Essays on Inequality, Spatial Interaction, and the Demand for Skills

DISSERTATION
of the University of St. Gallen,
Graduate School of Business Administration,
Economics, Law and Social Sciences (HSG)
to obtain the title of
Doctor Oeconomiae

submitted by

David Dorn

from

Zug

Approved on the application of

Prof. Dr. Alfonso Sousa-Poza

and

Prof. Dr. Monika Bütler

Dissertation no. 3613

Difo-Druck GmbH, Bamberg

The University of St. Gallen, Graduate School of Business Administration, Economics, Law and Social Sciences (HSG) hereby consents to the printing of the present dissertation, without hereby expressing any opinion on the views herein expressed.

St. Gallen, May 19, 2009

The President

Prof. Ernst Mohr, PhD

Acknowledgements

I would like to express my sincere appreciation to my thesis advisers Alfonso Sousa-Poza, Monika Bütler, and David Autor, for generously supporting me in my dissertation project. Alfonso Sousa-Poza introduced me to labor economics while I was still in the early stage of my undergraduate studies at the University of St. Gallen. During the many years that he supervised me as a student, diploma thesis writer, and doctoral candidate, he facilitated many steps of my career and taught me a wealth of skills for research work. Monika Bütler provided crucial support at decisive steps of my graduate studies. I am particularly grateful for her encouragement to attend the Gerzensee doctoral program, and for her advice and help in establishing contacts that led to my research visits in the United States. David Autor, who hosted me at the MIT Department of Economics, worked closely with me throughout my dissertation research and he co-authored the first and second essays of this thesis. I benefited greatly from his knowledge, advice, and hospitality and I thank him for his outstanding support.

During the thesis writing stage, I also had the pleasure to visit the University of Chicago's Harris School of Public Policy Studies and the Boston University Department of Economics. I am grateful to my respective hosts Kerwin Charles and Kevin Lang for their hospitality and many inspiring discussions. I would also like to extend my thanks to the many other people whose suggestions, advice, sharing of data, or logistical support contributed to this thesis; including Daron Acemoglu, Joshua Angrist, Abhijit Banerjee, David Card, Mark Doms, Seda Ertac, Brigham Frandsen, Luis Garicano, Maarten Goos, Michael Greenstone, Gordon Hanson, Fred Henneberger, Caroline Hoxby, Sari Pekkala Kerr, Ethan Lewis, Alan Manning, Michael Manove, Alexandre Mas, Blaise Melly, Konrad Menzel, Bruce Meyer, Manuel Oechslin, Mandy Pallais, Jessica Pan, Jesse Rothstein, Hans Schmid, Daniel Sheehan, Alp Simsek, Lisa Sweeney, William Wheaton, Maisy Wong, and Conny Wunsch. I greatly appreciate my family's unwavering support of all my plans.

May 2009 David Dorn

Abstract

This thesis studies determinants of economic inequalities between different education and racial groups in the United States of America.

The first essay analyzes the sources of a recent employment and wage growth at the lower tail of the U.S. employment and wage distributions. It shows that these developments are substantially accounted for by a growth in low education, in-person service occupations. A model of changing task specialization proposes that automation displaces 'routine' clerical and production tasks in the middle of the job distribution but not low-skill service jobs which may instead benefit from from increased demand when consumers shift their consumption to outputs whose production experienced little productivity growth. An empirical analysis at the level of local labor markets over the period of 1950 through 2005 confirms that markets which were initially specialized in routine-intensive occupations experienced a stronger growth of employment and relative wages in low-skill service jobs after 1980.

The second essay studies the reallocation of workers from middle-skill occupations towards the tails of the occupational skill distribution between 1980 and 2005. It shows that the average age of workers in contracting occupations is rapidly increasing as young workers rarely move into these jobs. While young workers with college education have moved both to the upper and lower tail of the occupational skill distribution, older workers and those without college education are increasingly found in lower-skill, lower-paying jobs.

The third essay studies racial residential segregation that results when white residents flee a neighborhood once its minority resident share exceeds a critical tipping point. It proposes a model where neighborhood tipping does not only result from racial preferences but also homeowners' pecuniary incentive to sell their houses prior to a racial change to avoid a loss in house value. Hence, high rates of homeownership among white residents make neighborhoods more likely to tip. An empirical analysis of neighborhood data from a large panel of cities in 1970 to 2000 confirms that homeowners are disproportionately likely to exit a neighborhood when tipping occurs. The departure of the relatively wealthy and well-educated homeowners contributes to a drop in human capital levels of tipping neighborhoods.

Contents

Chapter 1

Inequality, Spatial Interaction, and the Demand for Skills

One of the overarching topics in labor economics is the study of economic inequalities across different segments of the population. Some population groups are disproportionately likely to face adverse outcomes such as unemployment and poverty. Two groups that are particularly challenged are those with low educational attainment and members of racial and ethnic minorities. This dissertation addresses two distinct determinants of economic opportunities for low-skilled workers and minorities: The demand for employment of low-skilled workers in local labor markets and the process of neighborhood racial segregation in cities. The analysis focuses geographically on empirical patterns in local labor markets and neighborhoods in the United States though its economic arguments may apply to a broader range of countries.

Low-skilled workers—in the context of this thesis defined as workers without college education—typically command lower wages in the labor market than their better educated peers. Their economic situation notably deteriorated in the 1980s when real wages of low-skilled workers in the United States were falling, providing a sharp contrast to the rising wages of college-educated workers. A prominent explanation for this pattern of rising inequality is an adoption of computers in the workplace which complement skilled work but substitute for low-skilled labor. This notion of a monotone skill-biased technical change provides the bleak prediction that the demand for low-skilled workers will gradually fall. More recent empirical evidence is, however, at odds with this view. Since the 1990s, the wage structure is characterized by a pattern of polarized growth where both the wages at the top and at the bottom of the wage distribution outgrow the wages in the middle. A strikingly similar polarization trend appears in the occupation structure of the labor market where the most- and least-paid occupations are expanding in employment shares. The growth of wages and employment at the bottom of the labor market calls for a rethinking of the impact of technical change on the demand for low-skilled work.

The first essay of this thesis (chapter 2) studies these recent polarization patterns in the labor market of the United States. It challenges the notion of a uniformly declining demand for low-skilled work by documenting a growth of both employment and real wages in low-skilled service occupations such as waiters, cleaners, or hairdressers. The experience of these occupations stands in sharp contrast to the declining employment and stagnating wages in low-skilled production and clerical jobs. This essay proposes a theoretical framework of technological change and unbalanced productivity growth. It argues that automation does not generally displace all low-skilled work but specifically the 'routine' clerical and production tasks in the middle of the job distribution. The tasks of low-skill service occupations which involve communication and environmental adaptability are, however, difficult to displace by technology. Instead, if consumer preferences do not admit close substitutes for the tangible outputs of service occupations, increasing output of goods (i.e., non-service activities) will raise aggregate demand for service outputs, and ultimately employment and wages in service occupations. An empirical analysis of local labor markets over the period of 1950 through 2005 shows that markets which were initially specialized in routine-intensive occupations experienced a larger growth in low-skilled service employment after 1980. These local markets which were presumably particularly susceptible to computerization also had a larger reduction in routine employment and a greater adoption of personal computers in the workplace.

The shift of the occupational structure of the labor market towards more employment in the most- and least-paid occupations changes the set of job opportunities for workers. The impact of job polarization on workers may however differ considerably between workers of different age groups. If firms react to a decreasing demand for middle-income occupations by curtailing hiring into these jobs and if young workers seek careers in more promising growing occupations, then job polarization will be particularly pronounced among the youngest workers.

The second essay (chapter 3) analyzes the changing age structure of occupations and shifts in the occupational composition of workers by age and education groups. It shows a pronounced increase in the average age of workers in the declining middle-skill occupations as employment in these jobs falls particularly among young workers. The pattern of employment reallocation to the tails of the occupational distribution varies considerably by educational attainment and age. While young college-educated workers have become more concentrated in both the most- and least-paid occupations, employment losses among older college-educated workers in the middle of the occupational skill distribution are almost entirely countered by employment growth in lower-tail jobs. Among workers without college education, even the young tend to move down to less-skilled occupations such as the growing service jobs.

Racial and ethnical minorities such as African Americans and Hispanics are another group that is particularly often affected by adverse economic outcomes. While there are many explanations for disparities between non-Hispanic whites and minorities, there is a widespread concern that minorities may suffer from the highly visible residential segregation that characterizes urban neighborhoods in the United States. Minorities may for instance be caught in a trap where low human capital levels of neighborhood peers provide an unfavorable environment for the acquisition of education. Accordingly, the mechanisms that drive and sustain segregation are of high economic interest.

The third essay of this thesis (chapter 4) addresses a process of neighborhood tipping where a mixed-race neighborhood tips towards an all-minority equilibrium once its minority share rises above a critical level. It explores the possibility that the departure of whites from a tipping neighborhood is not only driven by a preference for white neighbors but also by the concern that house prices will fall once the neighborhood tips. Homeowners who are exposed to this financial risk therefore have a larger incentive than renters to leave a neighborhood when they anticipate an increased risk of tipping. Indeed, the empirical analysis shows that discontinuous drops in white population around critical minority share tipping points are confined to a reduction in white owner-occupied housing while there is no change in white renter-occupied housing. Consequentially, neighborhoods with higher homeownership rates are less stable in terms of racial composition than those with large renter shares; they experience larger decreases in white population and house prices at the tipping points. The selective departure of homeowners from tipping neighborhoods also leads to a reduction in wealthier and better educated residents and therefore provides a link between the segregation process and low human capital levels in minority-dominated neighborhoods.

Chapter 2

Inequality and Specialization: The Growth of Low-Skill Service Jobs in the United States

2.1 Introduction

A vast body of research documents a steep rise in wage inequality in the United States starting in the 1980s. This spreading of the wage distribution is evident in the upper panel of Figure 1, which plots changes in real hourly wages by percentiles of the hours-weighted earnings using data from the Census Integrated Public Use Microsamples for 1980, 1990 and 2000 (Ruggles, Sobek, Alexander, Fitch, Goeken, Hall, King and Ronnander, 2004). During the 1980s, wage growth was strongly monotone in wage percentiles, with either zero or negative growth in the bottom quartile of the distribution, modest wage growth in the second and third quartiles, and relatively sizable wage growth in the top quartile. This monotone pattern continued in part into the decade of the 1990s, but only in the upper half of the distribution. Wage growth below the median, by contrast, reversed course: wage gains were smallest at the median and monotonically increasing at lower percentiles, giving rise to a U-shaped pattern of wage growth that has been termed 'polarization.'[1]

These diverging patterns of wage growth in the 1980s and 1990s have clear counterparts in contemporaneous changes in the structure of skilled and unskilled employment. The upper panel of Figure 2 plots changes in the share of U.S. employment by occupational skill level, where the skill level of an occupation is proxied by the mean log wages of its workers in 1980.[2] Akin to the pattern for wages, employment growth in the 1980s was

[1]Goos and Manning coin the term 'polarization' in a 2003 working paper (Goos and Manning, 2003). Acemoglu (1999), Goos and Manning (2003, 2007), Autor, Katz and Kearney (2006, 2008), Spitz-Oener (2006), Dustmann, Ludsteck and Schoenberg (2007), and Smith (2008) present evidence that employment polarization has occurred during the last two decades in the UK, West Germany and US.

[2]We use a consistent ranking based on 1980 wages to fix a baseline occupational skill level.

Figure 1a

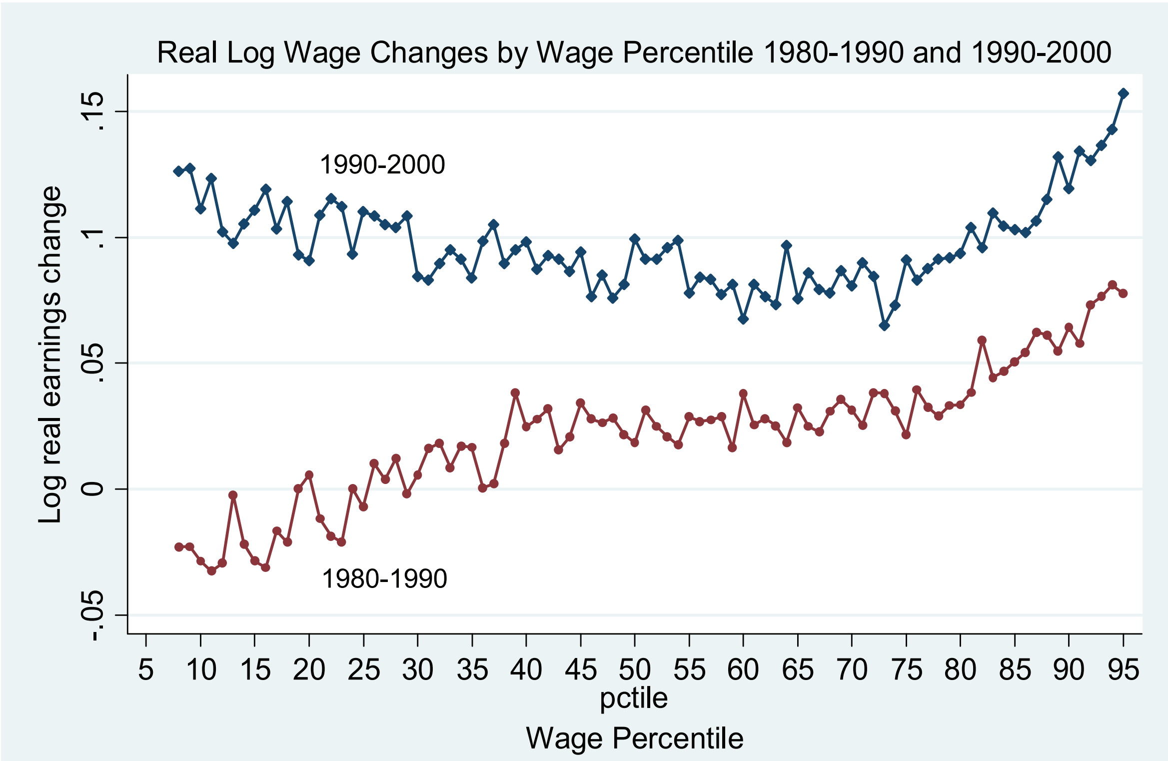
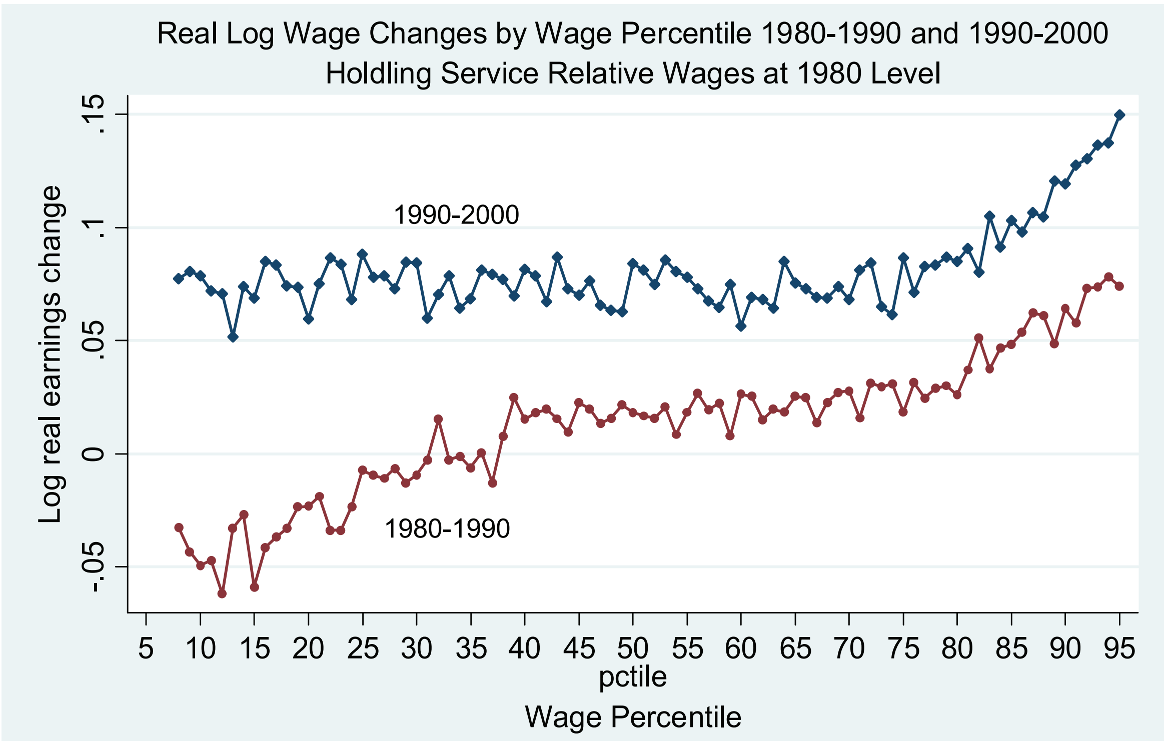


Figure 1b



strongly monotone in occupational skill levels: occupations with the lowest skill levels lost employment shares, those in the middle held constant or grew, and occupations in the top quintile expanded substantially. This monotone relationship gave way to 'polarized' employment growth during the 1990s, with occupations in both the bottom and top quintiles of the skill distribution gaining strongly in employment shares at the expense of the middle.

