that

$$f'(\gamma) = \left(\frac{w}{\alpha}\right)^{\rho/(1-\rho)} \left(-\frac{\rho^2}{1-\rho}\right)\gamma^{-\rho^2/(1-\rho)-1} - \alpha\rho\gamma^{\rho-1}$$

and that $f'(\gamma) < 0$ is equivalent to

$$\gamma > \left(\frac{w}{\alpha}\right)^{1/(1-\rho)} \left(\frac{-\rho}{1-\rho}\frac{1}{\alpha}\right)^{1/\rho}.$$

Therefore, this condition is satisfied if γ is large enough[1].

Firm size and employment in general equilibrium Let MPK stand for the marginal productivity of capital. Then the optimal conditions in a competitive equilibrium are as

[1]Note that it follow from the marginal condition of labor market, $(1-\alpha)(\frac{k}{\ell})^{\rho} = (\frac{w}{\alpha\gamma^{\rho}})^{\frac{\rho}{1-\rho}} - \alpha\gamma^{\rho}$, that $(\frac{w}{\alpha\gamma^{\rho}})^{\frac{\rho}{1-\rho}} - \alpha\gamma^{\rho} > 0$, or equivalently, $\gamma > \frac{w}{\alpha}(\frac{1}{\alpha})^{\frac{1-\rho}{\rho}}$.

follows:

$$\mathrm{MPK}_i = \eta z_i^{1-\eta}\left[(1-\alpha)\left(\frac{k_i}{q_\gamma}\right)^{\rho} + \alpha(\gamma\ell_i)^{\rho}\right]^{\eta/\rho-1}(1-\alpha)q_\gamma^{-\rho}k_i^{\rho-1} = r;$$

$$\mathrm{MPL}_i = \eta z_i^{1-\eta}\left[(1-\alpha)\left(\frac{k_i}{q_\gamma}\right)^{\rho} + \alpha(\gamma\ell_i)^{\rho}\right]^{\eta/\rho-1}\alpha\gamma^{\rho}\ell_i^{\rho-1} = w.$$

Firm i's demand for capital and labor may be written as

$$k_i = z_i\left(\frac{\eta(1-\alpha)}{r}\right)^{1/(1-\eta)} q_\gamma^{-\eta/(1-\eta)}\left[(1-\alpha)+\alpha\left(\frac{r}{w}\frac{\alpha\gamma q_\gamma}{1-\alpha}\right)^{\rho/(1-\rho)}\right]^{(\eta-\rho)/\rho(1-\eta)} \quad \text{and}$$

$$\ell_i = z_i\left(\frac{\eta\alpha}{w}\right)^{1/(1-\eta)} \gamma^{\eta/(1-\eta)}\left[(1-\alpha)\left(\frac{1-\alpha}{\alpha\gamma q_\gamma}\frac{w}{r}\right)^{\rho/(1-\rho)}+\alpha\right]^{(\eta-\rho)/\rho(1-\eta)},$$

respectively. Therefore, the net effect of technology on employment size is unclear. The reason is that technology has two effects on employment, as illustrated in the static case above.

Data Source

- Data used to calculate labor share for 2-digit sectors are from the industry accounts of the Bureau of Economic Analysis (BEA). Labor share for detailed manufacturing sectors is calculated using data from the Annual Survey of Manufactures. A simplified version of the ASM by 6-digit sectors is summarized in the NBER-CES data set. These data are available for 6-digit manufacturing sectors from 1958 to 2011.

- Concentration ratios for 2002, 2007, and 2012 are from the US Census Bureau's American FactFinder website and from various pre-1997 publications of the Census Bureau. This measure gives shares of the 4, 8, 20, and 50 largest firms in total sales (receipts, or value of shipments). For manufacturing, concentration ratios in terms of value added are also available.

Classification of multi-establishment enterprises The NAICS is designed to facilitate the collection, tabulation, presentation, and analysis of data relating to establishments (see Parker (2012)). For industry classification of multi-unit firms[2] with diverse production activities, the NAICS uses a multiple stage "hierarchical" approach. In the first stage, the firm is assigned to a major sector based on highest share of payroll. Within that major sector, the firm is then assigned to a subsector, based on the highest share across the subsectors within the major sector (cf. Awuku-Budu and Robbins (2014)). This process continues through to the most disaggregated level of industry classifications.

Labor Share in Census of Manufacturing versus NIPA

[2] "unit and "establishment" are used interchangeably. In this context,"firms", "enterprises", and "companies" amounts to the same things.

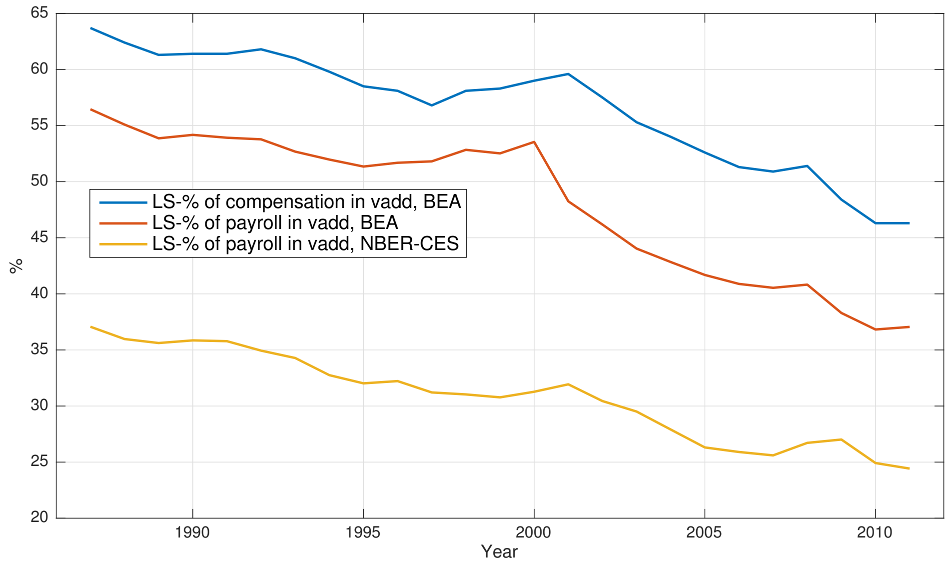


Figure A.1: MFG LS in ASM/NBER-CES and BEA-NIPA

Offshoring index

Following Feenstra and Hansen (1996) and Feenstra and Hansen (1998), I measure sector i's offshoring intensity as

$$\mathrm{OS}_i = \Sigma_j \frac{\text{Input from sector } j}{\text{Total intermediate input in } i} \times \text{Import intensity of } j.$$

The import intensity of j is measured as the fraction of imports in expenditures, which equals the value of shipments (plus imports and minus exports) in sector j. Offshoring intensity is calculated using the input–output tables of 1997 and 2015 for 66 private industries. Figure A presents the results. The measured offshoring intensity increases in most industries, but the increases are more prominent in manufacturing than in non-manufacturing sectors.

`offshoring-eps-converted-to.pdf`

Figure A.2: Offshoring intensity, 1997-2015

Table A.1: Concentration and Labor Share. Depend. var.: ΔLS, 1997-2007

	OLS				WLS			
	3 digit	4 digit	5 digit	6 digit	3 digit	4 digit	5 digit	6 digit
Δ*CR04*	-0.239**	-0.206***	-0.237***	-0.171***	-0.117	-0.175***	-0.191***	-0.156***
	(0.099)	(0.054)	(0.033)	(0.024)	(0.084)	(0.043)	(0.032)	(0.024)
R^2	0.23	0.16	0.22	0.10	0.09	0.16	0.16	0.08
Δ*CR08*	-0.213***	-0.126**	-0.175***	-0.146***	-0.125**	-0.132***	-0.147***	-0.125***
	(0.073)	(0.062)	(0.041)	(0.029)	(0.057)	(0.046)	(0.037)	(0.029)
R^2	0.31	0.05	0.09	0.05	0.20	0.09	0.08	0.04
Δ*CR20*	-0.318***	-0.149*	-0.210***	-0.157***	-0.132	-0.203***	-0.247***	-0.169***
	(0.115)	(0.077)	(0.049)	(0.036)	(0.084)	(0.063)	(0.053)	(0.040)
R^2	0.29	0.04	0.09	0.04	0.11	0.11	0.11	0.04
Δ*CR50*	-0.347*	-0.137	-0.211***	-0.124***	-0.168	-0.226***	-0.260***	-0.176***
	(0.167)	(0.100)	(0.062)	(0.048)	(0.113)	(0.084)	(0.067)	(0.054)
R^2	0.19	0.02	0.06	0.01	0.11	0.08	0.08	0.02
Obs.	21	86	183	464	21	86	183	464

Note: The single variable regression results are for manufacturing sectors, at various digit levels. The dependent variable is change in labor share and the independent variable is change in concentration from 1997 to 2007. Δ*CR04* refers to the change in *Share04*, which itself measure the share of value added accounted for by the largest 4 firms in a sector. In WLS regressions, the weight used is given by the average value added between 1997 and 2007.

Table A.2: Regression with panel data; Dependent Variable.: ΔLS

	(1)	(2)	(3)
Δ*CR04*	-0.183***	-0.283***	-0.084***
	(0.029)	(0.048)	(0.024)
Δ *k-l ratio (ln)*			-0.322
			(0.760)
Δ *per-capita vadd (ln)*			-30.249***
			(0.812)
Δ *% production worker*			0.656***
			(0.081)
Fixed effect	No	Yes	Yes
Obs.	927	927	927

Note: The panel regression results are for manufacturing sectors at the 6 digit level. There are two period, change from 1997 to 2002, and change from 2002 to 2007. LS is the ratio of payroll to value added.