A comparison of changes in wages and changes in employment over these two decades warrants two conclusions. First, the clear correspondence between price and quantity movements—i.e., changes in wages and employment by percentile—in both the 1980s and 1990s suggests that demand shifts play a central part in any economic explanation of the changing structure of wages over these decades. Second, understanding these demand shifts requires explaining a central difference between the two decades, namely, the twisting of the lower-tail of the wage and employment distributions in the 1990s and forward.[3]

This paper studies both theoretically and empirically the forces behind the changing shape of low-wage and low-skill employment in the U.S. labor market. A first contribution of the paper is to document a hitherto unknown fact: the twisting of the lower tail is substantially accounted for by a single, proximate cause, which is rising employment and wages in a category of work that the Census Bureau classifies as *service occupations*. Service occupations are jobs that involve assisting or caring for others, including: food service workers; security guards; janitors, cleaners and gardeners; home health aides; child care workers; and personal appearance and recreation occupations.[4] Though among the least educated and lowest paid categories of employment, the share of U.S. labor hours in service occupations grew by 30 percent between 1980 and 2005, after having been flat or declining in the three prior decades (Table 1).[5] Simultaneously, real wage growth in service occupations averaged seven percent per decade between 1980 and 2005, substantially exceeding wage growth in other blue collar occupations (Table 1).[6]

[3]Another key difference between the two periods is that the entire locus of wage growth is shifted upward in the 1990s. This movement corresponds to the rapid productivity increases commencing in the mid 1990s (Oliner and Sichel, 2000).

[4]It is critical to distinguish service occupations, a group of low-education occupations providing personal services and comprising 14.3 percent of labor input in 2005 (Table 1), from the service *sector*, a broad category of industries ranging from health care to communications to real estate and comprising 81 percent of non-farm employment in 2000 (source: www.bls.gov).

[5]Part-time jobs are relatively prevalent in service occupations, and hence the share of service *jobs* in US employment is even larger than their share in total labor input. For example, Hecker (2005) reports that service occupations accounted for nearly one in five jobs in 2004 whereas our calculations based on the 2005 American Community Survey find that service occupations contribute approximately one in seven hours of labor input. Moreover, service occupations account for a disproportionate share of employment among workers without college education. The share of non-college hours in service occupations rose by 50 percent between 1980 and 2005, from 13.8 to 20.7 percent, while declining in all other major occupational categories.

[6]Though farm occupations are estimated to have experienced comparable wage growth in this time interval, one should place little weight on these numbers. Census data are unlikely to capture farm

Figure 2a

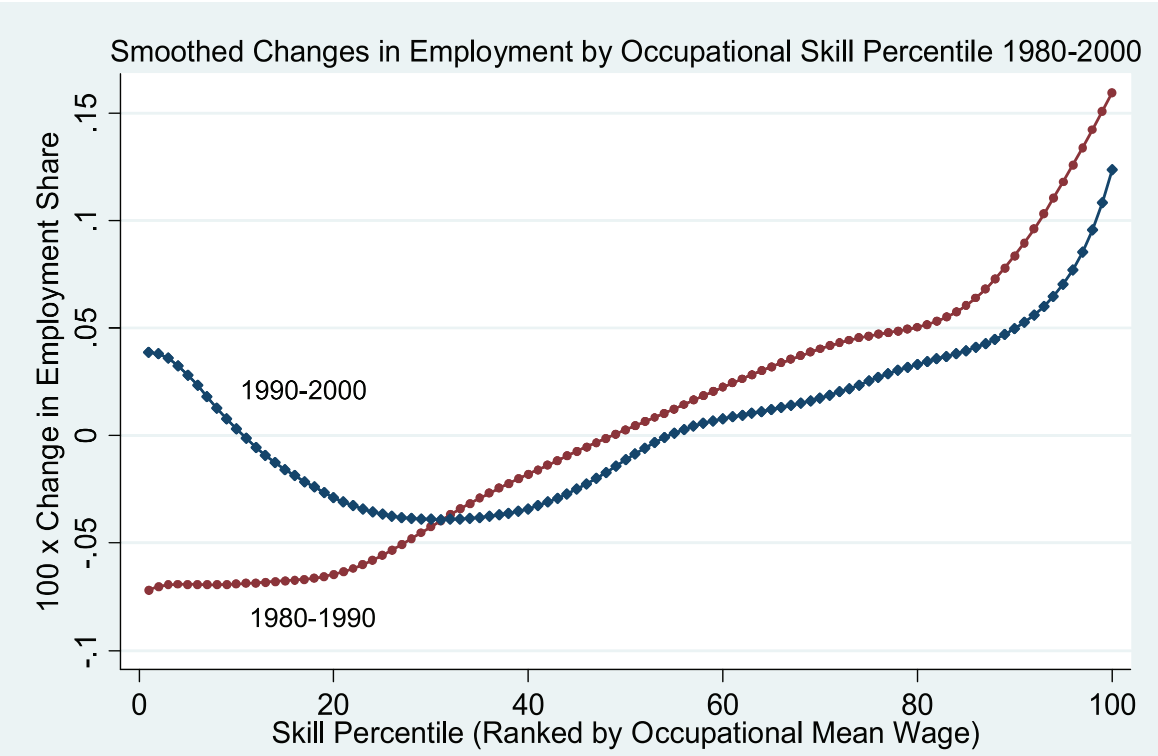


Figure 2b

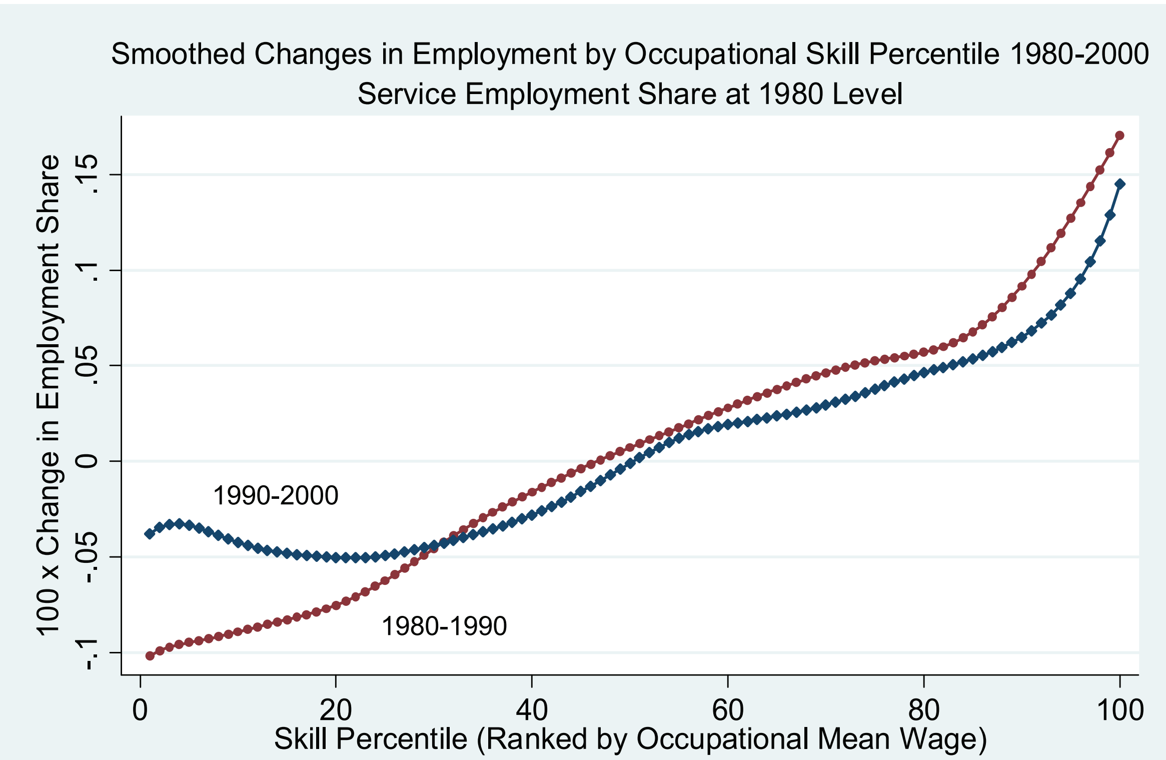


Table 1. Levels and Changes in Employment Share and Mean Real Log Hourly Wages by Occupation, 1950-2005

	Level						Decadal Growth Rate	
	1950	1970	1980	1990	2000	2005	1950-80	1980-05
			A. Share of Employment (%)					
Managers/Professionals	20.1	21.4	23.8	27.8	29.5	30.4	5.8	10.3
Technicians/Sales/Admin	21.7	26.6	28.9	30.8	30.1	29.5	10.1	0.8
Production/Craft/Repair	13.3	13.9	14.3	12.4	12.6	11.9	2.4	-6.9
Operators/Fabricat/Laborers	22.8	22.6	19.2	15.5	13.4	12.6	-5.5	-15.6
Farming/Fishery/Forestry	10.7	3.8	2.8	1.8	1.3	1.3	-36.3	-25.9
Service Occupations	11.4	11.7	11.0	11.8	13.0	14.3	-1.2	11.0
			B. Mean Real Log Hourly Wage (2005$)					
Managers/Professionals	2.24	2.88	2.87	2.95	3.07	3.19	0.21	0.13
Technicians/Sales/Admin	2.00	2.47	2.48	2.54	2.65	2.72	0.16	0.10
Production/Craft/Repair	2.20	2.70	2.73	2.68	2.71	2.73	0.18	0.00
Operators/Fabricat/Laborers	1.99	2.45	2.49	2.45	2.52	2.53	0.17	0.01
Farming/Fishery/Forestry	0.93	1.80	1.88	2.09	2.21	2.24	0.32	0.14
Service Occupations	1.51	2.06	2.16	2.21	2.32	2.33	0.22	0.07

Source: Census 1% samples for 1950 and 1970; Census 5% samples for 1980, 1990, 2000; American Community Survey 2005. Sample includes persons who were aged 18-64 and working in the prior year. Occupation categories are defined according to Census classification. Hourly wages are defined as yearly wage and salary income divided by the product of weeks worked times usual weekly hours. Employment share is defined as share in total hours worked. Labor supply is measured as weeks worked times usual weekly hours in prior year. All calculations use labor supply weights.

To see the magnitude of the contribution that service jobs make to employment and wage polarization, we consider a simple counterfactual case where employment and relative wage levels in service occupations are held at their 1980 levels. This counterfactual, shown in the lower panels of Figures 1 and 2, alters the picture of employment polarization considerably. Figure 2b shows that, holding service employment at its 1980 level, the upward twisting of the lower-tail of the employment distribution during the 1990s is largely eliminated. Moreover, this counterfactual exercise noticeably steepens the relationship between skill level and employment growth during the 1980s, reflecting the rapid growth

earnings accurately in recent decades since a substantial share of U.S. farm labor after 1980 is supplied by illegal immigrants.

of service occupations in this decade.[7] Figure 1b shows that, holding service occupation relative wages (rather than employment) constant at their 1980 level, has an analogous though less dramatic dampening effect on wage polarization during the 1990s—essentially eliminating the upward twist of the lower tail in the 1990s.[8] Similar to the case for employment, the positive relationship between wage levels and wage growth in the 1980s becomes steeper when relative wages of service occupation are held constant. These facts motivate our inquiry into the growth of service occupation employment. Because rising employment in service occupations appears central to the twisting of the lower-tail of the wage and employment distributions in the 1990s and forward, we believe that understanding their rise will provide conceptual leverage on the phenomenon of employment polarization more generally.

This paper explores the rise of service employment at the level of local labor markets. Our identification strategy exploits the fact that the output of service occupations is non-storable and non-transportable, and hence largely immune to trade and outsourcing.[9] Since consumers and producers of service occupation outputs must collate, it is fruitful to study the determinants of service employment at the detailed geographic labor market level, ideally within the local market in which service workers and service consumers both reside. We measure levels and changes in economic variables over 1980 through 2005 within 722 consistently defined, fully inclusive Commuting Zones using data from the Census IPUMS 5 percent samples for 1980, 1990 and 2000 and from the American Community Survey for 2005.[10]

The primary hypothesis that we pursue is that the rapid, secular rise in service employment since 1980 is attributable in part to non-neutral changes in productivity among job tasks spurred by advances in information technology. Concretely, this hypothesis stems

[7]The figure is generated using a simple variant of the DiNardo, Fortin and Lemieux (1996) density reweighting method. We pool Census data from either 1990 or 2000 with Census data from 1980 and estimate a weighted logit model for the odds than an observation is drawn from 1980 Census sample (relative to the actual sampling year) using as predictors a service occupation dummy and an intercept. Weights used are the product of Census sampling weights and annual hours of labor supply. We reweight observations in 1990 and 2000 using the estimated odds multiplied by the hours-weighted Census sampling weight. This procedure weights down the frequency of service occupations in 1990 and 2000 to match their 1980 frequency. Given the absence of other covariates in the model, the extra probability mass is implicitly allocated uniformly over the remainder of the distribution.