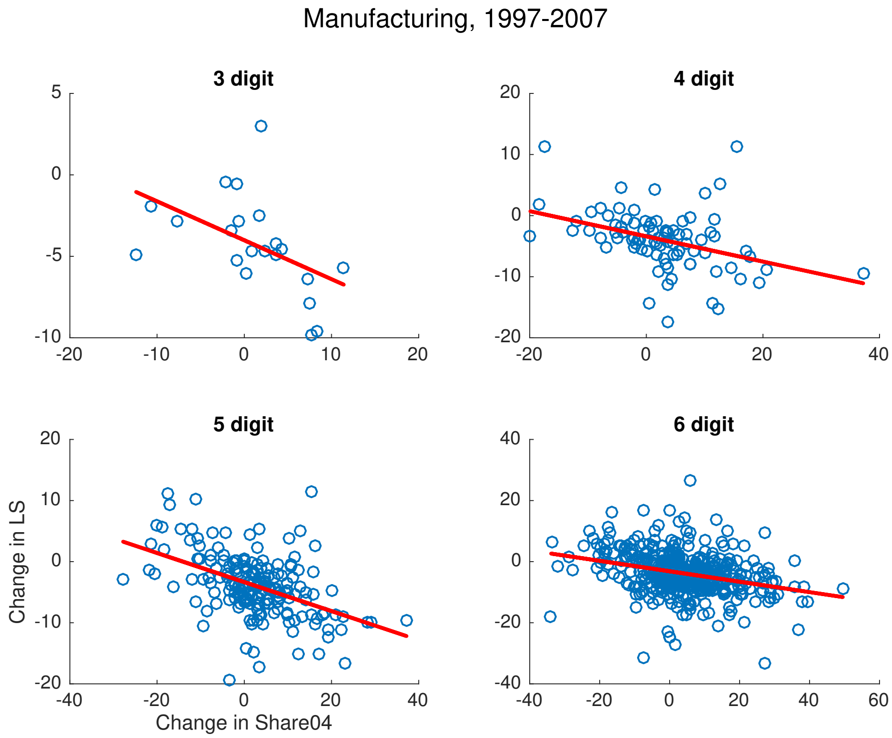


Figure A.3: ΔLS v.s. $\Delta Share04$, MFG

Note: Labor share on the vertical axis is calculated as the fraction of payroll in value added. Concentration on the horizontal axis is the value added share of 4 largest firms.

Table A.3: Dependent Variable:ΔLS

	1963-1967		1972-1977		1977-1982		1987-1992	
	OLS	WLS	OLS	WLS	OLS	WLS	OLS	WLS
$\Delta CR04$	-0.101**	-0.059**	-0.118***	-0.023	-0.190***	-0.240***	-0.174***	-0.128***
	(0.039)	(0.030)	(0.040)	(0.044)	(0.057)	(0.062)	(0.036)	(0.032)
R^2	0.02	0.01	0.02	0.001	0.03	0.03	0.05	0.03
$\Delta CR08$	-0.117***	-0.079***	-0.108***	-0.011	-0.164**	-0.147**	-0.195***	-0.121***
	(0.035)	(0.030)	(0.040)	(0.036)	(0.058)	(0.064)	(0.040)	(0.036)
R^2	0.02	0.02	0.02	0.03	0.02	0.001	0.05	0.03
$\Delta CR20$	-0.187***	-0.103***	-0.113**	-0.067	-0.105	-0.180**	-0.207***	-0.171***
	(0.046)	(0.036)	(0.045)	(0.053)	(0.067)	(0.073)	(0.049)	(0.044)
R^2	0.04	0.02	0.01	0.004	0.006	0.01	0.04	0.03
$\Delta CR50$	-0.167***	-0.076	-0.101*	-0.075	-0.134*	-0.274***	-0.229***	-0.141**
	(0.059)	(0.048)	(0.053)	(0.057)	(0.080)	(0.068)	(0.069)	(0.059)
R^2	0.02	0.007	0.001	0.004	0.007	0.03	0.02	0.01
Obs.	400	400	442	442	437	437	448	448

Note: The single variable regression results are based on 4 digit SIC manufacturing sectors. The dependent and independent variables are the change in labor share and concentration, from 1963 to 1967 in the first two columns, and from 1972 to 1977 in the last two columns. $\Delta CR04$ refers to the change in *Share04*, which itself measure the share of value of shipment by the 4 largest firms in a sector. The weights used in WLS regressions are the average value added of beginning and end years for each period.

Table A.4: Dependent Variable.: ΔLS, 2007-2012

	LS1-OLS	LS2-OLS	LS1-WLS	LS2-WLS
$\Delta CR04$	-0.292**	-0.386***	-0.267***	-0.348***
	(0.041)	(0.051)	(0.032)	(0.043)
R^2	0.23	0.25	0.29	0.28
$\Delta CR08$	-0.352***	-0.475***	-0.363***	-0.476***
	(0.049)	(0.061)	(0.038)	(0.051)
R^2	0.23	0.26	0.34	0.33
$\Delta CR20$	-0.400***	-0.558***	-0.442***	-0.581***
	(0.065)	(0.081)	(0.050)	(0.067)
R^2	0.18	0.22	0.31	0.30
$\Delta CR50$	-0.357***	-0.558***	-0.442***	-0.714***
	(0.092)	(0.115)	(0.050)	(0.096)
R^2	0.08	0.12	0.31	0.24
Obs.	175	175	175	175

Note: The single variable regression results are for manufacturing sectors at the 5 digit level. LS1 is the ratio of payroll to value added, and LS2 the ratio of compensation (=payroll+benefit) to value added. The average value added between 2007 and 2012 is used as the weight in WLS.

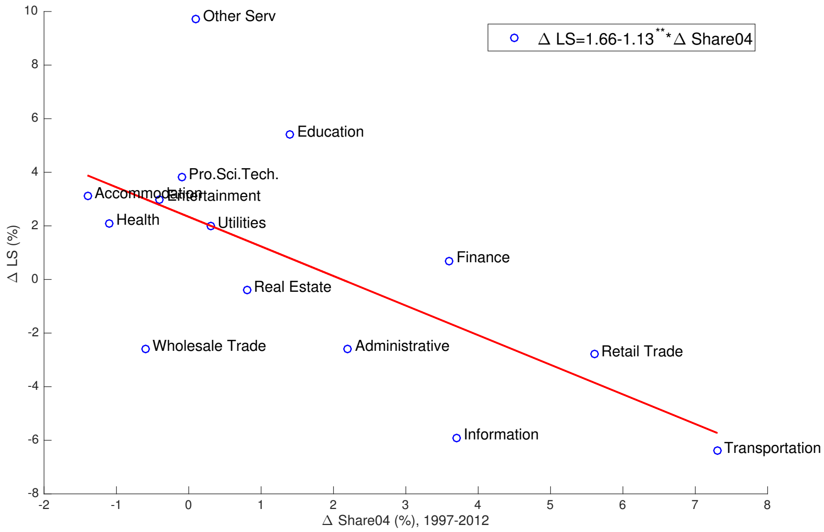


Figure A.4: ΔLS v.s. $\Delta Share04$, NON-MFG, 1997-2012

Note: LS in the vertical axis is the fraction of compensation of employees to value added. Concentration in the horizontal axis is the revenue share of 4 largest firms.

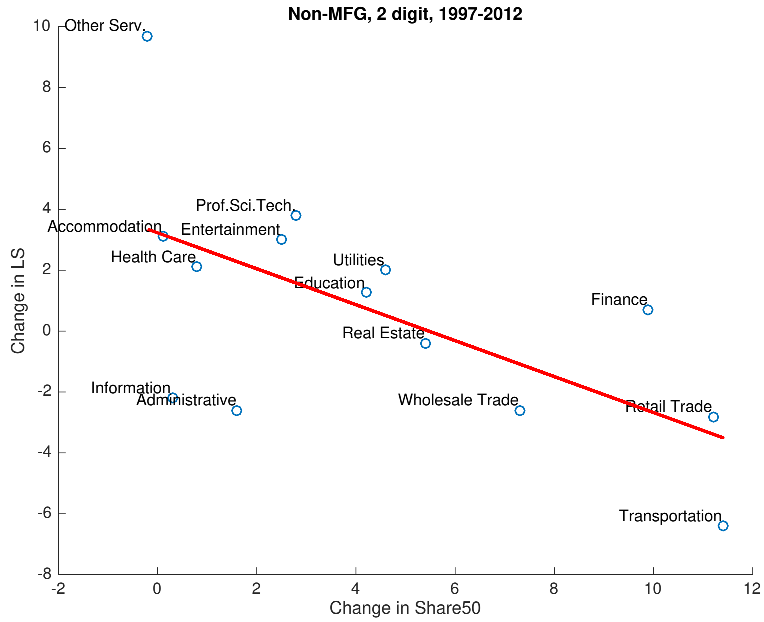


Figure A.5: ΔLS v.s. $\Delta Share50$, NON-MFG, 1997-2012

Note: Labor share on the vertical axis is measured the fraction of compensation of employees to value added. Concentration on the horizontal axis is the revenue share of 50 largest firms. In this graph, I made minimum adjustments to LS in 1997 based on data in Table A.11. In particular LS for *Information* in 1996, instead of 1997, and the average LS for *Education* between 1996 and 1998, instead of 1997, are employed.

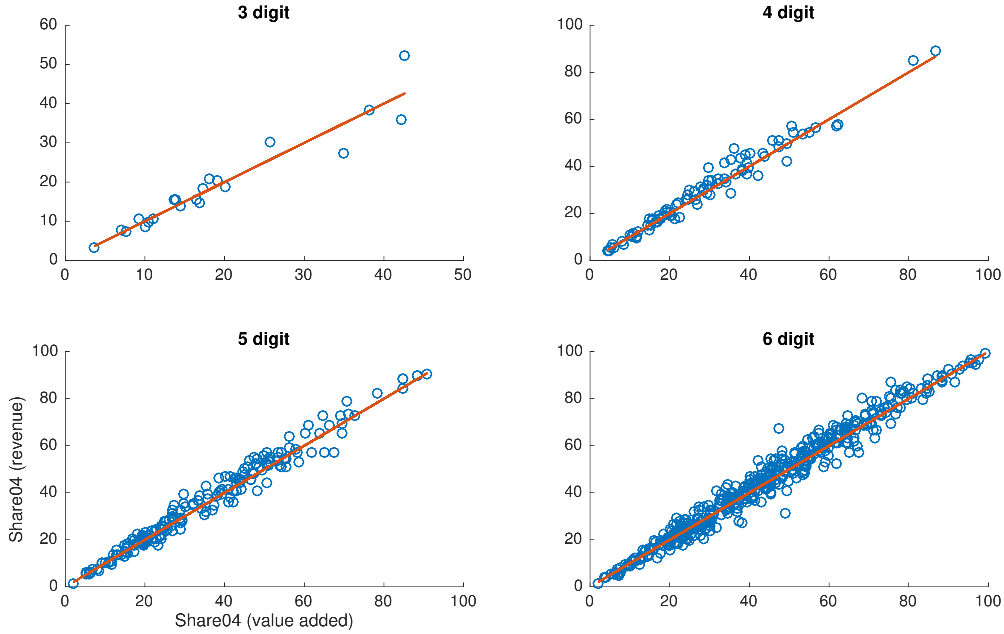


Figure A.6: Share of Top-4 Firms: Value Added versus Revenue, MFG 2002

Note: The vertical axis is the revenue share of 4 largest firms (in terms of revenue); and the horizontal axis is the value added share of 4 largest firms (in terms of value added).