[8]We fit a weighted OLS regression in each decade of real log hourly wages on a constant and a service occupation dummy using only observations from service occupations, production, craft and repair occupations, and operator, fabricator and laborer occupations. These regressions are weighted by the product of Census sampling weights and annual hours of labor supply (annual weeks worked times average weekly hours). To produce the figure, we adjust service occupation wages in 1990 and 2000 by subtracting off the estimated service occupation premium from the current decade and replacing it with the estimated 1980 service occupation premium.

[9]Indeed, many service activities—such as hair cutting, child care, and home health assistance—require physical contact between worker and customer.

[10]An important input into our empirical analysis is a time-consistent definition of local labor markets based on 'commuting zones' (Tolbert and Sizer, 1996). Commuting zones are built from clusters of counties with strong commuting ties and are intended to approximate local US labor markets.

from the observation that the physical and interpersonal activities performed in service occupations—such as personal care, table-waiting, order-taking, housekeeping, janitorial services—have proven cumbersome and expensive to computerize. The reason, explained succinctly by (Pinker, 2007, p. 174), is that, "Assessing the layout of the world and guiding a body through it are staggeringly complex engineering tasks, as we see by the absence of dishwashers that can empty themselves or vacuum cleaners that can climb stairs."[11]

This observation motivates our theoretical model. A central thrust of recent technological change has been the automation of a large set of 'middle education' routine cognitive and manual tasks, such as bookkeeping, clerical work and repetitive production tasks (Autor, Levy and Murnane, 2003; ALM, hereafter). These tasks are readily computerized because they follow precise, well-understood procedures. Computerization of routine tasks complements the 'abstract' creative, problem-solving, and coordination tasks performed by highly-educated workers (e.g., professionals and managers), for whom data analysis is an input into production. Paradoxically, computerization of routine tasks neither directly substitutes for nor complements the core jobs tasks of numerous low-education occupations, in particular those that rely heavily on physical dexterity and flexible interpersonal communications. We refer to these activities as 'manual tasks.' Service occupations are disproportionately comprised by such manual tasks, as we document below. We hypothesize that the rapid growth of service occupations commencing in the 1980s reflects an interaction between non-neutral technological progress—which raises productivity in routine tasks but does little to augment manual tasks—and consumer preferences. In particular, if consumer preferences do not admit close substitutes for the tangible outputs of service occupations—such as restaurant meals, house-cleaning, security services, and home health assistance—increasing output of goods (i.e., non-service activities) will raise aggregate demand for service outputs, and ultimately employment and wages in service occupations.

To explore these observations formally, we analyze a simple general equilibrium model of 'routine-task' replacing technological change, building upon ALM and Weiss (2008).[12] Technological progress in this model takes the form of an ongoing fall in the cost of computerizing Routine tasks, which are performed by both machinery and low-skilled labor in the production of Goods. Automation of these tasks—a form of capital deepening—raises the productivity of high-skilled ('college') workers who perform Abstract tasks but substi-

[11] The quotation continues, "...But our sensorimotor systems accomplish these feats with ease, together with riding bicycles, threading needles, sinking basketballs, and playing hopskotch. 'In form, in moving, how express and admirable' said Hamlet about man."

[12] We modify and extend the model of Weiss (2008) to encompass two types of low-skilled labor activities—routine and manual—and to permit self-selection of low-skilled workers among these tasks. These extensions highlight the dynamics of wages and employment of low-skilled workers as they self-select between goods and services sectors in response to ongoing technical change. The limiting cases of our model are qualitatively comparable to Weiss (2008). We thank Matthias Weiss for his input on the model.

tutes for the labor input of low-skilled ('non-college') workers who perform Routine tasks. Responding to falling wages in routine tasks, non-college workers may reallocate labor supply to Service occupations, which exclusively use Manual tasks and do not experience technological progress. This labor influx causes service output to rise but has ambiguous impacts on service wages.

We study the allocation of labor between goods and services, and the inequality of wages between high and low-skill workers, as automation drives the price of routine tasks towards zero. A key result of the model is that the limiting behavior of employment and wage inequality hinges critically on the elasticity of substitution between goods and services in consumption. If goods and services are gross substitutes, ongoing technical progress ultimately drives service consumption and service employment to zero. Wage inequality between college and non-college workers rises without bounds as the wages paid to routine tasks are eroded and the productivity of abstract labor is augmented.

If, instead, goods and services are weakly complementary, non-college labor may be drawn into service occupations as goods output rises. If so, wages paid to manual tasks—and hence non-college earnings—ultimately converge to a steady growth rate, which, depending upon the complementarity between goods and services, equals or exceeds the growth rate of college wages. Thus, inequality ultimately converges to a steady-state level or collapses.[13] Numerical simulations of the model show that if goods and services are complements, the time path of wage inequality may be non-monotone. Service output grows and service wages fall as low-skilled workers initially reallocate labor from goods to services—thus, from routine to manual tasks. When labor flows to services stabilize, low-skilled wages rise. Consequently, wage inequality between high and low-skilled workers may initially increase then plateau or fall. It bears emphasis that this mechanism does *not* operate through income effects. Indeed, consumers in the model have homothetic preferences. Rather, it derives from the interaction between productivity growth and imperfect substitutability in consumption.

A primary implication of the conceptual model is that both service occupations and wage inequality between 'Abstract' and 'Routine' occupations should rise in commuting zones undergoing displacement of routine tasks. Consistent with this notion, a careful, contemporaneous study by Mazzolari and Ragusa (2008) finds robust evidence that variation across Metropolitan Statistical Area (MSA) in the growth of wage inequality over 1980 through 2005 is strongly correlated with contemporaneous growth in service employment. This pattern suggests a potential link between labor demand shifts and the growth of service employment, as posited by our model. Because cross-MSA growth variation in wage inequality is primarily treated as exogenous by the Mazzolari-Ragusa study, it is not entirely clear—at least within our conceptual framework—how this correlation should be

[13] In the latter case, the low-skilled wage rises relative to the high-skilled wage and eventually surpasses it.

interpreted.[14]

To address the potential simultaneity between wage inequality and service employment, our identification strategy draws on the theoretical model of changing task specialization. If the secularly falling price of computing leads to displacement of routine labor input, the extent of routine task displacement in local labor markets should depend on the initial concentration of routine job activities in these markets. Using task measures from the Dictionary of Occupational Titles paired to Census data on occupational structure, we generate a simple index of the share of non-college labor employed in routine task-intensive occupations in each commuting zone at the start of the relevant time period. This routine-share measure proves strikingly predictive of the changes in employment and wage structure predicted by the model. In commuting zones with an initial concentration in routine-intensive occupations, we find substantially larger growth of employment in service occupations, coupled with differential reallocation of labor input away from routine-intensive occupations. These changes in task allocation occur both in aggregate and within major education groups, with the greatest reductions in routine labor input among non-college workers. The differential growth of service employment in routine-intensive commuting zones is accompanied by a distinct pattern of wage inequality: relative wages rise in both low-skilled service occupations and highly-skilled managerial, professional, technical, sales and administrative occupations; relative wages fall across the remaining set of low-skilled occupations, consistent with a reduction in demand for routine-intensive activities. In summary, these results reveal a process of employment and wage polarization within regional labor markets that parallels the polarization of employment observed in aggregate data.

This paper contributes to a long-standing literature debating the determinants of service employment in industrialized countries. Our model of rising service employment, driven by rapid productivity growth in goods production, may be viewed as a contemporary manifestation of Baumol's (1967) classic thesis that unbalanced technical progress leads to the expansion of sectors that have relatively slow productivity growth. The model is not, however, a simple restatement of Baumol's hypothesis. We demonstrate that unbalanced productivity growth is not itself sufficient to generate rising employment in technically lagging sectors. In fact, this result depends critically on the ratio of elasticities of substitution and consumption.[15] Perhaps of greater interest, the model underscores, consistent with Weiss (2008), that ongoing, skilled-labor augmenting technical

[14]Mazzolari and Ragusa also pursue instrumental variables estimates by projecting national earnings trends in high-skill occupations onto cross-MSA differences in initial employment shares in these occupations. This approach is appropriate where wage growth in high-wage occupations is an externally determined phenomenon, as is posited by their conceptual framework. This approach would not be logical in our model.

[15]In the setting we study, a necessary condition for the Baumol result to hold is that the degree of complementarity between labor and capital in goods production, weighted by the labor share in goods, is smaller than the complementarity between goods and labor-intensive services in consumption.

change does *not* necessarily imply ongoing rises in wage inequality. Consumer preferences can attenuate or even reverse the effect of factor augmenting technical change on wage inequality in the longer run—thus leading to employment and wage polarization.

Alongside unbalanced productivity growth, the recent rise of service employment and accompanying polarization may have other contributing causes, which we consider and test below. Influential work by Clark (1957) finds that the income elasticity of demand for services is greater than unitary, implying that preferences are non-homothetic. If so, rising prosperity will increase the share of income devoted to services, even with balanced productivity growth. We refer to this as the income-effect hypothesis. A related but distinct hypothesis, explored in papers by Manning (2004), Ngai and Pissarides (2007), Mazzolari and Ragusa (2008), and Leonardi (2008), focuses on substitution rather than income effects. These studies posit that rising returns to skill in advanced countries spur high skilled workers to substitute market for home-based production of household services. This permits high skilled workers to increase their labor supply and earnings, while simultaneously raising demand for service jobs.

While these alternative explanations appear both plausible and complementary to the hypothesis we purse, we stress two key points of differentiation. A first is that our theoretical framework does not rely on either income effects in consumption or substitution effects in labor supply to generate concurrent rises in high and low-skill employment and earnings in general equilibrium. Indeed, consumers in our model have homothetic preferences and do not engage in household production.[16] Second, to the degree that we can empirically test these hypotheses, we find at best modest empirical support. Growth of service employment within commuting zones is *negatively* correlated with changes in the hours worked of male and female college graduates, which appears inconsistent with the household substitution hypothesis. Similarly, rising high wages in commuting zones, as measured by growth in the 90th wage percentile, is only weakly correlated with increases in service employment, which appears inconsistent with the income effect hypotheses. Notably, both of these alternative explanatory variables—college labor supply and rising top wages—are positively correlated with our key explanatory measure, the routine employment share, which is in turn highly robust to their inclusion.

Alongside these demand side determinants of service occupation employment, we also carefully control for a host of other likely contributors, including: a rising supply of low-skilled immigrants who may reduce the market price of services (Cortes, 2008); dwindling manufacturing employment and rising unemployment, which may reduce job opportunities for less educated workers (Harrison and Bluestone, 1988); and increases in the educational attainment, elderly population share, and female labor force participation of

[16] Since technical change in our model raises the relatively earnings of high-skilled workers, allowing either non-homothetic preferences or the possibility of household 'marketization' would of course augment the model's prediction for growing service employment.

commuting zone residents. Each of these factors might be expected to contribute to rising employment in service occupations—and indeed, all are significantly correlated with increased penetration of service jobs in local labor markets. Nevertheless, controlling for these factors does not substantially affect the main inference: regions that were initially specialized in routine-intensive occupations experienced a disproportionate degree of employment and wage polarization commencing in the 1980s.

Our study is also related to papers by Doms and Lewis (2006) and Beaudry, Doms and Lewis (2006), who explore the determinants of computer adoption and changes in education returns across MSA during the period of 1980 through 2000.[17] These papers are motivated by a model of endogenous technology adoption proposed by Beaudry and Green (2003) in which geographic variation in computer adoption is driven by the relative abundance or scarcity of skilled workers, who are complemented by computer technology. Though computer adoption is not a primary focus of our paper, we do present results on this outcome and discuss their relationship to the Beaudry, Doms and Lewis (2006) results.

In the next section, we outline a model of unbalanced productivity growth and derive implications for trends in labor allocation and wage inequality. Section 3 describes the data sources and details how we measure local labor markets, job tasks and, in particular, routine task-intensity. Sections 4 and 5 present empirical tests of our hypotheses for service employment, task specialization, and wage polarization. Section 6 concludes.

2.2 Theoretical Framework

Building on work in ALM (2003) and Weiss (2008), this section offers a simple theoretical model to explore the effects of ongoing, routine task-replacing technological change on three general equilibrium outcomes: the allocation of labor among competing low-skilled activities (in particular, routine versus manual tasks); the scale of service employment; and the inequality of wages between high and low-skill workers.

2.2.1 Environment

We consider an economy with two consumption items, goods and services, $j = g, s$ and four factors of production. Three of these factors are labor (task) inputs: Manual, Routine and Abstract ($L = m, r, a$). These labor inputs are supplied by two types of workers, $i = H, U$. The fourth factor of production is computer capital. In each sector, a continuum of mass one of firms produce consumption goods.

[17] The city-level computer adoption measure employed below was developed by Doms and Lewis (2006) and generously provided to us by the authors. This measure is also used in Beaudry, Doms and Lewis (2006).

Production of Goods combines Routine labor, Abstract labor, and computer capital (K), measured in efficiency units, using the following technology:

$$Y_g = L_a^{1-\beta}\left[(1-\lambda)(\alpha_r L_r)^{\mu} + \lambda(\alpha_k K)^{\mu}\right]^{\beta/\mu}, \tag{2.2.1}$$

with $\beta, \mu \in (0,1)$. In this production function, the elasticity of substitution between Abstract labor and the Routine task input is 1 while the elasticity of substitution between Routine labor and computer capital is $\sigma_r = 1/(1-\mu)$ and, by assumption, is greater than 1. By implication, K is a *relative complement* to Abstract labor and a *relative substitute* for Routine labor.[18]

The second sector, which produces Services, uses only Manual labor, measured in efficiency units as L_m:

$$Y_s = \alpha_s L_m, \tag{2.2.2}$$

where $\alpha_s > 0$ is an efficiency parameter. We will normalize α_s to 1 in the rest of the paper, and so α_r may be thought of as a relative efficiency term.

There is a continuum of mass one of high-skilled workers, H, who are fully specialized in Abstract labor. Each H worker supplies Abstract labor inelastically to the Goods sector.

There is a continuum of mass one of low-skilled workers, U, each of whom supplies either Manual or Routine labor. Low-skill workers have homogeneous skill at performing manual tasks. If all U workers were to perform manual tasks, they would supply a unit mass of Manual labor.

Low-skilled workers have heterogeneous skills in performing Routine tasks. Let η equal a worker's skill in routine tasks, measured in efficiency units, with density and distribution functions $f(\eta)$ and $F(\eta)$. There is a mass of one of potential Routine labor input: $\int \eta f(\eta)\, d\eta = 1$. Each worker of type U supplies labor inelastically to the task offering the highest income level given her endowment, η.

It is convenient to choose a functional form for $f(\eta)$ to permit analytic solutions of the model. The choice of functional form is innocuous, however, since the long run equilibrium of the model (i.e., as $t \to \infty$) depends on technology and preferences, not on labor supply per se. Let η be distributed exponentially on the interval $[0, \infty]$ with $f(\eta) = e^{-\eta}$.