Table A.5: Dependent Variable: Share04 in terms of Revenue, MFG 2002

	3 digit	4 digit	5 digit	6 digit
Share04_vadd	0.917***	0.948***	0.959***	0.956***
	(0.043)	(0.014)	(0.008)	(0.007)
R^2	0.92	0.97	0.97	0.97
Obs.	21	86	183	467

Note: The dependent variable is the revenue share of the 4 largest firms (ranked by revenue); and the independent variable is the value added share for the 4 largest firms (ranked value added).

Table A.6: Concentration in the Manufacturing Sector-I

	Share04				Share08			
Year	3-D	4-D	5-D	6-D	3-D	4-D	5-D	6-D
1997	19.82	29.99	35.12	42.37	27.43	40.29	46.53	54.04
2002	20.81	32.19	37.52	44.84	29.38	43.01	49.05	57.40
2007	20.88	32.18	37.48	44.93	30.92	44.38	50.68	59.03
2012	21.64	32.63	38.26	44.47	32.56	45.20	51.90	59.20
	Share20				Share50			
Year	3-D	4-D	5-D	6-D	3-D	4-D	5-D	6-D
1997	39.82	54.59	60.98	69.59	52.61	67.00	73.23	79.15
2002	42.21	56.63	63.07	71.20	54.81	68.44	74.56	80.07
2007	44.27	58.65	65.29	73.22	56.53	70.60	76.77	83.61
2012	46.16	60.27	67.06	74.28	58.28	72.30	78.61	84.97

Note: Share04 refers to the weighted average of revenue share for the 4 largest firms, with revenue used as weights. 3-D means NAICS 3-digit sectors. The total number of 6-digit sectors decreased from 467 in 2007 to 362 in 2012. Industrial classification codes are consistent over time at other digit levels.

Table A.7: Concentration in the Manufacturing Sector-II

Year	Share04	Share08	Share20	Share50	Obs.
1963	39.19%	51.22	65.28	76.63	410
1967	38.97	51.42	65.51	77.47	408
1972	39.65	51.83	67.23	79.15	449
1977	39.13	52.64	67.96	79.86	444
1982	36.83	49.82	65.39	78.96	441
1987	40.11	52.23	67.23	79.23	451
1992	40.34	52.25	67.82	79.81	455
1997	42.37	54.04	69.59	79.15	470
2002	44.84	57.40	71.20	80.07	471
2007	44.93	59.03	73.22	83.61	467
2012	44.47	59.20	74.28	84.97	362

Note: Share04 measures the revenue share of the 4 largest firms. The indices are weighted average across SIC 4-digit sectors before 1992 and NAICS 6-digit sectors after 1997, weighted by revenue. The last column shows the total number of sectors.

Table A.8: Total NO of Firms (unit: thousand)

Year	1977	1982	1987	1992	1997	2002	2007	2012
Economy	3147.9	3604.0	4179.8	4377.1	4752.3	4908.7	5240.0	4979.5
MFG	261.2	272.9	290.8	296.0	303.2	283.4	267.8	234.4
WHO	277.9	307.6	337.0	354.8	372.9	349.6	341.4	310.8
RET	942.8	912.7	953.0	939.8	955.6	949.5	980.0	953.0
TCP	121.3	130.5	153.9	162.2	187.9	190.1	195.6	185.4
FIRE	298.0	299.5	347.2	358.1	393.7	429.8	489.7	435.9
SRV	1122.5	1288.5	1600.8	1741.6	1924.9	2055.0	2344.1	2355.5

Note: MFG-Manufacturing; WHO-Wholesale trade; RET-Retail trade; TCP-Transportation, communication and public utilities; SRV-Services.
Source: Business Dynamics Statistics

Relative Labor Share in the Manufacturing Sector

Table A.9: Share of Industry Statistics (%), Manufacturing

Firm groups	Emp.	Payroll	Val. add.	Rel. LP	Rel. Wage	Rel. LS
1997						
50 largest	11.7%	17.3	24.5	205	148	72
50th to 100th largest	4.4	5.3	7.7	175	120	69
101st to 150th largest	3.6	4.2	5.2	144	117	81
151st to 200th largest	2.8	3.0	3.8	136	107	79
201st and smaller	77.5	70.2	59.3	73	91	118
2007						
50 largest	9.9%	14.5	25.5	258	146	57
50th to 100th largest	5.3	6.3	9.1	141	120	85
101st to 150th largest	4.3	4.7	5.3	130	108	83
151st to 200th largest	2.3	2.8	3.6	185	107	65
201st and smaller	78.3	71.8	56.5	73	91	125
2012						
50 largest	10.8%	15.5	26.1	242	144	59
50th to 100th largest	6.1	7.3	8.6	172	119	69
101st to 150th largest	4.0	4.3	5.2	123	109	89
151st to 200th largest	2.0	2.4	3.7	157	122	78
201st and smaller	77.2	70.5	56.4	72	92	127

Note: Relative labor productivity is defined as share of value added *divided by* share of employment. Relative wage is the ratio of share of payroll to that of employment. Relative labor share is the ratio of payroll share to value added share.

Table A.10: Relative LS, Relative LP and Relative Wage for Top-100 (200) MFG Firms

	1967	1972	1977	1982	1987	1992	1997	2002	2007	2012
50 largest										
Rel. LS	98%	98	97	92	92	83	72	67	57	59
Rel. LP	128	134	143	150	158	169	205	209	258	242
Rel. Wage	126	131	139	138	144	141	148	140	146	144
100 largest										
Rel. LS	94	93	94	90	87	79	71	68	60	66
Rel. LP	132	136	140	147	158	171	200	194	228	205
Rel. Wage	123	127	132	133	138	136	140	132	137	135
200 largest										
Rel. LS	93	92	92	90	87	81	73	72	65	68
Rel. LP	131	132	137	141	150	160	181	174	200	191
Rel. Wage	120	122	125	126	130	129	132	125	130	129

Data source: Census of Manufacturers.

Table A.11: Labor Share (%), 2-digit Sectors

Year	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
1987	24.6	63.7	54.1	60.5	68.8	39.9	23	6.5	60.9	69.2	86.9	78.8	50.7	64.4	60.9
1988	24.9	62.4	53.9	60.5	67.5	40.5	23.1	6.7	61.4	69.5	86.5	79.8	51.6	65.2	61.4
1989	23.7	61.3	54.3	59.9	68.7	38.7	22.4	6.5	60.9	69.4	86.4	78.9	49.5	64.8	61.2
1990	23.9	61.4	54.9	60.6	69.9	39.1	22.2	6.4	60.8	69.7	87.5	79.2	48.9	65.7	62.2
1991	24.1	61.4	53.7	60.7	68.7	38.6	21.8	6.2	61.1	70	86	79.4	50	65.5	62.3
1992	24.6	61.8	54.4	60.9	69.2	38.2	22.5	6	61.5	70.3	87	79.9	49	65.7	62
1993	24.3	61	52.2	58.6	66.6	37.7	22.6	5.9	61.3	70.8	86.9	79.7	51.9	63.7	61.3
1994	23.3	59.8	50.2	58.2	65.1	38.1	22.1	5.9	61	70.7	86.1	79.6	52.1	63	60.2
1995	22.5	58.5	51.3	58.4	65.3	39.1	21.9	5.8	62.2	72.8	86.4	80.3	52.9	62.9	61.1
1996	22.9	58.1	50.7	57.7	65.9	39	22.4	5.9	62.8	73.4	86.9	81	53.6	62	61.7
1997	24.5	56.8	50.2	57.6	65.2	42.7	23.2	5.9	65	73.8	82.5	81.5	51.9	61.2	61.2
1998	26.3	58.1	51.1	57.1	64.9	43.5	24.3	6.3	67.1	76.8	86.4	82.1	55.3	62.8	61.5
1999	26.2	58.3	51.6	57.5	66.5	45.2	24.2	6.2	67.4	76.1	86.4	82.1	55.1	62.5	61.1
2000	27.9	59	52.5	58.3	66.8	52.1	24.6	6.4	70.5	77.6	86.5	82.1	55.8	61.1	61.4
2001	29.3	59.6	53.4	58.4	67.6	49.3	24.4	6.1	69	75.8	86.9	81.8	60	62.6	66.5
2002	30.3	57.5	53.3	58	67.6	41.6	23.6	6.1	65.2	75.7	89.9	81.7	59.27	60.9	66.1
2003	28.3	55.3	52.1	57.3	65.3	40.9	23.5	6	64.5	73.3	88.4	82.2	57.7	61.9	68.9
2004	27.4	54	50.8	57.2	63.2	38.3	24.5	6.2	64.2	73	87.2	82.4	56.9	62	69
2005	27.9	52.6	49.9	56.3	61.7	37.6	24.1	6.1	65.6	72.2	87.2	82.9	56	62.1	67.2
2006	26.5	51.3	49.6	56.1	59.1	38.3	24.8	6.3	67.1	73.2	86.8	83.1	56.1	61.5	67.2
2007	26.8	50.9	49.9	57.7	62.5	37.1	25.4	6.1	67.6	72	87	83.7	55.9	62.8	69.9
2008	27.9	51.4	49.6	58.5	60.5	35.3	25.8	5.8	65.7	71.6	86.5	82.8	57.5	63.9	72.5
2009	26.7	48.4	49.3	56.3	60.5	35.7	22.7	5.4	66.9	69.9	85.6	82.1	56.6	63.3	71.6
2010	25.3	46.3	47.5	55.2	57.8	34	23	5.3	66.8	69.8	85.8	82.4	55.1	62.7	70
2011	26.2	46.3	48.1	55.7	58.2	35.7	23.4	5.3	67.7	70.4	86.8	83	55.4	63.5	71
2012	26.5	46.3	47.6	54.8	58.8	36.8	23.1	5.5	68.8	71.2	87.9	83.6	54.9	64.3	70.9
2013	26.8	45.7	46.7	54.4	58	36.2	23.1	5.5	69.8	71.5	88.6	84	54.7	63.9	71.1
2014	26.3	46.2	46.9	54.5	57.7	38.3	23.1	5.7	70.4	72.3	89.1	84.1	54.8	64.6	71.7
2015	27	46.2	46.8	54.2	58.1	37.8	23.3	5.8	70.4	71.3	89	83.6	54	64.4	71.2

Note: Labor share is the share of compensation in value added. (1)-Utilities; (2)-Manufacturing; (3)-Wholesale Trade; (4)-Retail Trade; (5)-Transportation and Warehousing; (6)-Information; (7)-Finance and Insurance; (8)-Real Estate, rental and leasing; (9)-Professional, Scientific and Technical Services; (10)-Administrative and Waste Management Services; (11)-Educational Services; (12)-Health Care and Social Assistance; (13)-Arts, Entertainment, and Recreation; (14)-Accommodation and Food Services; (15)-Other Services.
Data Source: NIPA Value-added-by-Industry.

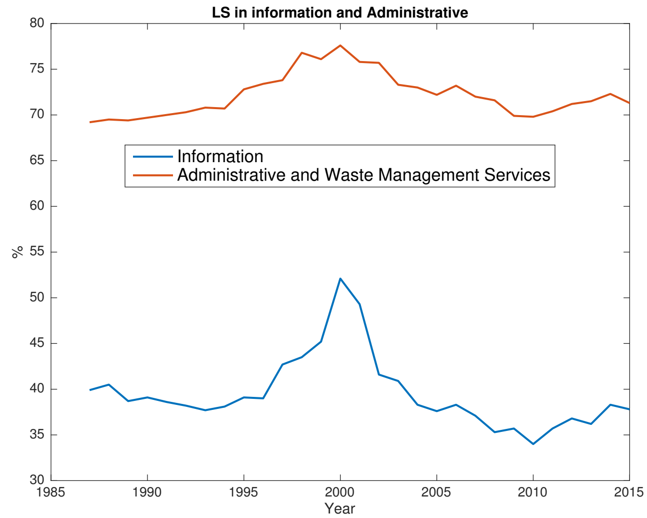


Figure A.7: Labor Share in Information and Administrative Sectors

There are surges in LS both in Information and Administrative (i.e. Administrative and Support and Waste Management and Remediation Services) sectors . This is partly caused by the realization of stock options, which is counted as labor compensation, during the internet bubble (Moylan (2008)). In 1998, labor compensation in the Administrative sector increased 11.7%, while the increase in value added is 7.4%, which results in an increase in labor share by 3% in that single year.

Information sector (NAICS 51) contains 6 sub-sectors: Publishing industries (except internet) (NAICS 511); Motion Picture and Sound Recording Industries (NAICS 512); Broadcasting (except internet) (NAICS 515); Telecommunications (NAICS 517); Data Processing, Hosting, and Related Services (NAICS 518); Other Information Services (NAICS 519). In 2012, the first and second sub-sectors account for 26%, and 15% of value added in the aggregate *Information* sector. Broadcasting and Telecommunication accounts for 52%, and the remaining 7% goes to the last two sub-sectors combined. Table A.12 lists labor share, measured as share of compensation in valued added in each sub-sectors around 2000.

Table A.12: Labor Share in Information Subsectors

	1998	1999	2000	2001	2002
Publishing industries, except internet	56.4%	50.6	65.6	65.7	51.0
Motion Pictures and sound recording industries	39.2	32.1	37.6	35.2	31.1
Broadcasting and telecommunications	35.2	38.5	39.1	38.3	36.7
Data processing, internet publishing, and other info. serv.	66.7	105.1	170.6	103.3	54.8

Table A.13: Relative Labor Share in Non-Manufacturing sectors, 2002

	(1)-Average				(2)-Weight=revenue				(3)-Weight=employment			
NAICS	1-4	5-20	21-50	≥ 51	1-4	5-20	21-50	$\geq$51	1-4	5-20	21-50	$\geq$51
1997												
2 digit	79.4%	88.0	84.3	104.9	56.3	86.9	67.9	109.0	79.8	93.2	85.7	104.0
3 digit	86.2	90.6	92.4	115.1	71.1	80.6	76.8	115.9	90.5	91.8	90.7	108.1
4 digit	83.4	91.8	99.4	113.5	80.8	83.4	90.1	118.3	90.0	92.7	97.4	109.1
5 digit	82.7	92.9	102.8	114.3	83.6	86.5	94.4	119.0	90.3	93.9	99.9	109.5
6 digit	84.4	95.3	103.9	113.8	83.7	86.3	94.8	120.3	90.8	94.0	100.0	110.0
2002												
2 digit	79.1%	83.4	84.4	107.2	65.6	72.5	78.4	113.0	79.2	89.8	86.9	106.1
3 digit	76.7	86.3	96.4	118.5	64.8	80.5	87.1	118.9	84.2	90.5	92.7	109.3
4 digit	81.7	90.4	99.7	113.7	81.3	85.8	91.9	116.8	87.4	92.3	96.4	108.9
5 digit	81.5	92.6	102.2	115.7	82.1	89.4	95.8	116.6	87.5	94.3	98.1	108.6
6 digit	83.3	94.5	102.7	115.3	83.2	88.4	96.6	118.6	88.5	94.1	98.3	109.7
2007												
2 digit	81.6%	84.4	79.2	108.0	66.1	72.3	70.7	112.8	87.0	89.8	82.4	105.9
3 digit	81.6	86.2	88.7	116.7	64.5	78.3	83.2	119.0	89.5	89.6	89.6	109.1
4 digit	83.4	89.6	97.8	112.7	82.0	84.6	87.4	116.0	89.0	92.4	93.6	108.7
5 digit	81.6	91.7	102.2	117.6	82.3	87.6	94.0	118.9	88.7	93.6	96.7	109.5
6 digit	83.0	93.6	103.4	117.1	82.6	87.8	94.5	120.6	89.8	93.5	96.9	110.4
2012												
2 digit	83.8%	80.8	87.2	107.7	59.6	63.1	78.2	115.7	87.0	87.5	89.5	106.1
3 digit	83.3	88.5	89.9	117.2	62.6	77.1	84.2	122.1	90.1	93.8	91.4	109.1
4 digit	82.8	91.6	95.6	114.3	83.5	85.5	88.8	117.3	89.9	94.2	94.2	108.5
5 digit	80.6	93.6	101.7	118.0	82.6	87.7	95.3	118.6	89.5	95.8	96.1	108.2
6 digit	82.3	95.1	103.1	116.7	82.0	87.5	96.4	119.7	89.6	95.2	96.8	108.9

Note: 1-4 denotes the 4 largest firms. For services sectors in 1997, the indices are for establishments subject to federal taxes due to data availability.

Table A.14: Relative Labor Share, 2-digit Non-Manufacturing Sectors, 2002

NAICS	Sector	Rel. LS				Rel. LP				Rel. Wage			
		1-4	5-20	21-50	≥51	1-4	5-20	21-50	≥51	1-4	5-20	21-50	≥51
22	Utilities	116.5%	98.0	105.3	90.7	105.4	108.2	97.8	92.5	122.4	106.1	103.0	84.0
42	Wholesale trade	10.4	35.0	61.2	123.8	834.1	343.2	210.8	79.3	87.1	120.2	128.9	98.1
44-45	Retail trade	95.7	97.9	90.1	102.2	96.2	92.7	89.0	103.7	92.0	90.7	80.2	106.0
48-49	Transportation	94.6	81.5	78.2	106.6	107.8	141.3	155.1	90.7	101.7	115.2	121.2	96.7
51	Information	84.6	74.0	88.9	130.7	125.9	165.2	107.5	70.7	106.5	122.3	95.6	92.4
52	Finance	91.0	76.1	70.9	118.4	124.2	128.0	159.1	82.0	113.1	97.4	112.8	97.0
53	Real Estate	93.9	64.1	86.4	106.9	87.4	163.2	165.3	92.6	82.0	104.6	142.8	99.0
54	Prof. Sci. Tech.	74.6	90.2	88.4	102.8	167.0	117.0	99.3	97.0	126.0	104.7	86.4	100.0
56	Administrative	68.9	124.8	100.8	99.5	281.7	80.2	121.1	96.4	195.0	100.1	122.1	95.9
61	Education	91.1	85.5	96.7	102.8	172.8	178.3	210.0	88.9	141.0	142.1	181.5	93.3
62	Health Care	73.1	91.2	91.0	103.3	139.9	138.0	111.8	96.5	108.6	126.0	102.6	98.5
71	Entertainment	53.1	69.6	69.7	99.8	218.9	102.0	142.6	93.9	95.8	70.0	97.4	102.7
72	Accommodation	94.0	104.1	99.9	99.8	121.0	126.5	127.4	94.3	113.7	131.7	127.3	94.1
81	Other Services			93.5	101.9			163.4	96.8			109.1	100.0

Note: Relative labor share for a subset of firms in a sector is calculated as the payroll share of these firms *divided by* their share of value added. 1-4 denotes the 4 largest firms.

Table A.15: Between-Within Decomposition of Labor Share, 1987-2013

	(1)-vadd share, %			(2)-labor share, %			(3)-decompstion	
Sector	1987	2013	change	1987	2013	change	between	within
Nonfarm Private	–	–	–	51.82	48.90	-2.92	0.84	-3.62
Utilities	3.38	2.09	-1.29	24.59	26.84	2.25	-0.33	0.06
Manufacturing	23.58	15.83	-7.75	63.69	45.67	-18.02	-4.32	-3.43
Wholesale Trd.	7.67	7.79	0.12	54.06	46.70	-7.36	0.06	-0.56
Retail Trade	9.28	7.54	-1.74	60.48	54.41	-6.07	-1.01	-0.51
Transportation	4.11	3.79	-0.32	68.80	58.02	-10.00	-0.21	-0.38
Information	5.98	6.16	0.18	39.94	36.15	-3.79	0.07	-0.23
Finance	7.58	8.84	1.26	56.12	56.35	0.22	0.70	0.02
Real Estate	15.21	16.79	1.58	6.51	5.52	-0.99	0.09	-0.16
Prof. Sci. Tech.	6.01	8.93	2.92	60.93	69.8	8.88	1.90	0.70
Administrative	2.35	3.85	1.50	69.18	71.54	2.36	1.08	0.08
Education	0.90	1.44	0.54	86.91	88.63	1.72	0.47	0.02
Health Care	6.58	9.24	2.66	78.75	84.00	5.25	2.17	0.42
Entertainment	0.90	1.28	0.33	50.69	54.66	3.97	0.21	0.05
Accommodation	3.21	3.60	0.39	64.38	63.93	-0.45	0.25	-0.02
Other Serv.	3.25	2.83	-0.42	60.87	71.08	10.21	-0.27	0.33

Note: Labor share is measured as the fraction of Compensation of employees in Value added.

Data source: BEA's Value-added-by-Industry Data.