Computer capital, K, is produced and competitively supplied using the following technology:

$$K = Y_k(t)\, e^{\delta t}/\theta. \tag{2.2.3}$$

where $Y_k(t)$ is the amount of the final consumption good allocated to production of K, $\delta > 0$ is a positive constant, and $\theta = e^{\delta}$ is an efficiency parameter. Productivity is rising

[18] In the Theory Appendix, we also consider the case where $\mu < 0$ and so L_r and K are gross complements.

at δ, reflecting technological progress. At time 1, one unit of the consumption good Y can be used to produce one efficiency unit of computer capital:

$$1 = e^{\delta}/\theta. \tag{2.2.4}$$

Competition will guarantee that the real price of computer capital (measured in efficiency units) is equal to marginal (and average) cost. So, at time $t = 1$, $p_k = 1$. As time advances, this price falls, with

$$p_k = \frac{Y_k}{K} = \theta e^{-\delta t}. \tag{2.2.5}$$

All workers/consumers have identical CES utility functions defined over consumption of Goods and Services:

$$\begin{aligned} u_i &= \left(c_{si}^{\rho} + c_{gi}^{\rho}\right)^{1/\rho}, && (2.2.6) \\ \text{where } \rho &< 1. && (2.2.7) \end{aligned}$$

The elasticity of substitution in consumption between goods and services is $\sigma_c = 1/(1-\rho)$. Consumers hold equal shares of all firms.

Consumers take prices and wages as given and maximize utility subject to the budget constraint that wages equal consumption. Firms maximize profits taking the price of consumption goods and wages as given. The CRS technology insures that equilibrium profits will be zero.

Of interest in this model is the long-run (as $t \to \infty$) allocation of low-skilled labor to goods and services, and the evolution of inequality, measured by the Manual to Abstract and Manual to Routine wage ratios.

2.2.2 Equilibrium

We normalize the price of good g to 1, i.e. $p_g(t) = 1$ for all t, without loss of generality. We can define the equilibrium as follows.

Definition 1 *An equilibrium in this economy is a tuple of aggregate allocations and prices* $(Y_s(t), Y_g(t), C_s(t), C_g(t), K(t), L_a(t), L_m(t), L_r(t), p_s(t), w_a(t), w_m(t), w_k(t))$ *and a cutoff skill for unskilled workers* $\eta^*(t)$ *such that*

1. *The representative consumer maximizes* (2.2.6) *subject to the budget constraint*

$$C_g(t) + C_s(t)\,p_s(t) \leq L_a(t)\,w_a(t) + L_m(t)\,w_m(t) + L_r(t)\,w_r(t).$$

2. *The firms that produce services and goods maximize profits, that is*

$$w_m(t) = \alpha_s p_s(t) \tag{2.2.8}$$

$$w_a(t) = \frac{d\left(L_a(t)^{1-\beta}\left[(1-\lambda)(\alpha_r L_r(t))^\mu + \lambda(\alpha_k K(t))^\mu\right]^{\beta/\mu}\right)}{dL_a(t)} \tag{2.2.9}$$

$$w_r(t) = \frac{d\left(L_a(t)^{1-\beta}\left[(1-\lambda)(\alpha_r L_r(t))^\mu + \lambda(\alpha_k K(t))^\mu\right]^{\beta/\mu}\right)}{dL_r(t)} \tag{2.2.10}$$

$$w_k(t) = \frac{d\left(L_a(t)^{1-\beta}\left[(1-\lambda)(\alpha_r L_r(t))^\mu + \lambda(\alpha_k K(t))^\mu\right]^{\beta/\mu}\right)}{dK(t)} \tag{2.2.11}$$

The firms that can convert output goods to capital goods (within the period) maximize profits, that is

$$w_k(t) \le \theta e^{-\delta t} \text{ (with equality if } K(t) > 0) \tag{2.2.12}$$

The unskilled workers allocate their labor between routine and manual tasks optimally, that is

$$w_m(t) \begin{cases} \ge \eta^*(t)\, w_r(t) & \text{if } L_m(t) = 1 \\ = \eta^*(t)\, w_r(t) & \text{if } L_m(t) \in (0,1) \\ \le \eta^*(t)\, w_r(t) & \text{if } L_m(t) = 0. \end{cases} \tag{2.2.13}$$

3. *Labor and goods markets clear, that is*

$$L_a(t) = 1,$$

$$L_m(t) = \int_0^{\eta^*} e^{-\eta} d\eta = 1 - e^{-\eta^*} \tag{2.2.14}$$

$$L_r(t) = \int_{\eta^*}^1 \eta e^{-\eta} d\eta = (\eta^* + 1)\, e^{-\eta^*} \tag{2.2.15}$$

$$C_s(t) = Y_s(t) = \alpha_s L_m(t)$$

$$C_g(t) + K(t)\,\theta e^{-\delta t} = Y_g(t). \tag{2.2.16}$$

2.2.3 Capital Demand

First note that there are no dynamic linkages, hence the equilibrium at each t can be separately characterized given the level of productivity $\theta e^{-\delta t}$.

We claim that the choice of capital in this economy solves

$$\max_{K(t)\in\mathbb{R}_+} L_a(t)^{1-\beta}\left[(1-\lambda)(\alpha_r L_r(t))^\mu + \lambda(\alpha_k K(t))^\mu\right]^{\beta/\mu} - \theta e^{-\delta t} K(t). \tag{2.2.17}$$

This can be seen by combining Eqs. (2.2.11) and (2.2.12) and noting that the choice of capital satisfies the first order condition for the above concave maximization problem.

Note that, by the market clearing condition (2.2.16), the objective function for Problem (2.2.17) is equal to C_g. Therefore, the choice of capital in equilibrium maximizes net output in the economy (which is consumed by the representative agent). We denote the optimal value of Problem (2.2.17) $F(L_a(t), L_r(t), t)$. We have that $F(L_a(t), L_r(t), t)$ is strictly increasing and differentiable in $L_a(t)$ and $L_r(t)$ with derivatives

$$w_r = \frac{dF(L_a(t), L_r(t), t)}{dL_r(t)} \tag{2.2.18}$$

$$w_a = \frac{dF(L_a(t), L_r(t), t)}{dL_a(t)} \tag{2.2.19}$$

where the equivalence with wages w_r and w_a comes from the equilibrium conditions (2.2.10) and (2.2.9) along with the envelope theorem for Problem (2.2.17). We will not explicitly solve for F since the exact algebraic expression is messy. Instead we will derive its asymptotic properties (sufficient for our analysis) for each of the cases we analyze below.

2.2.4 Demand for Manual Labor

We next derive a demand and a supply curve for $L_m(t)$ given price p_s, which will characterize the static equilibrium. The consumer optimization implies

$$p_s = \left(\frac{L_m(t)}{F(1, L_r(t), t)}\right)^{-1/\sigma_c}. \tag{2.2.20}$$

Note that, given the cutoff $\eta^*(t)$, we have that $L_m(t)$ and $L_r(t)$ are given by Eqs. (2.2.14) and (2.2.15), hence they are related with

$$\begin{aligned} L_r(t) &= (1 - \log(1 - L_m(t)))(1 - L_m(t)) \\ &\equiv g(L_m(t)), \end{aligned} \tag{2.2.21}$$

where $g : [0,1] \to [0,1]$ is a strictly decreasing function with $g(0) = 1$ and $g(1) = 0$. Plugging this in Eq.(2.2.20) gives

$$p_s = \left(\frac{F(1, g(L_m(t)), t)}{L_m(t)}\right)^{1/\sigma_c}, \tag{2.2.22}$$

which gives a demand equation for $L_m(t)$. Note that F is strictly increasing in the second variable and g is strictly decreasing, so the demand curve is strictly decreasing. Note that the demand curve starts from $p_s(L_m = 0) = \infty$ and goes down to $p_s(L_m = 1) = (F(1, 0, t))^{1/\sigma_c}$ (which is 0 when $\mu < 0$, but may be positive when $\mu > 0$).

2.2.5 Supply of Manual Labor

To derive a supply equation for $L_m(t)$, we use Eqs. (2.2.8) and (2.2.18) in the equation

$$w_m(t) = \eta^*(t)\, w_r(t).$$

to get

$$p_s(t) = \eta^*(t)\, \frac{dF(1, L_r(t), t)}{dL_r(t)}.$$

Plugging in $L_r(t) = g(L_m(t))$ and also

$$\eta^*(t) = \eta(L_m) \equiv -\log(1 - L_m(t)),$$

we have

$$p_s(t) = -\log(1 - L_m(t))\, \frac{dF(1, g(L_m(t)), t)}{dL_r(t)}. \tag{2.2.23}$$

The supply equation will typically be increasing, but it may not be increasing everywhere. It starts from $p_s(L_m = 0) = 0$ and limits to $p_s(L_m = 1) = \infty$ hence the supply and demand curves always intersect.

Putting the demand and supply equations together, we have

$$F(1, g(L_m(t)), t)^{1/\sigma_c} = -L_m(t)^{1/\sigma_c} \log(1 - L_m(t))\, \frac{dF(1, g(L_m(t)), t)}{dL_r(t)}. \tag{2.2.24}$$

which characterizes the equilibrium value of $L_m(t)$. The following proposition shows that an equilibrium always exists.

Proposition 1 *An equilibrium exists. The equilibrium level of $L_m(t)$ is characterized as the solution to Eq. (2.2.24). Once $L_m(t)$ is determined, the remaining variables are determined from the equilibrium conditions in Definition 1.*

Typically, there will be a unique intersection for supply and demand curves and we will be able to analyze the dynamics (as technology progresses) by looking at how the intersection point moves. We will study the dynamics in a simulation. Next, we will analyze the limiting behavior of this economy as $t \to \infty$.

2.2.6 Asymptotic Equilibrium

Assume (it is easy to verify this assumption) that $L_m(t)$ asymptotes to a constant in the limit, $\lim_{t\to\infty} L_m(t) = L_m^*$. Note that the Theorem of the Maximum applied to Problem (2.2.17) implies that the optimum level of $K(t)$ is increasing in t. Moreover, at $t = \infty$, cost of capital would be zero and $K = \infty$ would be optimal, hence optimal $K(t)$

will be arbitrarily large for sufficiently large t, i.e., we have $\lim_{t\to\infty} K(t) = \infty$. To make progress for solving Eq. (2.2.24) in the limit, we need to evaluate the limit values for $F(1, g(L_m(t)), t)$ and $\frac{dF(1,g(L_m(t)),t)}{dL_r(t)}$.

Capital Input

Rewrite Problem (2.2.17) as

$$\max_{K(t)\in\mathbb{R}_+} \lambda^{\beta/\mu}(\alpha_k K(t))^{\beta} \frac{\left[(1-\lambda)(\alpha_r L_r(t))^{\mu} + \lambda(\alpha_k K(t))^{\mu}\right]^{\beta/\mu}}{\lambda^{\beta/\mu}(\alpha_k K(t))^{\beta}} - \theta e^{-\delta t} K(t). \tag{2.2.25}$$

Note that the term $\frac{\left[(1-\lambda)(\alpha_r L_r(t))^{\mu} + \lambda(\alpha_k K(t))^{\mu}\right]^{\beta/\mu}}{\lambda^{\mu/\beta}(\alpha_k K(t))^{\beta}} \downarrow 1$ as $K(t) \to \infty$. This suggests that we introduce another maximization problem

$$G(1,t) = \max_{K(t)} \lambda^{\beta/\mu}(\alpha_k K(t))^{\beta} - \theta e^{-\delta t} K(t), \tag{2.2.26}$$

and denote its solution by $\tilde{K}(t)$. We claim that, in the limit, the value and the optimal solution to this maximization problem behaves like those of the optimization problem in (2.2.25). More specifically, we claim

$$\lim_{t\to\infty} \frac{F(1, g(L_m(t)), t)}{G(1,t)} = 1 \text{ and } \lim_{t\to\infty} \frac{K(t)}{\tilde{K}(t)} = 1. \tag{2.2.27}$$

To prove this statement formally, consider the first order condition for Problem (2.2.25)

$$\beta\lambda\alpha_k^{\mu} K(t)^{\mu-1}\left[(1-\lambda)(\alpha_r L_r(t))^{\mu} + \lambda(\alpha_k K(t))^{\mu}\right]^{(\beta-\mu)/\mu} = \theta e^{-\delta t}.$$

Similarly, consider the first order condition for Problem (2.2.26)

$$\beta\lambda^{\beta/\mu}\alpha_k^{\beta}\tilde{K}(t)^{\beta-1} = \theta e^{-\delta t}.$$

Dividing the last two displayed equations, taking the limit and noting that $K(t) \to \infty$ proves our claim in Eq. (2.2.27). Note that by straightforward algebra, $G(1,t)$ and $\tilde{K}(t)$ can be calculated as

$$\tilde{K}(t) = \left(\lambda^{\mu/\beta}(\alpha_k)^{\beta}\frac{e^{\delta t}}{\theta}\right)^{1/(1-\beta)} \text{ and } G(1,t) = (1-\beta)\lambda^{\mu/\beta}\alpha_k^{\beta}\tilde{K}(t)^{\beta}.$$

Combining the last equation and Eq. (2.2.27), we have

$$\lim_{t\to\infty} \frac{F(1, g(L_m(t)), t)}{c_1 K(t)^{\beta}} = 1, \tag{2.2.28}$$

where $c_1 \equiv (1-\beta)\lambda^{\mu/\beta}\alpha_k^{\beta}$ is some constant. Eq. (2.2.28) characterizes the behavior of F in the limit. In words, in the limit, routine labor become less and less important in production (since $\mu > 0$) and F behaves as a production function that does not use routine labor at all.

Next, we consider $\frac{dF(1,g(L_m(t)),t)}{dL_r(t)}$. Since $K(t) \to \infty$, we have

$$\begin{aligned}
&\lim_{t\to\infty} \frac{\frac{dF(1,g(L_m(t)),t)}{dL_r(t)}}{\beta(1-\lambda)\alpha_r^{\mu}\lambda^{(\beta-\mu)/\mu}L_r(t)^{\mu-1}(\alpha_k K(t))^{(\beta-\mu)}} \\
= &\lim_{t\to\infty} \frac{\beta(1-\lambda)\alpha_r^{\mu}L_r(t)^{\mu-1}\left[(1-\lambda)(\alpha_r L_r(t))^{\mu}+\lambda(\alpha_k K(t))^{\mu}\right]^{(\beta-\mu)/\mu}}{\beta(1-\lambda)\alpha_r^{\mu}\lambda^{(\beta-\mu)/\mu}L_r(t)^{\mu-1}(\alpha_k K(t))^{(\beta-\mu)}} \\
= &\ 1,
\end{aligned}$$

where the first line uses the expression in (2.6.5) and the last line uses the fact that $\lim_{t\to\infty} K(t) = \infty$. Hence we have

$$\lim_{t\to\infty} \frac{\frac{dF(1,g(L_m(t)),t)}{dL_r(t)}}{c_2 g(L_m(t))^{\mu-1}K(t)^{\beta-\mu}} = 1, \tag{2.2.29}$$

where $c_2 \equiv \beta(1-\lambda)\alpha_r^{\mu}\lambda^{(\beta-\mu)/\mu}\alpha_k^{\beta-\mu}$ is some constant and we have used $L_r(t) = g(L_m(t))$. This characterizes the limiting behavior for $\frac{dF(1,g(L_m(t)),t)}{dL_r(t)}$.

Labor Supply Asymptotics

We now use Eqs. (2.2.28) and (2.2.29) in Eq. (2.2.24) to solve for the asymptotic equilibrium level of $L_m(t)$. We can rewrite Eq. (2.2.24) as

$$\begin{aligned}
&\left[\frac{F(1,g(L_m(t)),t)}{c_1 K(t)^{\beta/(1-\beta)}}\right]^{1/\sigma_c} c_1^{1/\sigma_c}K(t)^{\beta/\sigma_c} \\
= &-L_m(t)^{1/\sigma_c}\log(1-L_m(t))\,c_2 K(t)^{\beta-\mu}\,g(L_m(t))^{\mu-1}\left[\frac{\frac{dF(1,g(L_m(t)),t)}{dL_r(t)}}{c_2 g(L_m(t))^{\mu-1}K(t)^{\beta-\mu}}\right]
\end{aligned}$$

which, with some algebra and using Eq. (2.2.21), can be simplified to

$$\begin{aligned}
&\frac{c_1^{1/\sigma_c}}{c_2}\frac{\left[\frac{F(1,g(L_m(t)),t)}{c_1 K(t)^{\beta/(1-\beta)}}\right]^{1/\sigma_c}}{\left[\frac{\frac{dF(1,g(L_m(t)),t)}{dL_r(t)}}{c_2 g(L_m(t))^{\mu-1}K(t)^{\beta-\mu}}\right]}K(t)^{\beta/\sigma_c-(\beta-\mu)} \\
= &-\log(1-L_m(t))\,L_m(t)^{1/\sigma_c}(1-\log(1-L_m(t)))^{\mu-1}(1-L_m(t))^{\mu-1}.
\end{aligned}$$

When we take the limit as $t \to \infty$, the terms in brackets go to 1, hence

$$\begin{aligned} & \frac{c_1^{1/\sigma_c}}{c_2} \lim_{t\to\infty} K(t)^{\beta/\sigma_c - (\beta-\mu)} && (2.2.30) \\ = \; & \lim_{t\to\infty} -\log(1-L_m(t))\, L_m(t)^{1/\sigma_c} (1-\log(1-L_m(t)))^{\mu-1} (1-L_m(t))^{\mu-1}. \end{aligned}$$

Since $K(t) \to \infty$, the left hand side either goes to 0 or ∞ depending on the sign of $\beta/\sigma_c - (\beta-\mu)$. The right hand side goes to 0 if $L_m(t) \to 0$, and to ∞ if $L_m(t) \to 1$.[19] Hence, the fact that the equality above holds in the limit implies

$$\lim_{t\to\infty} L_m(t) = \begin{cases} 0 \text{ if } \frac{1}{\sigma_c} < \frac{\beta-\mu}{\beta} \\ 1 \text{ if } \frac{1}{\sigma_c} > \frac{\beta-\mu}{\beta}. \end{cases} \quad (2.2.31)$$

In words, if share of machines in goods production is sufficiently small ($\beta < \mu$) or if goods and services are sufficiently complementary $\left(\frac{1}{\sigma_c} > \frac{\beta-\mu}{\beta}\right)$, then then in the limit all unskilled labor is drawn to manual tasks. Else if $\beta > \mu$ and $\frac{1}{\sigma_c} < \frac{\beta-\mu}{\beta}$, that is, the share of machine in goods production is large and goods and services are sufficiently substitutable, then routine tasks continue to be important in the limit and all labor is drawn to routine tasks.

Wage Inequality Asymptotics

We calculate the limiting behavior for abstract, manual and routine wages. For manual wages, we have

$$w_m(t) = p_s(t) = \left(\frac{F(1, g(L_m(t)), t)}{L_m(t)}\right)^{1/\sigma_c},$$

where we have used the demand equation. Hence, using Eq. (2.2.28), we have

$$\lim_{t\to\infty} \frac{w_m(t)}{c_1^{1/\sigma_c} \left(K(t)^\beta / L_m(t)\right)^{1/\sigma_c}} = 1, \quad (2.2.32)$$

[19] Proving that the RHS limits to ∞ as $L_m(t) \to 1$ requires some careful algebra. First, note that, as $L_m(t) \to 1$ $\lim_{t\to\infty} \frac{(1-\log(1-L_m(t)))^{\mu-1}}{-\log(1-L_m(t))^{\mu-1}} = 1$. Then, in this case the RHS limit can be rewritten as

$$(-\log(1-L_m(t)))^\mu L_m(t)^{1/\sigma_c} (1-L_m(t))^{\mu-1}.$$

Recall that we are analyzing the case $\mu > 0$. Hence the first term in this expression goes to ∞ at exponential rate. If $\mu < 1$, then the last term goes to ∞ as well and the limit is ∞ as claimed. Else if $\mu > 1$, the last term goes to 0, but it goes to zero at a polynomial rate. Since the first term goes to ∞ at an exponential rate and the last term goes to zero at polynomial rate, the product goes to ∞ as claimed. This step can more rigorously be proven using the L'Hospital Rule.

For abstract wages, we have

$$w_a(t) = \frac{dF(1, g(L_m(t)), t)}{dL_a(t)} = (1-\beta) F(1, g(L_m(t)), t),$$

hence using Eq. (2.2.28), we have

$$\lim_{t\to\infty} \frac{w_a(t)}{(1-\beta) c_1 K(t)^{\beta}} = 1. \tag{2.2.33}$$

Now using the fact that

$$w_m(t) = w_r(t)\,\eta(L_m)$$

in equilibrium, we also derive the limiting behavior for routine wages as

$$\lim_{t\to\infty} \frac{w_r(t)}{c_1^{1/\sigma_c} K(t)^{\beta/\sigma_c} \Big/ \left[L_m(t)^{1/\sigma_c} \times -\log(1-L_m)\right]} = 1. \tag{2.2.34}$$

We are also interested in relative wages. From $w_m(t) = w_r(t)\,\eta(L_m)$, we clearly have

$$\frac{w_m(t)}{w_r(t)} = \eta(L_m) = \begin{cases} 0 \text{ if } \frac{1}{\sigma_c} < \frac{\beta-\mu}{\beta} \\ \infty \text{ if } \frac{1}{\sigma_c} > \frac{\beta-\mu}{\beta}. \end{cases}$$

Also, from Eqs.(2.2.32) and (2.2.33), we have

$$\lim_{t\to\infty} \frac{w_a(t)}{w_m(t)} = \lim_{t\to\infty} \frac{(1-\beta) c_1 K(t)^{\beta}}{c_1^{1/\sigma_c} \left(K(t)^{\beta} / L_m(t)\right)^{1/\sigma_c}} = \begin{cases} \infty \text{ if } \sigma_c > 1 \\ (1-\beta) \text{ if } \sigma_c = 1 \\ 0 \text{ if } \sigma_c < 1. \end{cases}$$

Hence, we summarize our findings for wages and relative wages in this case ($\mu > 0$) as follows. We have that wages for manual and abstract labor always go to infinity. The relative wage of manual labor to routine labor $w_m(t)/w_r(t)$ go to infinity if $\frac{1}{\sigma_c} > \frac{\beta-\mu}{\beta}$ and to zero otherwise (which is, not surprisingly, the same condition which determines the limiting value of $L_m(t)$). Finally, relative wages for abstract to manual labor depend on σ_c: If $\sigma_c < 1$, then $w_a(t)/w_m(t)$ is 0; if $\sigma_c = 1$, then $w_a(t)/w_m(t)$ is $(1-\beta)$, and if $\sigma_c < 1$, then $w_a(t)/w_m(t)$ is 0. We summarize our findings in the following proposition.

Proposition 2 *When $\mu > 0$, we have $L_m(t) \to 1$ if $\frac{1}{\sigma_c} > \frac{\beta-\mu}{\beta}$ and $L_m(t) \to 0$ if*

$\frac{1}{\sigma_c} < \frac{\beta-\mu}{\beta}$. *For the limit wages, we have*

$$\begin{aligned}
\lim_{t\to\infty} \frac{w_m(t)}{w_r(t)} &= \begin{cases} \infty \text{ if } \frac{1}{\sigma_c} > \frac{\beta-\mu}{\beta} \\ 0 \text{ if } \frac{1}{\sigma_c} < \frac{\beta-\mu}{\beta}. \end{cases} \\
\lim_{t\to\infty} \frac{w_a(t)}{w_r(t)} &= \infty \\
\lim_{t\to\infty} \frac{w_a(t)}{w_m(t)} &= \begin{cases} 0 \text{ if } \sigma_c < 1, \\ \infty \text{ otherwise.} \end{cases}
\end{aligned}$$

2.2.7 Summary and Empirical Implications

In summary, the ongoing substitution of computer capital for routine labor input in our model (driven by the falling price of computer power) spurs low-skilled workers to reallocate labor input from routine tasks in goods production to manual tasks in the production of services. Employment and wages in middle-skill clerical and routine production jobs declines. Employment in low-skill service occupations rises. Wage inequality rises between high and middle-skill workers due to the combination of rising productivity of abstract tasks and and a falling price of routine tasks. Inequality between high and low-skill workers may ultimately converge to a state or may expand indefinitely. Specifically:

1. When the share of routine tasks in goods production is sufficiently small ($\beta < \mu$) or the elasticity of substitution between goods and services is sufficiently small ($1/\sigma_c > [(\beta - \mu)/\beta]$), then all unskilled labor gets allocated to manual tasks, and the wages of routine labor relative to manual labor go to 0.

2. When the share of routine tasks in goods production is sufficiently large ($\beta > \mu$) and the elasticity of substitution between goods and services is sufficiently large ($1/\sigma_c < [(\beta - \mu)/\beta]$), then all unskilled labor is allocated to routine tasks in the limit. The manual wage to routine wage ratio limits to 0. The abstract wage to routine wage ratio in this case always limits to infinity (since we necessarily have $\sigma_c > 1$). Hence, in the limit, the abstract wage is greater than the routine wage which is in turn greater than the manual wage.

3. The relative wage of abstract to manual labor limits to infinity if $\sigma_c > 1$, to zero if $\sigma_c < 1$, and to $1 - \beta$ if $\sigma_c = 1$.

One element intentionally left absent from the model is the opportunity for workers to invest in human capital.[20] While in reality, rising earnings inequality spur further skills investment, we omit this possibility from the model to emphasize that even with human

[20] Indeed, in our data, the non-college share of worked hours falls from 58 to 38 percent between 1980 and 2005.

capital stocks held constant, ongoing skilled–labor augmenting technical change need not imply ongoing growth of inequality.

Can this aggregate model be applied to the analysis of employment and wages in detailed geographic areas, such as cities or commuting zones? The answer depends on whether these areas can plausibly be treated as approximating separate markets. If yes, the model predicts that markets with higher initial concentration in routine tasks—corresponding to higher values of β in local goods production—will see greater growth of service employment and greater polarization of wages as computerization progresses.[21] If no, we must consider to what extent the model applies in local labor markets that interact in a full spatial equilibrium.

There is one key factor that aids the identification of the model in the more general, spatial equilibrium case: the output of service occupations is non-traded, and hence inter-region trade is not expected to enforce a uniform service wage across geographic areas. In the short run, local demand shocks should affect local service occupation wage levels. And the rate at which these regional wage differences are arbitraged depends upon the responsiveness of labor movements to cross-region wage variation. Much evidence suggests that mobility responses to labor demand shocks across US cities and states are typically slow and incomplete (Topel, 1986; Blanchard and Katz, 1992; Glaeser and Gyourko, 2005). Mobility is particularly low for the less-educated, who comprise the majority of service occupation workers (Bound and Holzer, 2000). It is therefore plausible that local demand shocks may affect service wages even over the medium term.

The non-tradeability of service outputs has a second useful implication: because demanders and suppliers of service occupations must collocate, the geographic analysis can potentially identify the local determinants of the demand for service jobs, even in the case when service wage levels are not set locally. Consequently, we expect the 'quantity' implications of the theoretical framework to hold at the local labor market level, even in full spatial equilibrium. The wage side of the analysis must be treated as more speculative.

2.3 Data Sources and Measurement

2.3.1 Data Sources

Our empirical analysis draws on the Census Integrated Public Use Micro Samples (Ruggles et al., 2004) for the years 1950, 1970, 1980, 1990, and 2000 and the American

[21]Formally, we could rewrite equation (2.2.1) at the city (or commuting zone) level with a city-specific routine task intensity: $y_{jg} = \alpha_g R^{b_j} A^{1-b_j}$ where j denotes cities and a higher value of b_j indicates greater initial routine task intensity. If all other preference and labor supply parameters are comparable across cities (that is, uncorrelated with b_j), a uniform decline in the routine task price that is common across cities will induce greater growth in wage inequality and service employment in high b cities.

Community Survey (ACS) for 2005.[22] The Census samples for 1980, 1990 and 2000 include 5 percent of the US population, the 1970 Census and ACS sample include 1 percent of the population, and the 1950 Census sample includes approximately 0.2 percent of the population.[23] Large sample sizes are essential for an analysis of changes in labor market composition at the detailed geographic level.

A time-consistent definition of local labor markets is a requirement for analyzing geographic variation over time. Previous research has often used Metropolitan Statistical Areas (MSAs) as a proxy for local labor markets (e.g., Beaudry et al., 2006; Mazzolari and Ragusa, 2008). MSAs are defined by the US Office for Management and Budget for statistical purposes; they consist of a large population nucleus and adjacent communities that have a high degree of social and economic integration with the core city. The geographic definition of MSAs is periodically adjusted to reflect the growth of cities. Despite efforts to improve the time-consistency of MSA definitions (e.g., Jaeger et al. 1998), the information provided by the Census Public Use Micro Samples does not allow for a consistent measurement of MSAs. This lack of geographic consistency is problematic for an analysis of changes in employment composition. Of particular concern is that the employment characteristics of the suburban areas that are gradually added to MSAs are likely to systematically differ from the characteristics of the core cities. In addition, MSAs do not cover the rural parts of the US.

This study pursues an alternative approach for the definition of local labor markets based on the concept of Commuting Zones (CZs). Tolbert and Sizer (1996) used county-level commuting data from the 1990 Census data to create 741 clusters of counties that are characterized by strong commuting ties within CZs, and weak commuting ties across CZs. Our analysis focuses on the 722 CZs that cover the entire mainland of the US, including both metropolitan and rural areas. Relative to other geographic units used for analysis of local labor markets, commuting zones have two advantages: they are based primarily on economic geography rather than incidental factors such as minimum population or state boundaries; and they cover the entire US. In addition, it is possible to use Census Public Use Micro Areas (PUMAs) to consistently match Census geography to CZs for the full period of our analysis.[24] We are not aware of prior economic research that makes use of this geographic construct.

We matched the geographic information that is available in the Census Public Use samples to the CZ geography. The most disaggregated geographic unit reported in the

[22]We do not use Census data for the year 1960 because detailed geographic information is not available.

[23]The 1950 sample-line subsample on which we rely is only one-fifth as large as the full 1 percent public use sample. We use the sample-line file because it contains education and occupation variables, which are key to our analysis.

[24]We use the Tolbert and Sizer (1996) definition of commuting zones based on commuting patterns in the 1990 Census. Tolbert and Killian (1987) earlier developed commuting zones using the 1980 Census. These commuting zones are largely but not fully identical with the 1990 definitions.

Census samples is the PUMA or, prior to 1980, the similarly defined county groups or state economic areas. A PUMA is a subarea of a state that comprises a population of 100,000 to 200,000 persons but has otherwise no clearly inherent economic interpretation. The 2000 Census splits the US into more than 2,000 PUMAs. The Census Bureau reports how the population of a PUMA is distributed over counties. If a PUMA overlaps with several counties, our procedure to match PUMAs to CZs assumes that all residents of that PUMA have the same probability of living in a given county. The aggregation of counties to CZs then allows computing probabilities that a resident of a given PUMA falls into a specific CZ. In every Census year, a clear majority of PUMAs can be matched to a single CZ, while the residents of the remaining PUMAs are attributed to several CZs using probability weights based on the relative share of a PUMA's population that falls into a given CZ. This technique allows us to calculate the population characteristics of residents of each CZ consistently in each year of our data. The detailled procedure for the geographic matching is documented in the Thesis Appendix.