Table A.16: Between-Within Decomposition of Labor Share (adjusted for Capital Depreciation), 1987-2013

Sector	(1)-vadd share, %			(2)-labor share, %			(3)-decompstion	
	1987	2013	change	1987	2013	change	between	within
Nonfarm Private	–	–	–	59.81	57.20	-2.61	0.45	-2.95
Utilities	2.98	1.73	-1.25	32.19	37.98	5.79	-0.43	0.13
Manufacturing	23.25	14.97	-8.28	74.54	56.51	-18.03	-5.56	-3.32
Wholesale Trade	8.15	8.49	0.34	58.74	50.13	-8.61	0.19	-0.71
Retail Trade	10.19	8.11	-2.08	63.58	59.15	-4.43	-1.29	-0.40
Transportation	3.85	3.75	-0.10	84.71	68.53	-16.18	-0.08	-0.59
Information	5.43	5.39	-0.04	50.71	48.28	-2.43	-0.02	-0.13
Finance	7.82	9.07	1.25	62.85	64.28	1.43	0.79	0.13
Real Estate	13.54	15.14	1.60	8.44	7.16	-1.28	0.13	-0.18
Prof. Sci. Tech.	6.53	9.52	2.99	64.68	76.62	11.94	2.11	1.00
Administrative	2.56	4.16	1.60	73.36	77.50	4.14	1.23	0.15
Edu. & Health Care	7.96	11.41	3.45	86.54	92.63	6.09	3.08	0.60
Entertainment	0.87	1.29	0.42	60.37	63.20	2.83	0.27	0.03
Accommodation	3.38	3.93	0.55	70.56	68.54	-2.02	0.38	-0.07
Other Services	3.50	3.04	-0.46	65.21	77.40	12.19	-0.32	0.42

Note: Value added is adjusted for consumption of fixed capital (depreciation). Education and health care are merged since the adjusted labor share in educational services exceeds 100% in some years.
Data source: BEA Value-added-by-Industry Data; NIPA Table 3.4: ESI Current-Cost Depreciation of Private Fixed Assets.

Table A.17: Share (%) of Proprietors' Income in Value-added

Sector	1998	2013	change	Sector	1998	2013	change
Nonfarm Private	6.53	6.14	-0.39	Real Estate	3.84	2.43	-1.39
Utilities	0.50	-4.49	-5.99	Prof. Sci. Tech.	19.63	16.52	-3.09
Manufacturing	1.20	1.35	0.15	Admin. & Manage.	4.10	5.12	1.02
Wholesale Trade	3.87	3.82	-0.05	Education	3.16	3.25	0.09
Retail Trade	7.34	7.91	0.57	Health Care	11.67	9.55	-2.12
Transportation	9.43	7.43	-2.00	Entertainment	15.16	14.85	-0.31
Information	1.67	2.62	0.95	Accommodation	6.21	4.49	-1.72
Finance	4.35	4.18	-0.17	Other Services	30.14	31.13	0.99

Note: Average share from 2010 to 2013, instead of 2013, for the information sector is used. Data for proprietors's income in administrative & waste management services and management of companies and enterprises are merged.
Data source: BEA Value-added-by-Industry Data, NIPA Table 6.12D: Nonfarm Proprietors' Income by Industry.

Table A.18: Concentration and Relative Labor Share, Top-50 firms

	CR (Share50)				RLS-Top50			
	1997	2002	2007	2012	1997	2002	2007	2012
Wholesale Trade	20.3	27.2	24.9	27.6	39.9	36.4	30.8	26.9
Retail Trade	25.7	31.7	33.3	36.9	96.4	95.2	95.4	92.9
Transportation	30.7	33.0	42.7	42.1	103.7	86.6	81.9	80.1
Utilities	64.5	69	70.1	69.1	94.9	104.2	102.4	107.3
Information		62	62	62.3		81.2	81.2	80.7
Finance	38.6	44.9	46	48.5	77.9	77.5	78.7	72.1
Real Estate	19.5	24.4	26.1	24.9	73.4	78.7	69.1	74.2
Prof. Sci. Tech.	16.2	16.5	18.6	19.0	84.6	85.8	84.7	93.8
Administrative	22.1	21.9	23.0	23.7	96.4	101.8	109.2	125.2
Education	19.6	23.2	23.5	23.8	79.7	90.8	87.7	90.6
Health Care	18.8	17.2	17.4	19.6	86.5	84.0	86.8	86.9
Entertainment	21.8	23.5	24.1	24.3	75.3	65.2	65.8	70.2
Accommodation	21.1	23.1	23.7	21.2	100.3	100.7	100.9	102.2
Other Services	12.8	14	13.8	12.6	94.9	88.3	91.4	100.4

Note: For most services sectors in 1997, statistics are only available for establishments subject to federal income taxes (instead of all establishments) in Service sectors. To be consistent, the same criteria is applied to 2002, 2007, and 2012.

Table A.19: Revenue Share and Relative Labor Share of Top-20 Firms

	CR (Share20)				RLS-Top20			
	1997	2002	2007	2012	1997	2002	2007	2012
Manufacturing	–	–	–	–	–	–	–	–
Wholesale Trade	12.9	18.7	16.6	18.1.6	42.1	25.2	22.4	11.1
Retail Trade	18.5	23.9	25.4	27.8	99.9	96.9	102.4	97.2
Transportation	21.8	25.2	34.9	33.7	111.3	89.2	85.3	81.1
Utilities	40.6	44.9	44.5	48	85.7	103.5	110.8	116.0
Information		48.5	49.9	50.7		79.1	77.5	69.21
Finance	22.6	28.2	28.5	31.6	80.9	81.4	81.3	68.6
Real Estate	14.1	17.1	16.3	15.8	80.9	75.4	76.7	78.8
Prof. Sci. Tech.	11.6	11.3	12.7	12.6	84.1	84.7	80.3	91.5
Administrative	14.2	14.9	15.2	16.7	88.6	102.3	117.5	133.9
Education	13.3	16	16.1	16.7	76.6	88.2	85.2	95.1
Health Care	14.2	13.3	13.1	14.9	85.9	81.9	82.3	82.3
Entertainment	15.1	14.7	15.6	15.9	70.6	62.4	62.5	63.6
Accommodation	14.8	16.5	17.4	15.1	99.3	101.1	102.5	103.4
Other Services	8.5	10	10	8.3	97.3	86.2	92.9	104.5

Note: For most services sectors in 1997, statistics are only available for establishments subject to federal income taxes, rather than all establishments. For consistency, the same criteria is applied to 2002, 2007, and 2012.

Table A.20: Revenue Share and Relative Labor Share for Top-4 Firms, 6-digit Average

	Share04				Relative LS			
	1997	2002	2007	2012	1997	2002	2007	2012
Wholesale Trade	24.3%	31.4	29.1	30.8	55.4%	51.3	51.2	47.0
Retail Trade	18.5	26.8	31.0	34.6	89.8	89.4	90.4	84.6
Transportation	23.6	24.5	30.8	35.2	94.7	89.6	92.5	89.7
Utilities	25.8	23.2	23.0	24.2	86.1	90.8	86.5	111.7
Information		49.83	52.4	52.0		87.1	92.7	92.5
Finance	26.0	32.0	36.1	35.4	94.1	96.2	92.5	97.8
Real Estate	18.8	24.0	25.1	23.3	92.9	76.4	72.8	83.2
Prof. Sci. Tech.	15.0	15.1	17.9	18.3	82.1	78.6	80.6	81.8
Administrative	21.8	23.1	24.4	24.4	94.3	94.5	95.5	99.3
Education	16.6	19.4	19.4	19.4	80.1	90.2	92.5	91.1
Health Care	16.2	15.1	15.1	17.0	92.0	90.1	97.6	97.4
Entertainment	19.2	20.5	21.5	21.6	86.1	79.6	87.8	80.0
Accommodation	13.8	17.4	18.5	16.0	103.4	103.5	107.9	105.7
Other Services	13.8	14.7	15.0	14.1	100.3	94.0	94.8	97.2

Note: Both concentration and Relative Labor Share are weighted average across NAICS 6-digit sectors, with revenue as weights.

Labor share (measured as the share of compensation in revenue). *Data Source:* Compustat.