Our sample of workers consists of individuals who were between age 16 and 64 and who were working in the year preceding the survey. Residents of institutional group quarters such as prisons and mental institutions are dropped along with unpaid family workers. Labor supply is measured by the product of weeks worked times usual number of hours per week. For individuals with missing hours or weeks, labor supply weights are imputed using the mean of workers in the same education-occupation cell, or, if the education-occupation cell is empty, the mean of workers in the same education group. All calculations are weighted by the Census sampling weight multiplied with the labor supply weight and the weight derived from the geographic matching process.

The computation of wages excludes self-employed workers and individuals with missing wages, weeks or hours. Hourly wages are computed as yearly wage and salary income divided by the product of weeks worked and usual weekly hours. Topcoded yearly wages are multiplied by a factor of 1.5 and hourly wages are set not to exceed this value divided by 50 weeks times 35 hours. Hourly wages below the first percentile of the national hourly wage distribution are set to the value of the first percentile. The computation of full-time full-year weekly wages is based on workers who worked for at least 40 weeks and at least 35 hours per week. Wages are deflated using the Personal Consumption Expenditure Index.

The Census classification of occupations changed over time, particularly between 1970 and 1980 and between 1990 and 2000. We use a slightly modified version of the crosswalk developed by Meyer and Osborne (2005) to create time-consistent occupation categories. Our changes create a balanced panel of 330 occupations for the years 1980 to 2005 that allows to follow a consistently defined set of occupations over time. The occupation categories of the 1950 to 1970 Census are also matched to this occupation system but not all 330 occupations are observed in every year. The designation of occupations as "service occupations" is based on the occupational classification of the 2000 Census. We subdivide service occupations into nine groups: food preparation and service workers; building and grounds cleaning workers and gardeners; health service support workers (such as health and nursing aides, but excluding practical or registered nurses); protective service workers; housekeeping, cleaning and laundry workers; personal appearance workers (such as hairdressers and beauticians); child care workers; recreation and hospitality workers (such as guides, baggage porters, or ushers); and other personal service workers. Protective service occupations are further subdivided into policemen and fire fighters, and guards. Because police officers and firefighters have much higher educational attainment and wage levels than all other service workers, we exclude them from our primary definition of service occupations (though our results are not sensitive to their inclusion). The detailed definition of the occupational classification is provided in the Thesis Appendix.

2.3.2 Measuring the 'Routine Employment Share'

Our empirical work below analyzes the degree to which commuting zones that are initially specialized in routine task activity experience polarization of employment and wages as the price of computing secularly declines. This analysis requires a summary index of employment in routine activities within commuting zones. We infer this information from the occupational composition of employment. To measure routine task-intensity in each occupation, we draw on data from ALM, who merge job task requirements—manual, routine and abstract—from the fourth edition of the US Department of Labor's *Dictionary of Occupational Titles* (US Department of Labor, 1977; 'DOT' hereafter) to their corresponding Census occupation classifications.[25] For each occupation k, we form an index of routine task-intensity, RTI:

$$RTI_k = \ln\left(\hat{R}_{k,1980}/\hat{M}_{k,1980}\right), \tag{2.3.1}$$

where $\hat{R}$ and $\hat{M}$ are, respectively, the intensity of routine and manual task input in each occupation in 1980, measured on a 0 to 10 scale.[26] This measure is rising in the relative importance of routine tasks within an occupation and falling in the relative importance of manual tasks. Since RTI does not have a cardinal scale, we standardize it with a mean of

[25] Following Autor, Katz and Kearney (2006), we collapse ALM's original five task measures to three task aggregates: the manual task index corresponds to the DOT variable measuring an occupation's demand for "eye-hand-foot coordination"; the routine task measure is a simple average of two DOT variables, "set limits, tolerances and standards," measuring an occupation's demand for routine cognitive tasks, and "finger dexterity," measuring an occupation's use of routine motor tasks; the abstract task measure is the average of two DOT variables: "direction control and planning," measuring managerial and interactive tasks, and "GED Math," measuring mathematical and formal reasoning requirements. Further details on these variables are found in Appendix Table 1 of ALM. The ALM measures are also employed by Goos and Manning (2007) and Peri and Sparber (2007) among others.

[26] For the 5 percent of microdata observations with the lowest manual task score (which is zero for most of these observations), we use the manual score of the 5th percentile.

zero and an employment weighted, cross-occupation standard deviation of unity in 1980.

This simple measure appears to capture well the job categories that motivate our conceptual framework. Table 2 shows that among the 10 most routine task-intensive occupations, 6 are clerical and accounting occupations and several others represent repetitive physical motion activities. Among the 10 least routine task intensive occupations, 4 are service occupations, and the remainder involve driving motor vehicles.[27] Appendix Table 1 lists the full set of Census service occupations and their rankings. Of these 31 occupations, 17 fall in the quintile of lowest of RTI scores and 23 of fall below the median. Thus, in the cross-section, this index appears to perform well.

To apply this index to commuting zones, we must aggregate the occupation level data to the geographic level. For ease of interpretation, we use a simple binary approach in which occupations are classified as routine task-intensive ($ROCC_k = 1$) if they fall in the top-third of the employment-weighted distribution of the RTI in 1980:

$$ROCC_k = \begin{cases} 1 & \text{if} \sum_{i=1}^{k} L_{i,1980} \leq \frac{1}{3} \sum_{i=1}^{K} L_{i,1980} \\ 0 & \text{otherwise} \end{cases} . \tag{2.3.2}$$

In this expression, L is equal to hours of labor supply in an occupation in 1980 and K is the count of occupations. We then assign each commuting zone, j, an aggregate routine-share measure (RS) equal to the fraction of employment that falls in routine task-intensive occupations in a given year:

$$RS_{jt} = \frac{\sum_{i=1}^{k} L_{jkt} \times ROCC_k}{\sum_{i=1}^{K} L_{jkt}}. \tag{2.3.3}$$

By construction, the mean of this measure is 0.33 in 1980.[28]

We perform two summary analyses to assess whether the aggregate trends in task input match the basic assumptions of the model. Table 3 provides means and standard deviations of the three DOT task variables—routine, manual, abstract—for the years 1980 through 2005. Here, each variable is standardized with mean zero and cross-occupation standard-deviation of one in 1980. Consistent with expectations, abstract tasks show a secular rise over 1980 through 2005 and routine tasks show a secular decline. The magnitudes of these changes is large. The mean abstract task score in 2005 lies 1.5 standard deviations above its 1980 mean, while the mean routine task score falls 2.3 standard

[27]Motor vehicle operation closely fits our definition of manual tasks, requiring little formal education but considerable ability to respond flexibly to a changing environment. Such occupations are classified as transportation and material moving rather than service in the Census. These occupations do not, however, possess the strong supplier-demander collocation attribute of service occupations. Thus, they are not well suited to our geographic analysis.

[28]We have experimented with alternative commuting zone routine intensity measures, including counting the share of employment in the top 20 or top 50 percent of routine occupations (rather than the top third) or simply taking the mean routine-intensity score in each commuting zone. All of these measures perform similarly in our analysis.

Table 2. Occupations with Highest and Lowest RTI Scores

A. Occupations with Highest RTI Scores	
1	Secretaries and Stenographers
2	Bank Tellers
3	Pharmacists
4	Payroll and Timekeeping Clerks
5	Motion Picture Projectionists *
6	Boilermakers
7	Butchers and Meat Cutters
8	Accountants and Auditors
9	Actuaries
10	Proofreaders

B. Occupations with Lowest RTI Scores	
1	Parking Lot Attendants
2	Fire Fighting, Prevention and Inspection *
3	Bus Drivers
4	Taxi Cab Drivers and Chauffeurs
5	Public Transportation Attendants and Inspectors *
6	Police and Detectives, Public Service *
7	Truck, Delivery, and Tractor Drivers
8	Garbage and Recyclable Material Collectors
9	Crossing Guards *
10	Railroad Coupler, Brake, and Switch Operators

Notes: Asterisk denotes service occupations according to Census occupation classification. The Routine Task Index (RTI) measures the log routine/manual task ratio for each detailled occupation. For occupations with equal RTI score, the tie is split by giving a higher ranking to the occupation with larger share in total US employment in 1980. Residual occupations groups ("miscellaneous"/"other"/"not elsewhere classified") are excluded.

deviations below its 1980 mean. The mean manual task score falls by 0.3 standard deviations between 1980 and 1990 before leveling off and eventually rising over the following fifteen years. It bears emphasis that the over-time variation in these measures is driven exclusively by shifts in occupational composition (since DOT characteristics for each occupation, based on the 1977 DOT file, are static). If, plausibly, within-occupation changes in task content trend in the same direction as between-occupation changes, our measures will understate the extent of task change.[29]

As a geographic level analogue to these occupational-level measures, Appendix Table 2 summarizes commuting zone level trends in the RS measure. The overall RS measure falls between 1980 and 2005, with the most rapid decline between 2000 and 2005.

[29]Similar results are reported by ALM, though their occupation-level data only extend to 1998.

Table 3. Levels and Changes in Standardized Task Measures, 1980-2005

	Standardized Task Score				Ten Times Average Annual Change			
	1980	1990	2000	2005	1980-1990	1990-2000	2000-2005	1980-2005
Abstract Tasks	0.00	0.89	1.33	1.47	0.89	0.44	0.27	0.59
	(1.00)	(1.17)	(1.31)	(1.30)				
Routine Tasks	0.00	-0.96	-1.75	-2.25	-0.96	-0.79	-1.01	-0.90
	(1.00)	(0.86)	(0.88)	(0.87)				
Manual Tasks	0.00	-0.32	-0.34	-0.20	-0.32	-0.02	0.28	-0.08
	(1.00)	(0.89)	(0.83)	(0.83)				

n = 722 Commuting Zones in each decade, weighted by start of period commuting zone share of national population. Abstract, Routine and Manual task measures are based on the Dictionnary of Occupational Titles (DOT) and defined according to Autor-Levy-Murnane (2003). Task scores by commuting zones are standardized to a mean of zero and a standard deviation of one in 1980.

Disaggregating the RS measure by education group reveals that employment in routine task-intensive occupations is always highest among workers with a high school degree or some college education ('middle educated' workers in our model), and lower for college graduates and, particularly, high school dropouts. Notably, he aggregate decline in RS over 1980 through 2005 occurs for all four education groups, with the largest declines for the education groups initially most specialized in routine occupations. Taken together, these patterns suggest that the RS measure may serve as a reasonable proxy for the task constructs posited by the model.

2.4 Predicting the Growth of Service Employment

A primary implication of our conceptual model is that commuting zones that are initially specialized in routine task activity will experience differential growth of service employment as routine tasks are supplanted by computerization. The scatter plot in the upper panel of Figure 3, which depicts the bivariate relationship between commuting zone Routine Share (RS) in 1980 and the change in the share of non-college labor input in service occupations over the subsequent 25 years, provides strong initial support for this prediction. Each plotted point in this figure represents one of 722 commuting zones, while the regression line corresponds to the following weighted OLS regression of the change in the service employment share on the initial RS, where weights are equal to commuting zone shares of national population in 1980:

$$
\begin{array}{ccccccc}
\Delta SVC_{j,1980-2005} & = & -0.039 & + & 0.323 \times RS_{j,1980} & + & e_{jt} \\
& & & & (t = 18.1) & & R^2 = 0.31
\end{array}
\qquad (2.4.1)
$$

The explanatory power of this bivariate relationship is substantial. The coefficient of 0.323 on the RS measure implies that a commuting zone with the mean Routine Share in 1980 is predicted to increase its share of non-college labor in service employment by 6.9 percentage points between 1980 and 2005.[30] Given an 85th/15th percentile range of the RS variable of approximately 0.10, the model predicts that the 85th percentile commuting zone increased its non-college service share by 3.2 percentage points more than the 15th percentile commuting zone.

To provide a more concrete sense of the geography underlying this pattern, the lower panel of Figure 3 plots the relationship between initial routine share and the growth of service employment over 1980 through 2005 for the 40 commuting zones in the sample with populations over 1 million which are identified by the names of their largest city in the figure. The bivariate relationship found in the full sample of 722 commuting zones remains robust and of comparable magnitude in this vastly reduced sample. The city names appearing next to each plotted point also highlight an important attribute of cities that are initially intensive in routine employment: these are not for the most part fading industrial cities; rather, they tend to be relatively intensive in finance, technology, government, and education.[31] The reason that these high-skill industries are associated with high levels of routine-intensive employment is that they employ large fractions of workers in supporting occupations, such as clerical work and accounting, which are themselves highly routine-intensive.

Table 4 explores the simple bivariate relationship between the routine employment share and growth of service employment over five and one-half decades (1950 to 2005) using specifications of the following form:

$$
\Delta SVC_{js\tau} = \alpha_t + \beta_\tau \times RS_{jst} + \gamma_s + e_{js\tau}. \qquad (2.4.2)
$$

In this equation, τ represents a decadal change, t denotes the start year of the corresponding decade τ, and s denotes the state in which the commuting zone is located.[32] The inclusion of a vector of state dummies, γ, means that the coefficient of interest, β, is identified by within-state cross-CZ variation.[33] A striking pattern that emerges from this

[30] $\Delta SVC = -0.039 + 0.323 \times 0.333 = 0.069$

[31] The three cities with the highest concentration of routine employment in 1980 are New York, San Francisco, and Washington, DC. The three cities at the bottom of the ranking are Pittsburgh, New Orleans, and San Antonio.

[32] The dependent variable for 1950 to 1970 is divided by two and the one for 2000 to 2005 is multiplied by two to place them on the same decadal time scale.