Figure A.8: LS in Selected MFG firms, 1970-2016

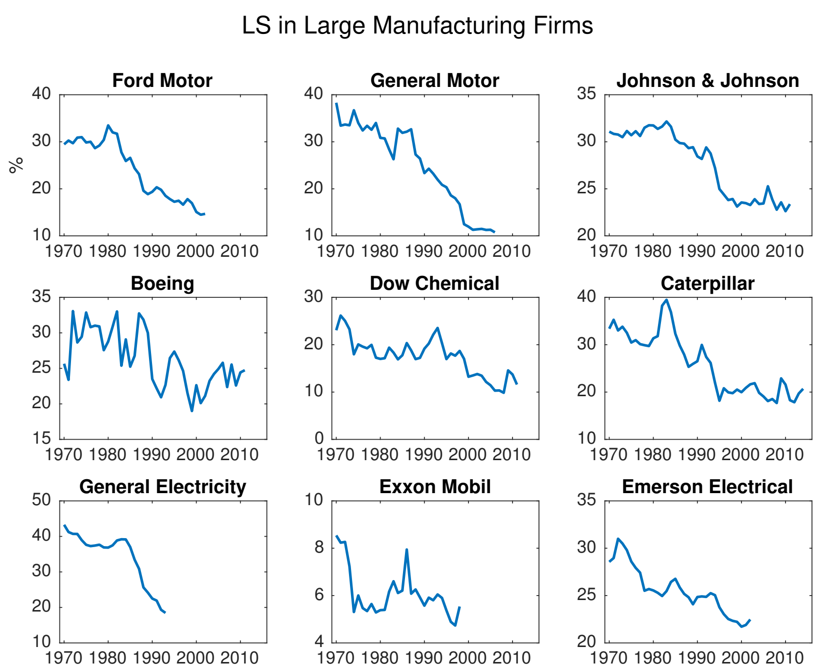


Figure A.9: LS in Selected Transportation firms, 1970-2016

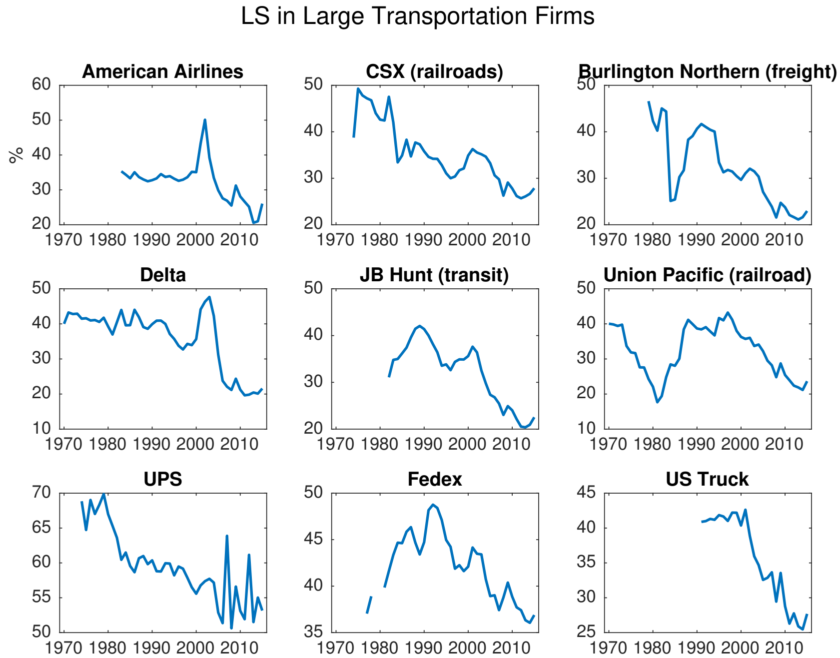


Figure A.10: LS in Selected Finance firms, 1970-2016

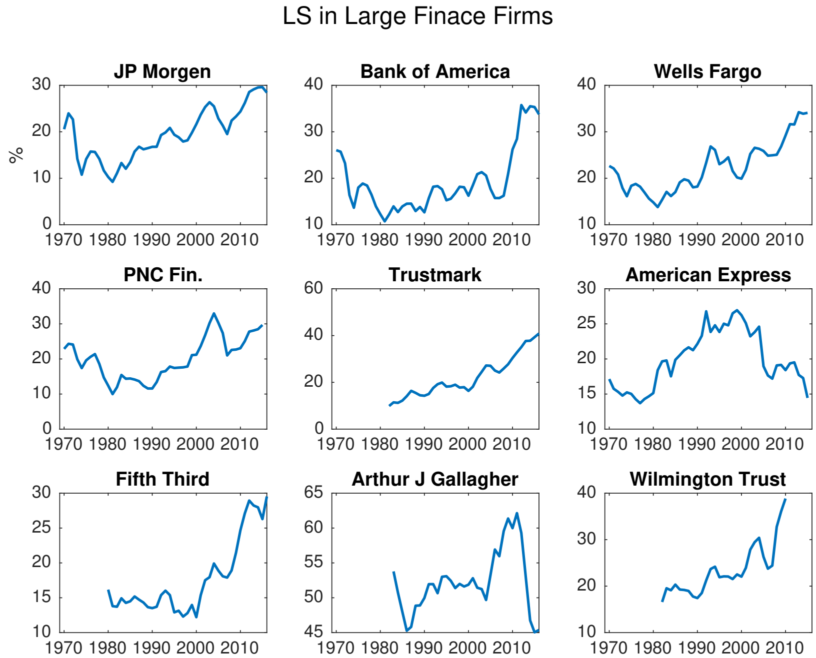
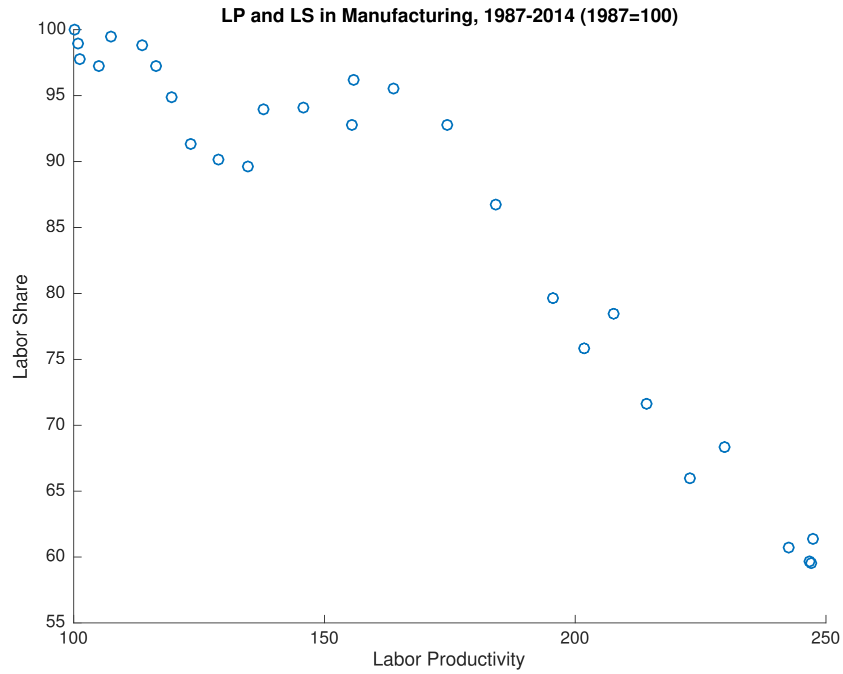


Figure A.11: Labor Productivity and Labor Share, Manufacturing 1987-2014 (1987=100)



Note: A circle represents a year.

Data Source: The 'Labor Productivity and Costs' program of Bureau of Labor Statistics.

Appendix B

Supplementary Material to Chapter 2

B.1 Homicide Reporting and Characteristics

B.1.1 Homicide and institutional background

In Qing China, local governors assumed administrative authority and also adjudicated legal cases. Under the traditional Chinese legal practice of 'life for life' sentencing, the offender in a homicide would typically receive the death penalty. During normal times, local governors were required to report all homicide and other death-penalty cases to the central government (using a standardized template called *Tiben*), since death penalties had to be reviewed and approved by the emperor.

The *Case Summary Books* are all based on the cases that were submitted by local governors in the previous year–whether or not they had been reviewed by the emperor and regardless of the case's final outcome. These books are currently maintained at the First National Historical Archives and are available for 39 (nonconsecutive) years. For each of these 39 years, the associated *Case Summary Books* include case totals by province (in China Proper, excluding Tibet, Xinjiang, Mongolia, and Manchuria; the same exclusion applies to all the statistics studied in this paper) and by type of offense. We subtract the reported non-homicide case total from the number of total cases in order to obtain each year's national number of homicides. Non-homicide cases account for about 9% of the reported 'homicide/robbery' cases. We employ this procedure for each individual province.

The *Case Summary Books* do not include cases from the national capital (Beijing) or those handled directly by the Ministry of Justice. Each year, the Ministry of Justice organized two rounds of case reviews: the *Qiushen* or Autumn Deliberations, which reviewed death-penalty cases submitted from the provinces; and the *Chaoshen* or Imperial Court Deliberations, which took place shortly after the Autumn Deliberations and reviewed death-penalty cases that originated in Beijing or were handled by the Ministry of Justice. These deliberations involved the classification of each case into one of several categories (e.g., 'facts confirmed' or *Qingshi*, probated, and undecided). Cases labeled 'facts confirmed' would be delivered to the emperor for his final ruling on whether to immediately implement the death penalty. Cases approved by the emperor and concluded by the Ministry were recorded in *Qing Shilu* (the *Qing Chronicles*). The data in the *Chronicles* indicate that *Chaoshen* cases accounted for about 6% of the total before 1790 and for 3% thereafter. We adjust each year's original *Case Summary Books* homicide total by these ratios to obtain the estimated national homicide total for each of the Books' 39years.

During the Qing Dynasty, the time allowed for the legal process before a *Tiben* filing with the Grand Secretariat was six months–excluding evidence collection time and business travel time for the relevant officers. Because these activities (and other official delays) were excluded from the time limit, it could actually take much longer before a local governor would conclude a case. FigureB.1, which is based on 49,627 cases from the *Case Summary Books*, plots the distribution of the time lag between a crime's occurrence and the case's conclusion by local governors for both homicide and non-homicide cases.[1] For homicide, more than 75% of the cases were concluded within 18 months and more than 90% of them within 24 months; in contrast, the time-lag distribution for important but non-homicide cases (illustrated by the dash-outlined bars) was relatively flatter. The implication is that homicide cases were taken more seriously by officials and handled with greater urgency, which increases our confidence in the reliability of *Tiben*-based homicide estimates.

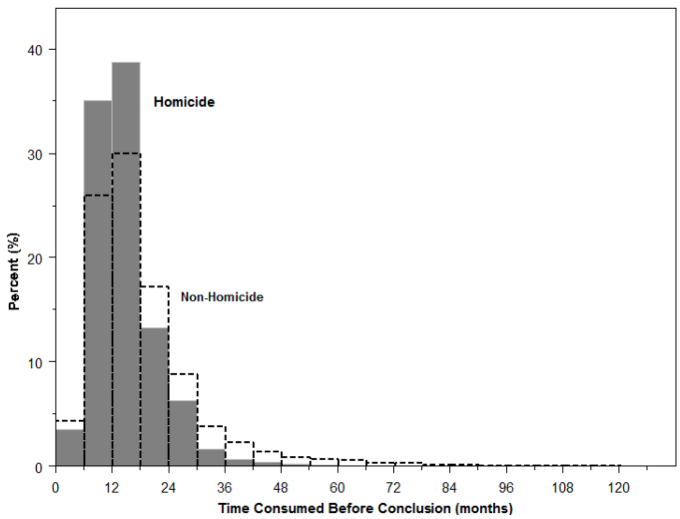


Figure B.1: Time lag, in months, between a crime's occurrence and the case's conclusion (at the province level)

There were exceptions to the procedures just described. For some homicide cases, there could be several rounds of back-and-forth communication between the local and central governments, resulting in multiple sets of memorials for the same case. Jiang (1988) reports that, in the 114 years under the consecutive reigns of Qianlong, Daoguang, and Guangxu, cases with repeated rounds of memorials accounted for less than 5% of the total Red Book reports. Note also that such cases would not affect the *Case Summary Books* statistics because the counts therein included only new cases that were concluded at the province level and submitted to the Grand Secretariat in the prior year.

[1]The figure's horizontal axis is truncated at 120 months.

During times of civil war or major rebellions, local governors and generals might be given special authority to execute serious criminals (who violated military rules or caused major social disorder) without due process. As a result, these periods led to under-representation (in the memorial Archives) of the violence actually committed. Provided such periods of conflict did not last too long, these extrajudicial executions are not likely to distort the overall violence trend–especially since we approximate homicide occurrences for these years using data from adjacent years.

It should be noted that 'Summary Execution without Due Course' (*Jiu Di Zheng Fa*) was practiced mostly during and after the Taiping Rebellion of the 1850s–1860s. Summary executions occurred primarily for non-homicide cases (e.g., robbers and rebels). And even if a homicide offender was executed without due process, the case would still be reported and included in the Case Summary Books by the Ministry of Justice–although that procedure was largely ignored during the chaotic period from the 1850s to the 1880s, when most such executions occurred.

In traditional China, infanticide was not treated as homicide in Qing China, so it is not included either in the Case Summary Books or in the extant homicide case archives. According to Cockburn (1991), infanticide accounted for 10%-20% of all homicides in Kent (England). During the Middle Ages, high rates of infanticide in the Christian West reflect the Church's prohibition against–and severe punishment of–infidelity and pre-marital sex. Although China was not a Christian nation and hence not subject to religion-based infanticidal behavior, the Chinese people inhabited a patriarchal social system with a strong and long-standing preference for boys over girls. As a result, infanticide was concentrated on newborn baby girls; hence infanticide, as a percentage of total homicide, should be higher in China than in Europe. There were also other China-specific biases that led to the under-reporting of homicide. For example, the Confucian tradition viewed children as the father's property and wives as the property of their respective husbands. That tradition explains why the killing of children (including adult children) by the father, and the killing of a wife or concubine by her husband, were often not treated as serious offenses and thus were not reported or prosecuted according to normal legal procedures (Cheung, 1972). Similarly, the killing of slaves or maids by their masters was treated more lightly than other homicides. For all these reasons, including unreported infanticide, we inflated the unadjusted homicide rate series for China (as just derived) by25%.

B.1.2 Estimates of annual red books and homicide cases

No *Case Summary Books* statistics on homicide cases (and other criminal acts) are available for the period 1661-1743. We therefore rely on the estimates given by Fang (1934) as follows. We mentioned above that two copies of *Tiben* memorials are reserved in the Red Books Archives. The *Tiben* memorials submitted to the Grand Secretariat were mostly about homicide cases, although some non-homicide (but serious) crimes were also included. Fang calculated, for each month, the total number of Red Books returned from the Department of Punishment to the Archives;[2] he thus derived summary totals for each of the 540 months spanning nearly the entire Qing Dynasty (except for 18 months that he labeled 'incomplete'). For certain years, data were given for some months but not for all 12months.

Based on these 540 monthly observations from Fang (1934), we run a simple regression with dummies for each month and year (to control for seasonal effects)–and also with a dummy for whether the observation is incomplete. The coefficient for a year's dummy rep-

[2] The Grand Secretariat had six departments with access to the Red Books Archives. For the purposes of this paper, we focus on those cases returned from the Department of Punishment.

resents the year's monthly average of Red Books, so multiplying this coefficient by 12 yields the estimated total number of Red Books for that year.[3] We repeat this procedure for the years 1661-1898 and thus derive the annual Red Books estimates reported in Table2.1 and plotted in Figure2.1.

From 1744 to 1850, the Red Book counts and the national homicide totals (from *Case Summary Books*) are both available for 16 of these 107 years. The 16 years are 1748, 1751, 1754, 1755, 1759, 1760, 1762, 1809, 1823, 1826, 1830, 1834, 1835, 1837, 1844, and 1850. We use 1850 as the stop year because–even though statistics from the *Case Summary Books* are reliable with regard to homicide cases during 1851-1860, the Taiping Rebellion may have distorted the relation between their homicide count and the Red Book count. For those 16 years, the average ratio of the Red Books estimates to the national homicide totals is2.95. This ratio is relatively stable across these years, with a variance of 0.29. Note that this number is higher than 2 since many non-memorial files were also recorded in the Red Book Archives.

Between 1744 and 1850, there are 27 years with homicide data from both the *Case Summary Books* and the land/marriage memorials. For these 27 years[4], the average ratio of the national homicide total (from the *Case Summary Books*) to the number of land/marriage memorials is 1.66 with a variance of 0.07.

[3]For the year 1712, Fang (1934) estimated the number of Red Books for only one month and labeled that year as incomplete. However, the estimate for that month is more than twice as high as for months in 1711 and1713. Since reporting errors may have led to this discrepancy, we instead take the respective totals for 1707 and 1713 and use linear interpolation to obtain an estimate for1712. This approach does not affect the overall homicide trend discussed in thetext.

[4]We did not use data for 1789 and 1846 even though homicide statistics from both the *Case Summary Books* and the Land & marriage memorials are available. After comparing total memorials in these two years with their respective neighboring years, we suspected significant archive-associated losses for 1789 and 1846 and therefore excluded those years.

B.2 Post-1860s institutional change

Several changes were made after 1860 in response to the Taiping Rebellion and other conflicts. First, the emperors began issuing orders to pardon certain categories of alleged criminals, such as some homicide cases awaiting the Ministry of Justice's review;[5] In order to evaluate the impact of increased pardons on different types of crime, in FigureB.2, we plot the number of cases per year for different types of crime (i.e., robbery, adultery and rape, murder, ordinary dispute-related killings, and offenses against Confucian values) as summarized in the surviving *Case Summary Books*. It is clear that reports of pardonable offenses, such as 'robbery' and 'ordinary dispute-led killing', declined much more after 1860 than did such serious and non-pardonable crimes such as 'murder' and 'adultery and rape'. We therefore conclude that these changes in pardon policy rendered the *Case Summary Books* less useful, for our purposes, as regards years after1860. Yet because the land/marriage *Tiben* case counts were not affected, they remain reliable indicators for inferring rates of homicide violence.

[5] As Emperor Qianlong stated: 'Pardons of criminals may sometimes be issued during national celebrations or during drought, flood or other natural calamity as a way to provide relief' (Xue, 1905, vol.2). As for pardons of death-penalty offenders, the 1861 rules allowed for sentencing to be reduced by one 'degree' or more–for example, from 'death by hanging' to 'remote-area exile'–and for the death penalty to be removed except for certain types of offense (e.g., treason, anti-government rebellion, killing of parents by their children, killing of grandparents, organizing civil war, husband killing, killing of a household head by a slave or maid, massacring a family, murder, manslaughter, robbery, and witchcraft; see *Qing Huidian*, vol.731). After 1860, pardons by the emperor became more frequent and the *Case Summary Books* practice was changed to exclude certain cases (e.g., criminal sentences reduced by the Ministry of Justice without the emperor's review).

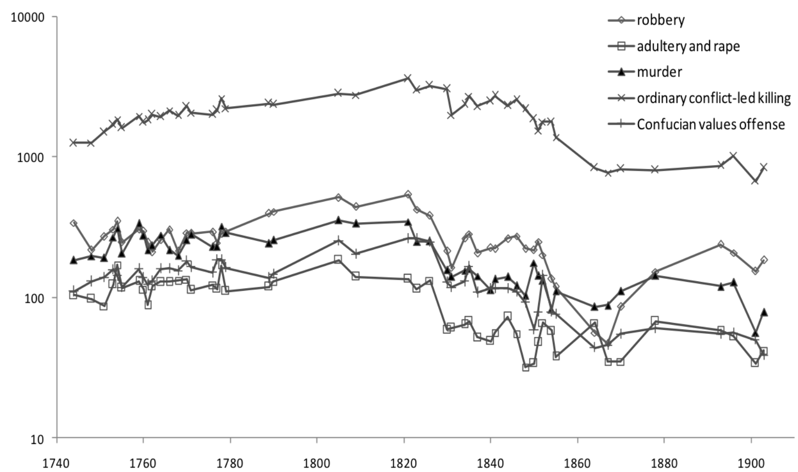


Figure B.2: Trends for different types of criminal offenses (based on *Case Summary Books*)

Note: Because some pardons were granted to perpetrators of both homicide and non-homicide capital offenses, the latter are not excluded in these plots.

B.2.1 Provincial Data and Tables

Table B.1: Homicide Cases and Homicide Rates across Provinces

Province	No. of Cases	Rate	Obs.	Province	No. of Cases	Rate	Obs.
Anhui	141	0.48	26	*Shaanxi*	121	1.21	29
	(46)	(0.12)			(56)	(0.51)	
Jiangsu	138	0.39	27	*Yunnan*	73	0.78	29
	(28)	(0.10)			(27)	(0.20)	
Zhejiang	110	0.44	32	*Guizhou*	102	1.50	30
	(28)	(0.12)			(41)	(0.41)	
Fujian	117	0.81	31	*Henan*	170	0.67	31
	(48)	(0.40)			(32)	(0.12)	
Sichuan	357	1.70	31	*Guangdong*	172	0.88	32
	(148)	(0.56)			(51)	(0.31)	
Hunan	134	0.77	30	*Shanxi*	183	1.40	30
	(47)	(0.22)			(52)	(0.47)	
Hubei	136	0.77	28	*Shandong*	150	0.49	30
	(33)	(0.20)			(44)	(0.12)	
Guangxi	76	0.87	28	*Jiangxi*	142	0.70	29
	(32)	(0.27)			(31)	(0.15)	
Gansu	78	0.45	30	*Zhili*	199	0.98	26
	(35)	(0.18)			(65)	(0.33)	

Note: Homicide cases and rates refer to annual total of homicide and homicides per 100,000 population, respectively, for each province from 1744 to 1849. Numbers in brackets stand for standard deviation

Table B.2: Summary Statistics for Provincial Panel Data

Variable	Mean	Std. Dev.	Min.	Max.	Obs.
Homicide rate	0.83	0.43	0.16	2.60	439
Panel Date					
PopDense	124	87	26	422	445
GPrice	161	50	82	329	429
PriceCV	0.15	0.05	0.04	0.41	445
Mkt	0.27	0.45	0	1	445
War	0.06	0.33	0	2.98	445
Cross-sectional Date					
Chong	0.76	0.12	0.52	0.94	15
Gentry	4.20	1.44	2.22	8.39	15

Note: Homicide rate refers to homicide cases per 100,000 population. Panel A contains variables with panel data. PopDense is population divided by total area (square kilometers), with population data from Cao (2001). GPrice is average grain price of each province in the unit of tael per shi; grain price data if from the Grain Price Database for Qing Dynasty, Institute of Modern History at the Academia Sinica. PriceCV is coefficient of variation for grain prices across prefectures, averaged over the last 5 years. Mkt values 1 if PriceCV is below the 25th percentile, and 0 otherwise. War denotes the percentage of counties in war, data from The Chronological Timetable of War for Qing China. Gentry denotes number of local gentry in Qing dynasty (unit: 10,000) from Zhang (1991). Chong is the average value of Chong ratings (geographic/strategic importance) across prefectures for each province.