[33] If a commuting zone contains adjacent counties that cross state boundaries, we implicitly redefine

Figure 3a

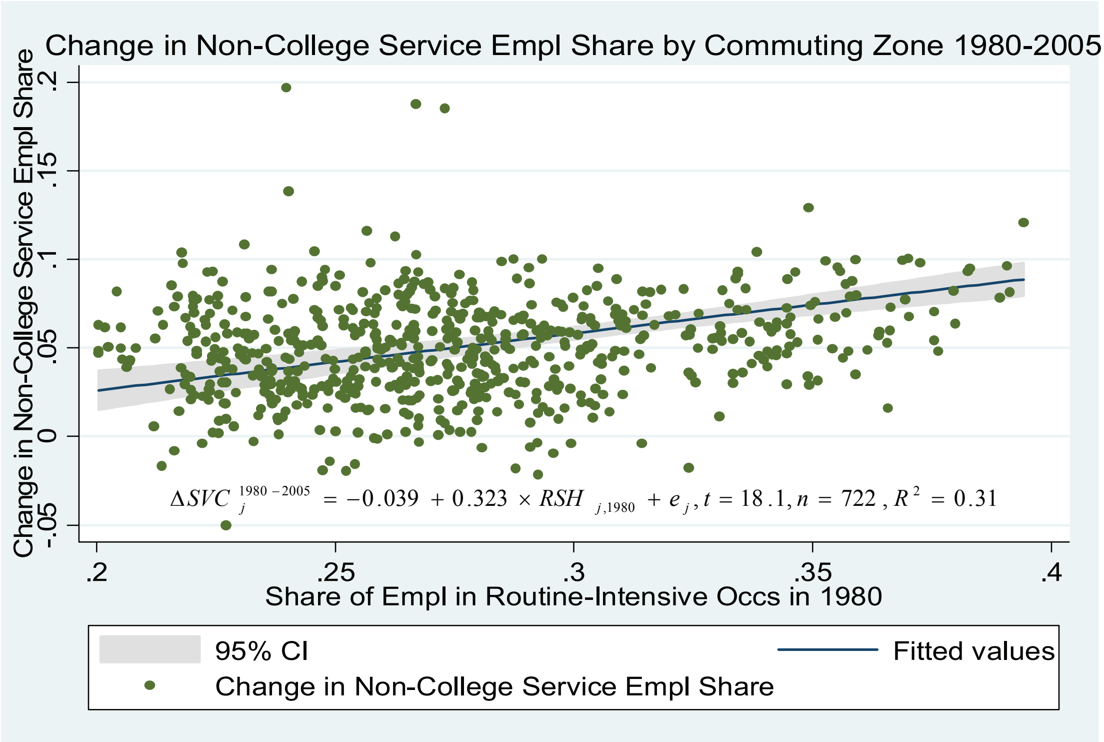


Figure 3b

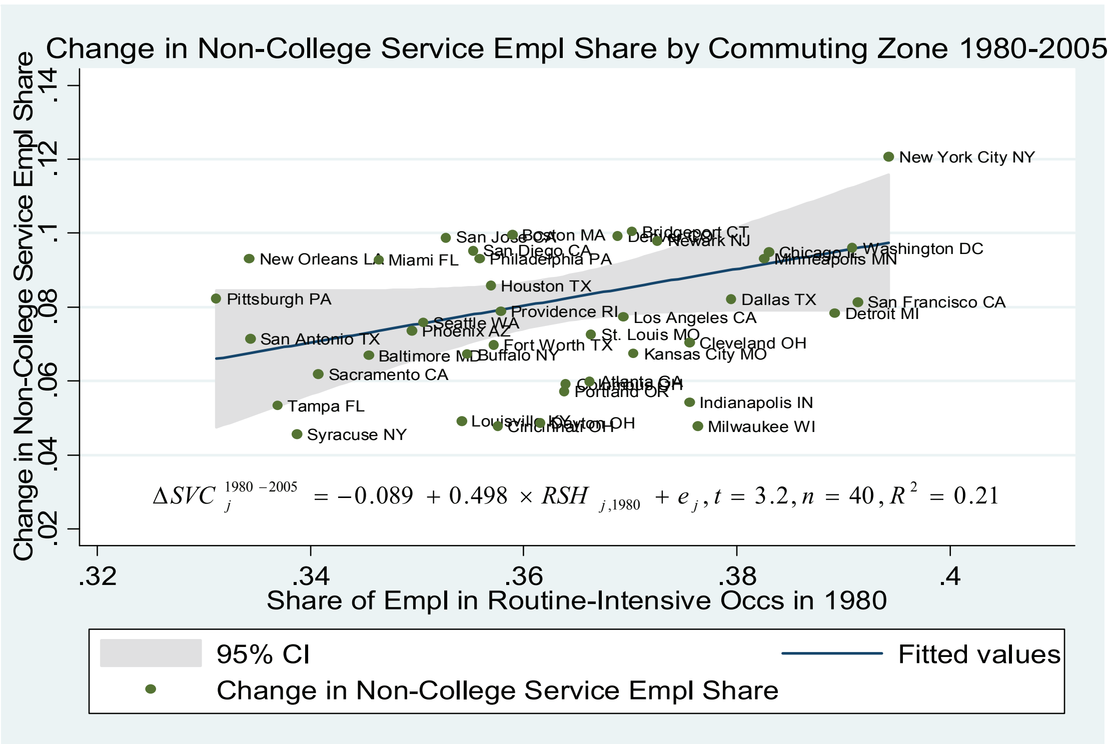


Table 4. Routine Employment Share and Growth of Service Employment within Commuting Zones, 1950 - 2005.
Dependent Variable: 10 × Annual Change in Share of Non-College Employment in Service Occupations

	1950 - 1970		1970 - 1980		1980 - 1990		1990 - 2000		2000 - 2005	
Share of Routine Occs.$_{-1}$	-0.122	**	0.032		0.082	**	0.084	*	0.321	**
	(0.020)		(0.034)		(0.024)		(0.037)		(0.087)	
Constant	0.022	**	-0.032	**	-0.014	*	-0.003		-0.042	
	(0.003)		(0.009)		(0.007)		(0.011)		(0.027)	
State dummies	Yes		Yes		Yes		Yes		Yes	
R^2	0.495		0.435		0.528		0.596		0.334	

N= 722 commuting zones. Robust standard errors in parentheses are clustered on state. Models are weighted by start of period commuting zone share of national population. Routine occupations are defined the occupations with largest routine task / manual task ratios that account for one third of overall employment in 1980. ~ $p \leq 0.10$, * $p \leq 0.05$, ** $p \leq 0.01$.

table is that the strong, positive predictive relationship between the routine employment share and growth of service employment is not detected prior to the decade of the 1980s, and actually has the opposite sign in the 1950 to 1970 period.[34] Beginning in 1980, this relationship becomes positive and significant, and its magnitude rises in each subsequent time interval.

2.4.1 Controlling for Skill Supply, Labor Market Conditions, and Demographics

We next explore a host of explanatory factors that may potentially determine geographic variation in the growth of service employment using an augmented version of equation (2.4.2). In particular, we estimate stacked first-difference models of the form

$$\Delta SVC_{js\tau} = \alpha + \beta_1 \times RS_{jst} + \beta_2 \times RS_{jst} \times I\left[t \geq 1980\right] + \beta_3 \Delta X_{js\tau} + \delta_\tau + \gamma_s + e_{js\tau}, \quad (2.4.3)$$

state boundaries so that the commuting zone is located in the state contributing a larger share of its population.

[34]One speculative explanation for the negative relationship between RS and the growth of service employment is based on the observation that US farm employment contracted rapidly in these two decades, falling from 11 to 3 percent of employment. Logically, farm-intensive commuting zones had low levels of the RS in 1950. The movement of labor from farming into services in these CZs may potentially explain the negative relationship between the RS and growth of service employment in this period.

where the sample includes each decadal change from Table 4 over the period 1950 to 2005, and we include a full set of time period effects, state effects, and measures of contemporaneous changes in a number of relevant human capital, labor market, and demographic variables.

The first column of the Table 5a pools all five and one-half decades of data to estimate the RS-service employment slope over the full period. Consistent with the results in Table 4, the strong, positive relationship between routine employment share and growth of service employment is non-existent prior to the 1980s. Column 2 shows that this finding is not sensitive to the inclusion of the state dummy variables, which function as state-specific trends in the first-differenced specification.

Subsequent columns of Table 5a sequentially control for a number of key factors that may contribute to growth of service employment within CZ's. Column 3 adds two variables intended to capture shifts in the demand and supply of services: the change in the college-educated share of the population and the change in the share of the population that is non-college immigrants. These controls enter with the expected sign: a rise in the highly-educated population or an increase in immigrant penetration predicts growth in service employment among non-college workers (Cortes, 2006).

Column 4 adds two variable that measure local labor market conditions: the change in the local unemployment rate and the change in the share of non-college employment in manufacturing. The service employment share rises significantly with the unemployment rate and also increases when manufacturing employment falls. This evidence suggests that service employment is less cyclical than non-service employment and that workers may choose service occupations when higher paying work is unavailable.

Column 5 considers two additional variables that may shift demand for service work: the employment to population rate of females and the population share of seniors (age 65+). If services substitute for household production, a rise in female labor supply may increase service demand (as well as potentially increase labor supply to service occupations). Contrary to expectations, increased female employment is associated with a lower growth of service employment.[35] A growing share of senior citizens in the population—who may have relatively high demand for services—is predictive of growth in service employment.

Notably, inclusion of these explanatory variables leaves the significant, positive relationship between the routine employment share and growth of service employment largely unaffected. When all explanatory variables are simultaneously included (column 6), the point estimate on the RS falls by about 45 percent, but the precision of the point estimate rises.

It also bears note that the Table 5a specifications likely 'over-control' for alternative causal factors, since many of these explanatory variables—immigration, unemployment,

[35]But this relationship flips sign when we condition on other explanatory variables, particularly the unemployment rate (see column 6).

Table 5a. Routine Employment Share and Growth of Service Employment within Commuting Zones, 1950 - 2005: Stacked First Differences.
Dependent Variable: 10 × Annual Change in Share of Non-College Employment in Service Occupations

	1950 - 2005											
	(1)		(2)		(3)		(4)		(5)		(6)	
Share of Routine Occs.$_{-1}$ × 1980-05	0.179	**	0.166	**	0.100	**	0.147	**	0.160	**	0.090	**
	(0.026)		(0.025)		(0.024)		(0.023)		(0.026)		(0.019)	
Share of Routine Occs.$_{-1}$	-0.024	*	-0.051	**	-0.054	**	-0.059	**	-0.055	**	-0.054	**
	(0.011)		(0.014)		(0.014)		(0.015)		(0.015)		(0.013)	
Δ College/Non-college pop					0.013	*					0.018	**
					(0.006)						(0.006)	
Δ Immigr/Non-college pop					0.105	**					0.107	**
					(0.037)						(0.037)	
Δ Manufact/empl							-0.070	**			-0.092	**
							(0.026)				(0.022)	
Δ Unempl rate							0.321	**			0.393	**
							(0.045)				(0.047)	
Δ Female empl/pop									-0.063	**	0.096	**
									(0.021)		(0.021)	
Δ Age 65+/pop									0.056		0.174	**
									(0.054)		(0.051)	
1970-1980 dummy	-0.008	**	-0.006	*	-0.011	**	-0.015	**	-0.004		-0.027	**
	(0.003)		(0.003)		(0.002)		(0.003)		(0.004)		(0.002)	
1980-1990 dummy	-0.037	**	-0.030	**	-0.016	*	-0.027	**	-0.026	**	-0.023	**
	(0.008)		(0.007)		(0.008)		(0.006)		(0.008)		(0.006)	
1990-2000 dummy	-0.039	**	-0.031	**	-0.019	**	-0.026	**	-0.032	**	-0.011	~
	(0.007)		(0.007)		(0.007)		(0.005)		(0.007)		(0.006)	
2000-2005 dummy	-0.026	**	-0.019	*	-0.003		-0.024	**	-0.020	*	-0.006	
	(0.008)		(0.008)		(0.008)		(0.006)		(0.008)		(0.006)	
Constant	0.012	**	0.014	**	0.016	**	0.016	**	0.018	**	0.009	*
	(0.003)		(0.004)		(0.003)		(0.004)		(0.004)		(0.004)	
State dummies	No		Yes		Yes		Yes		Yes		Yes	
R^2	0.368		0.397		0.414		0.450		0.403		0.474	

N=3610 (5 time periods x 722 commuting zones). Robust standard errors in parentheses are clustered on state. Models are weighted by start of period commuting zone share of national population. Routine occupations are defined the occupations with largest routine task / manual task ratios that account for one third of overall employment in 1980. ~ $p \le 0.10$, * $p \le 0.05$, ** $p \le 0.01$.

and falling manufacturing employment—may stem (in part) from a common cause: labor demand shifts against routine-intensive occupations. Indeed, when the Table 5a models are re-estimated in Table 5b using as controls start-of-period *levels* of the six additional explanatory variables rather than contemporaneous changes, the size and significance of the routine share measure in predicting growth of service employment is only slightly affected by their inclusion.[36] In net, initial employment concentration in routine-intensive occupations is a far stronger predictor of growth in service employment than any other human capital, labor market, or demographic variable that we have explored.[37]

2.4.2 Which Service Occupations and Which Workers?

We now explore whether the robust, geographic link between routine task-intensity and growth of service employment is pervasive among service employment categories and among demographic sub-groups of non-college workers. Estimates of equation (2.4.3) fit separately for each major service occupation group (Table 6) reveal that the aggregate relationship between the routine employment share and subsequent growth of service employment is driven by a broad set of service occupations, including food service, personal appearance, child care, recreation, and building cleaning and gardening.[38] In fact, point estimates are positive for all nine service occupation categories for 1980-2005 period, and are statistically significant in six of them. Notably, while healthcare support occupations are the third largest contributor to service employment growth over 1980 to 2005 (after food service and janitorial services), their growth is not strongly predicted by routine task-intensity. Plausibly, rising demand for healthcare support services derives from other sources, particularly the aging of the US population.

Complementing these results for occupations, Table 7 estimates equation (2.4.3) for four demographic sub-groups of non-college workers. The relationship between the RS and rising service employment is largest for foreign borns but also positive for US-born workers and for workers of both genders. Females are the only demographic group for whom the relationship between RS and growth in service employment was already positive prior to 1980, and though that relationship accelerated notably after 1980, the respective

[36]We nevertheless report the contemporaneous change specification in Table 5a to demonstrate robustness.

[37]This result is particularly noteworthy given the strong correlations between the RS and many of the explanatory variables. In a multivariate cross-sectional regression for 1980 shown in Appendix Table 3, the RS is rising in the relative supply of college-educated residents, the immigrant share of the non-college population, and the female employment rate. It is falling in the CZ unemployment rate and the elderly share of population.

[38]The 1950 to 1980 comparisons of detailed service occupation employment are somewhat unreliable because the 1950 Census classifies many service workers in broad "not elsewhere classified" categories. This gives rise to large spurious increases in many subcategories of service employment over 1950 to 1980, balanced by an offsetting drop in "miscellaneous service occupations." This consistency issue affects comparisons at this very disaggregated level only, and does not contaminate the overall measure of service occupation employment.