Appendix C

Supplementary Material to Chapter 3

Union density, 1930-2007

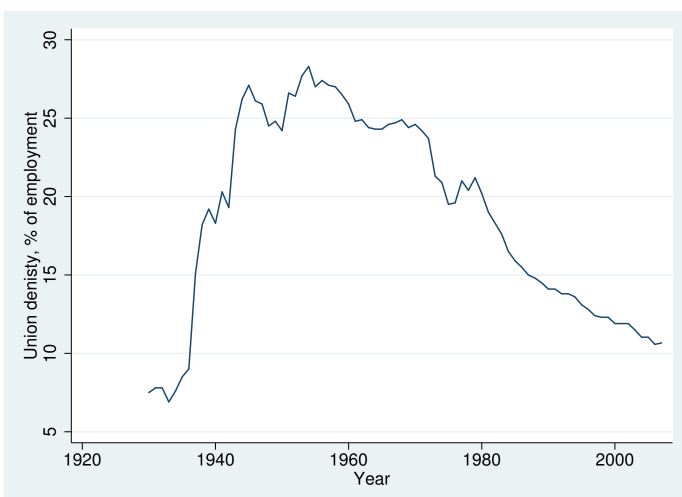


Figure C.1: Union Density

Labor Wedge, 1947-2011

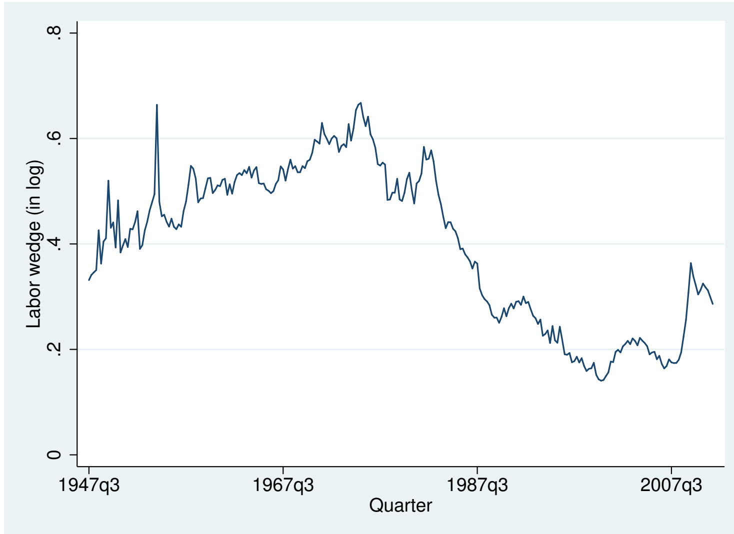


Figure C.2: Labor Wedge, 1947-2011

Structural break for the labor wedge series, The break point: 1977-Q3

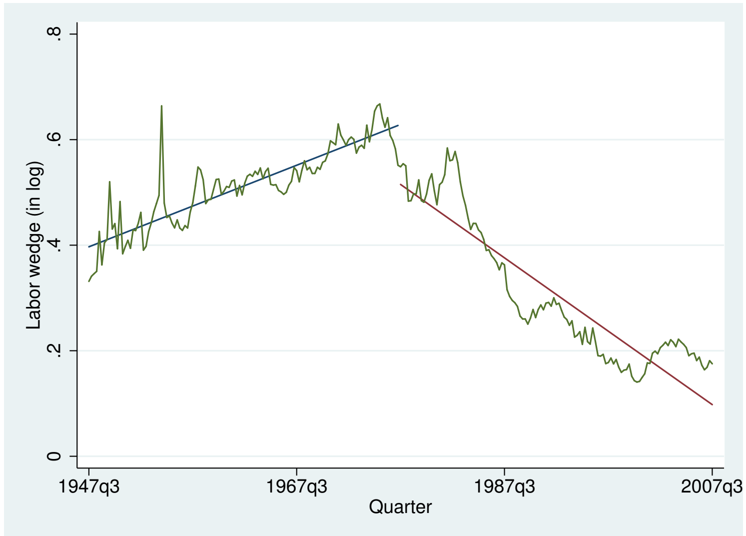


Figure C.3: Structural break test for the labor wedge series

Working hours are measured in 2 ways. In the baseline measure, it equals to average weekly working hours (from CPS) times the ratio of employment to working age population (i.e. population from 16 to 64 years old). The ratio of employment to labor force is used in the second measure.

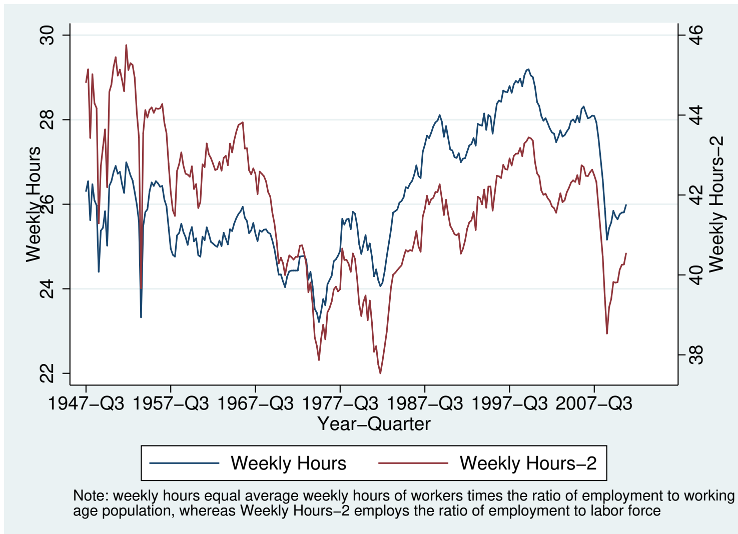


Figure C.4: Working Hours

C-Y ratio is also measured in 2 ways. It is the share of Personal Consumption Expenditure of GDP in the baseline measure. As an alternative, consumption is measured as personal consumption expenditure plus government consumption expenditure.

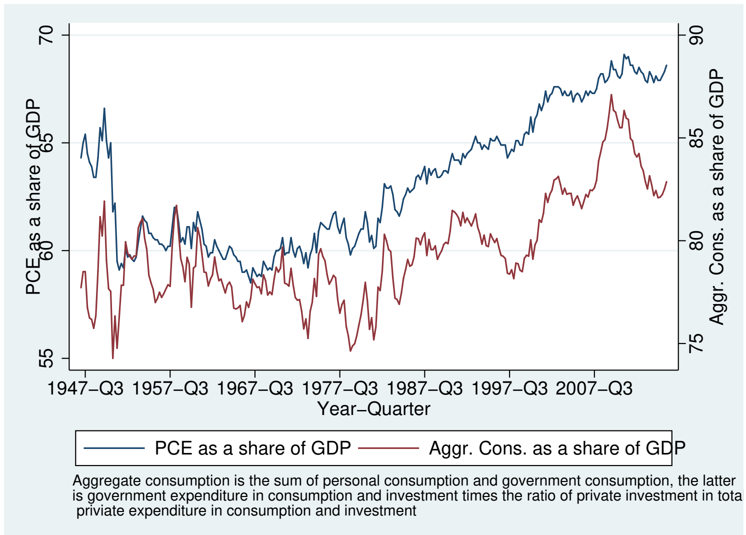


Figure C.5: Consumption-Income Ratio

Labor wedge with multiple sectors Consider a more general utility function with different sectors. Let utility from consumption be

$$\begin{aligned} \max \quad & \Sigma_{i=1}^{N}\theta_i \ln c_i \\ s.t. \quad & \Sigma_{i=1}^{N} p_i c_i = C \\ & \Sigma_{i=1}^{N}\theta_i = 1 \end{aligned}$$

where C denotes total expenditure on all consumption goods. Denote λ the Lagrangian multiplier for budget constraint. The first oder conditions are given be

$$\begin{aligned} & \frac{\theta_i}{c_i} = \lambda p_i \quad i = 1, ..., N \\ & \Sigma_{i=1}^{N} p_i c_i = C \end{aligned}$$

It follows that $\lambda = C$. Then aggregate utility from consumption is

$$\begin{aligned} \Sigma_{i=1}^{N}\theta_i \ln c_i &= \Sigma_{i=1}^{N}\theta_i \ln \frac{\theta_i C}{p_i} \\ &= \ln C + const. \end{aligned}$$

Though widely used in economics, we should caution that these results follow directly from the Cobb-Douglas utility function across different consumption goods. Deviation from C-D utility function might generate an aggregate utility function where $\frac{\partial U}{\partial C}$ depends on the vector of prices, $\{p_i\}_{i=1}^{N}$, which complicates the calculation of sectoral labor wedge.

Employment in competitive and unionized sector Labor and capital demand in unionized sector is

$$p^u(1-\alpha)\chi * A^u * (k^u)^{(1-\alpha)\chi-1} * (n^u)^{\alpha*\chi} = r$$

$$p^u\alpha\chi * A^u * (k^u)^{(1-\alpha)\chi} * (n^u)^{\alpha*\chi-1} = w^u$$

It follows from the two optimality conditions that labor demand is given by

$$n^u = (\chi A^u)^{\frac{1}{1-\chi}} * (\frac{1-\alpha}{r})^{\frac{(1-\alpha)\chi}{1-\chi}} * (\frac{\alpha}{w^u})^{\frac{1-(1-\alpha)\chi}{1-\chi}}$$

Similarly, we can solve the labor demand in the competitive sector, which is

$$n^c = (\chi A^c)^{\frac{1}{1-\chi}} * (\frac{1-\alpha}{r})^{\frac{(1-\alpha)\chi}{1-\chi}} * (\frac{\alpha}{w^c})^{\frac{1-(1-\alpha)\chi}{1-\chi}}$$

The ratio of employment in the unionized sector to that in the competitive sector is

$$\frac{n^u}{n^c} = (\frac{A^u}{A^c})^{\frac{1}{1-\chi}}(\frac{w^c}{w^u})^{\frac{1-(1-\alpha)\chi}{1-\chi}}$$

Note that the exponent $\frac{1-(1-\alpha)\chi}{1-\chi} > 1$, if $A^u = A^c$, a relatively moderate difference in w^c and w^u translates into a large difference in the $\frac{n^u}{n^c}$ ratio.

Table C.1: Union density across sectors

Naics	Industry	Union density, %	Sector
52	Finance and Insurance	2.0	c
54	Professional, Scientific, and Technical Services	2.1	c
11	Agriculture, Forestry, Fishing and Hunting	2.6	c
81	Other Services (except Public Administration)	3.2	c
72	Accommodation and Food Services	3.3	c
55	Management of Companies and Enterprises	4.5	c
53	Real Estate and Rental and Leasing	5.7	c
42	Wholesale Trade	6.4	c
56	Administrative Support and Waste Management ...	6.6	c
44-45	Retail Trade	7.5	c
71	Arts, Entertainment, and Recreation	8.2	c
62	Health Care and Social Assistance	11.0	c
21	Mining, Quarrying, and Oil and Gas Extraction	14.0	c
51	Information	14.8	c
31-33	Manufacturing	18.8	u
23	Construction	20.2	u
92	Public Administration	31.2	u
22	Utilities	33.3	u
61	Education Services	34.8	u
48-49	Transportation and Warehousing	39.1	u

Note: Union density is the average from 1983 and 2007.

Table C.2: Union Wage Premium

	1983-1984	1985-1995	1996-2007
Non-Union	2.84	2.81	2.87
Union	3.05	3.04	3.07
Premium	0.21	0.23	0.20

Data Source: CPS, Western and Rosenfeld (2011). Listed are hourly wages in log.