Table 5b. Routine Employment Share and Growth of Service Employment within Commuting Zones, 1950 - 2005: Stacked First Differences.
Dependent Variable: 10 × Annual Change in Share of Non-College Employment in Service Occupations

	1950 - 2005					
	(1)	(2)	(3)	(4)	(5)	(6)
Share of Routine Occs.$_{-1}$ × 1980-05	0.179 **	0.166 **	0.130 **	0.132 **	0.176 **	0.108 **
	(0.026)	(0.025)	(0.023)	(0.022)	(0.019)	(0.014)
Share of Routine Occs.$_{-1}$	-0.024 *	-0.051 **	-0.062 **	-0.039 **	0.003	0.034
	(0.011)	(0.014)	(0.019)	(0.013)	(0.019)	(0.022)
College/Non-college pop$_{-1}$			0.007 ~			0.010 *
			(0.004)			(0.004)
Immigr/Non-college pop$_{-1}$			0.026 ~			0.009
			(0.014)			(0.015)
Manufact/empl$_{-1}$				-0.031 **		-0.011
				(0.010)		(0.008)
Unempl rate$_{-1}$				-0.100 *		-0.281 **
				(0.042)		(0.069)
Female empl/pop$_{-1}$					-0.098 **	-0.163 **
					(0.027)	(0.029)
Age 65+/pop$_{-1}$					-0.045	-0.026
					(0.030)	(0.023)
1970-1980 dummy	-0.008 **	-0.006 *	-0.005	-0.005 ~	0.005	0.011 *
	(0.003)	(0.003)	(0.004)	(0.003)	(0.005)	(0.005)
1980-1990 dummy	-0.037 **	-0.030 **	-0.018 **	-0.016 *	-0.014 *	0.024 **
	(0.008)	(0.007)	(0.006)	(0.007)	(0.007)	(0.009)
1990-2000 dummy	-0.039 **	-0.031 **	-0.024 **	-0.02 **	-0.006	0.032 **
	(0.007)	(0.007)	(0.005)	(0.007)	(0.008)	(0.011)
2000-2005 dummy	-0.026 **	-0.019 *	-0.015 *	-0.008	0.008	0.042 **
	(0.008)	(0.008)	(0.007)	(0.007)	(0.010)	(0.011)
Constant	0.012 **	0.014 **	0.014 **	0.022 **	0.034 **	0.056 **
	(0.003)	(0.004)	(0.004)	(0.004)	(0.007)	(0.008)
State dummies	No	Yes	Yes	Yes	Yes	Yes
R^2	0.368	0.397	0.404	0.405	0.417	0.442

N=3610 (5 time periods x 722 commuting zones). Robust standard errors in parentheses are clustered on state. Models are weighted by start of period commuting zone share of national population. Routine occupations are defined the occupations with largest routine task / manual task ratios that account for one third of overall employment in 1980. ~ $p \leq 0.10$, * $p \leq 0.05$, ** $p \leq 0.01$.

Table 6. Routine Employment Share and Growth of Employment in Detailled Service Occupations within Commuting Zones, 1950 - 2005: Stacked First Differences.

Dependent Variable: 10 × Annual Change in Share of Non-College Employment in Specific Service Occupation

	Food Service	Building Clean/ Garden	Health Support	House Clean/ Laundry	Child Care	Personal Appear-ance	Security Guards	Recreat-ion	Misc Personal Svcs
	A. Regression Analysis								
Share of Routine Occs._{-1} × 1980-05	0.059 **	0.036 **	0.006	0.011	0.015 *	0.022 **	0.005 ~	0.012 **	0.001
	(0.012)	(0.007)	(0.011)	(0.010)	(0.006)	(0.003)	(0.003)	(0.004)	(0.009)
Share of Routine Occs._{-1}	-0.010 *	0.004	-0.007 ~	-0.006	-0.007 *	-0.006 **	0.004 *	-0.008 **	-0.015 *
	(0.005)	(0.004)	(0.004)	(0.006)	(0.003)	(0.002)	(0.002)	(0.002)	(0.007)
Constant	0.009 **	0.004 **	0.006 **	0.003 **	0.003	0.001	-0.001	-0.001	-0.010 **
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)
State dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.110	0.256	0.117	0.470	0.131	0.157	0.097	0.264	0.585
	B. Share in Total Non-College Employment								
Empl. share 1980	4.18%	3.11%	1.88%	1.41%	0.51%	0.75%	0.63%	0.15%	0.31%
Empl. share 2005	6.55%	4.69%	3.04%	1.86%	1.00%	0.94%	0.88%	0.43%	0.44%
Change 1980-2005	2.37%	1.58%	1.15%	0.44%	0.49%	0.19%	0.25%	0.28%	0.13%

N=3610 (5 time periods x 722 commuting zones). Robust standard errors in parentheses are clustered on state. Models are weighted by start of period commuting zone share of national population. Routine occupations are defined the occupations with largest routine task / manual task ratios that account for one third of overall employment in 1980. ~ p ≤ 0.10, * p ≤ 0.05, ** p ≤ 0.01.

coefficient is imprecisely estimated and not significant.

2.4.3 Changes in Service Employment in Matched Commuting Zones

While our main findings are highly robust to inclusion of various controls and to numerous cuts of the data, it remains the case that the results derive from a comparison of commuting zones that differ initially on numerous observable dimensions. As a further means to maximize the validity of comparisons across commuting zones, we matched commuting zones on their *expected* routine employment share and then use only unexpected (residual) variation in the routine share to predict growth in service employment. To implement this test, we first regress the routine share measure in 1980 on all of covariates used above:

$$RS_{j,1980} = \alpha + \beta \times X_{j,1980} + \gamma_s + e_{j,1980}. \tag{2.4.4}$$

The explanatory power of this model is quite high, as shown in Appendix Table 3, yielding an R-squared of 0.82. One might legitimately suspect that the unexplained variation remaining in RS would not therefore be predictive of subsequent changes in service occupation employment.

We test this proposition by grouping the 722 commuting zones into evenly sized terciles based on their *predicted* routine share in 1980, $\widehat{RS}_{j,1980}$, from the model above. We then

Table 7. Routine Employment Share and Growth of Service Employment within Commuting Zones, 1950 - 2005.
Dependent Variable: 10 × Annual Change in Share of Population Group's Non-College Employment in Service Occupations

	Males		Females		US Borns		Foreign Borns	
Share of Routine Occs.$_{-1}$ × 1980-05	0.129	**	0.084		0.046	*	0.323	**
	(0.028)		(0.070)		(0.022)		(0.082)	
Share of Routine Occs.$_{-1}$	-0.017		0.097	**	-0.043	*	-0.100	~
	(0.014)		(0.034)		(0.017)		(0.059)	
Constant	0.005	~	-0.049	**	0.015	**	0.024	~
	(0.003)		(0.009)		(0.003)		(0.013)	
Control variables	Yes		Yes		Yes		Yes	
State dummies	Yes		Yes		Yes		Yes	
R^2	0.238		0.514		0.284		0.052	

N=3610 (5 time periods x 722 commuting zones). Robust standard errors in parentheses are clustered on state. Models are weighted by start of period commuting zone share of national population. Routine occupations are defined the occupations with largest routine task / manual task ratios that account for one third of overall employment in 1980. ~ p ≤ 0.10, * p ≤ 0.05, ** p ≤ 0.01.

regress commuting zone level changes in service employment over 1980 through 2005 within each tercile of predicted routine-share on the unexplained component of routine-employment, $\widetilde{RS}_{j,1980} = RS_{j,1980} - \widehat{RS}_{j,1980}$:

$$\Delta SVC_{j\tau} = \alpha_\tau + \lambda \times \widetilde{RS}_{j,1980} + \omega_{j\tau}. \tag{2.4.5}$$

Estimates in Table 8 demonstrate a remarkably consistent pattern: the *unexplained* component of routine employment specialization in 1980 is strongly predictive of subsequent growth in service employment within commuting zones. Point estimates for λ are comparable in magnitude across all three terciles, and are significant in one of the three subsamples despite the smaller sample sizes. These results again underscore that initial occupational specialization predicts subsequent changes in labor market outcomes that are not otherwise predicted by measures of skill supply, industrial structure, or demographic composition.

2.4.4 Testing Alternative Demand-Side Explanations

Growing employment in low-skilled service jobs could also arise from two other demand-side forces noted in the Introduction: rising high incomes, which may generate additional demand for outputs of service occupations if they produce luxury goods; and rising mar-

Table 8. Growth of Service Employment by Tercile of Predicted Routine Share, 1980 - 2005.
Dependent Variable: 10 × Annual Change in Share of Non-College Employment in Service Occupations

	C'zones by Predicted 1980 Routine Share			
	All	Top 1/3	Middle 1/3	Bottom 1/3
Residual Share of Routine Occs.$_{-1}$	0.122 *	0.100	0.119 *	0.118
	(0.054)	(0.068)	(0.051)	(0.083)
N	2166	723	720	723
R^2	0.061	0.089	0.028	0.021

Predicted routine share is based on the predicted values from a regression of routine share on college/non-college population, share of immigrants among non-college population, manufacturing share, unemployment rate, female labor force participation, population share above age 65, and state dummies. Residual routine share equals the residuals of that regression. Robust standard errors in parentheses are clustered on state. Models are weighted by start of period commuting zone share of national population and contain a constant and time dummies. ~ $p \leq 0.10$, * $p \leq 0.05$, ** $p \leq 0.01$.

ket returns to skill, which may spur high skill workers to increase their labor supply and purchase additional household services to compensate for foregone household production.

We explore the evidence for these income and substitution effects by augmenting the baseline regression model for 1980 through 2005 with two additional measures. To proxy for wage structure shifts that may generate income effects, we control for changes in the 90th percentile of the log weekly wage distribution among full-time, full-year workers in each commuting zone.[39] To measure the evolving market participation of high-skilled workers, we control for changes in mean annual hours worked by college-graduates in each commuting zone. These outcomes are, as expected, strongly predicted by the initial routine-share variable, $RS_{j,1980}$. High wages and high-skilled hours rose by significantly more over 1980 through 2005 in commuting zones that were initially specialized in routine-intensive employment.[40]

Somewhat surprisingly, estimates in Table 9 find that these proxies for income and substitution effects are not strong correlates of changes in service employment. Growth in the 90th percentile weekly wage is only weakly correlated with rising service employment, and this relationship turns negative when the RS variable is added. Annual hours

[39] This is arguably the best proxy of the market price of 'high' skills available from the Census data since it captures the wage commanded by workers with strong labor force attachment. Other wage measures yield similar results but have lower explanatory power.

[40] A table of results is available from the authors.

Table 9. Predicting Changes in Service Occupation Employment using Proxies for Income and Substitution Effects. Dependent Variables: 10 × Annual Change in Share of Non-College Employment in Service Occupations, 1980-2005

	(1)		(2)		(3)		(4)		(5)		(6)	
Δ ln(P90) Weekly Wage	0.013		-0.012									
	(0.015)		(0.015)									
Δ Avg Annual Hours per Coll Grad ÷ 2080					-0.111	**	-0.120	**				
					(0.032)		(0.032)					
Δ Avg Annual Hours per Male Coll Grad ÷ 2080									-0.064	**		
									(0.018)			
Δ Avg Annual Hours per Female Coll Grad ÷ 2080											-0.081	**
											(0.024)	
Share of Routine Occs_{-1}			0.140	**			0.138	**	0.145	**	0.121	**
			(0.024)				(0.025)		(0.025)		(0.023)	
R^2	0.166		0.194		0.185		0.216		0.205		0.211	

N=2166 (3 time periods x 722 commuting zones). Robust standard errors in parentheses are clustered on state. Models are weighted by a worker's share in total labor supply in a given year. All models include an intercept, state dummies, and time dummies . ~ p ≤ 0.10, * p ≤ 0.05, ** p ≤ 0.01.

worked by college graduates are *negatively* related to service employment growth, and this pattern is true whether we measure college labor labor overall or separately by gender. Moreover, the routine share measure remains in all cases highly predictive of rising service employment, and is essentially unaffected by inclusion of these controls.

In summary, the results in Tables 4 through 9 provide robust support for a key prediction of the conceptual model: geographic areas that were specialized in routine-intensive occupations prior to the era of rapid computerization experienced significantly greater growth of service employment in the ensuing decades. This predictive relationship is pervasive across categories of service work, and affects employment trends among different groups of non-college workers, i.e., male and female, foreign and US-born. The link between the routine employment share and rising service employment does not take hold until the 1980s, and it accelerates in the two subsequent decades. Most notably, this simple measure of occupational structure appears to capture a significant dimension of local economic activity that is not well measured by a host of other labor supply, labor demand, and demographic proxies, including education, immigration, unemployment, female labor force participation, and population aging, as well as proxies for other demand-side forces.[41] Subsequent sections explore further predictions of the model, using the RS measure as a

[41] With sufficient degrees of freedom, it would clearly be feasible to construct a multivariate index of occupational structure that is more predictive of subsequent changes in commuting zone characteristics—and in particular the growth of service employment—over 1980 through 2005. A complete set of occupation dummies would, for example, absorb all variation in RS.

key predictive variable.

2.5 Task Specialization, Computer Adoption, and Wage Inequality

Our conceptual model makes four further predictions about the relationship between initial specialization in routine occupations and subsequent commuting zone level outcomes. First, displacement of routine labor input should lead to shifts in job specialization, as workers—particularly the less-educated—move out of routine-intensive occupations. Second, computer adoption should be more extensive in these regions, since higher routine task-intensity implies greater demand for computer capital. Third, changing task prices should spur rises in earnings inequality—particularly in the upper-half of the distribution—as the abstract task price rises relative to the routine task price. Finally, wages in service occupations may rise relative to other activities performed by less skilled workers in the same commuting zones if goods and services are complements in consumption. As noted in section (2.2), the 'price' implications of our model are less robust than the 'quantity' implications since they hinge on imperfect arbitrage on wage rates across commuting zones. Thus, the third and fourth implications appear less clear cut.

2.5.1 Task Specialization

Our conceptual framework implies that the differential rise in service employment evident in routine task-intensive regions is one manifestation of a general phenomenon of shifts in task specialization away from routine-intensive labor. We test this implication by estimating a variant of equation (2.4.3) in which the dependent variable is the *change* in the routine employment share within a commuting zone, both overall and within broad education categories. Table 10 shows that during the 1980 to 2005 period, commuting zones with high routine employment shares occupations saw larger declines in routine-intensive employment—a relationship that is robust to the full set of contemporaneous labor market and demographic controls used in prior models (column 2). In particular, the coefficient of 0.082 in column 1 indicates that the 80th percentile commuting zone experienced about 0.8 percentage points larger a fall in routine employment per decade than the 20th percentile commuting zone (a 2.0 percentage point differential over 25 years). Given an aggregate decline of 1.6 percentage points in employment shares in routine-intensive occupations in this period, this magnitude is sizable. Notably, there is a negative significant relationship between start of period RS and movements out of routine-intensive occupations even prior to the 1980s. But the magnitude of this relationship

Table 10. Changes in Share of Routine Occupations within Commuting Zones, 1950 - 2005.
Dependent Variable: 10 × Annual Change in Education Group's Share of Employment in Routine Occupations

	All		College		Non-College	
	(1)		(2)		(3)	
Share of Routine Occs.$_{-1}$ × 1980-05	-0.089	**	-0.009		-0.184	**
	(0.017)		(0.027)		(0.025)	
Share of Routine Occs.$_{-1}$	-0.174	**	-0.188	**	-0.130	**
	(0.015)		(0.023)		(0.016)	
State dummies	Yes		Yes		Yes	
R^2	0.733		0.316		0.654	

N=3610 (5 time periods x 722 commuting zones). Robust standard errors in parentheses are clustered on state. Models are weighted by start of period commuting zone share of national population. Routine occupations are defined the occupations with largest routine task / manual task ratios that account for one third of overall employment in 1980. ~ p ≤ 0.10, * p ≤ 0.05, ** p ≤ 0.01.

increases by more than 50 percent during the post-1980 period.

Subsequent panels of Table 10 examine this relationship separately for college and non-college workers. The decline in routine task-intensive employment for college workers in high RS commuting zones commences prior to the 1980s, and does not accelerate thereafter. By contrast, the differential rate of decline in routine-intensive employment among non-college workers more than doubles after 1980. Consonant with the conceptual model, the recent movement out of routine-intensive occupations is concentrated among less educated workers.[42]

2.5.2 Computer Adoption

The conceptual model unambiguously predicts that the decline in routine *labor* input within commuting zones should be accompanied by the adoption of information technology (which substitutes for routine labor)—and that this process should be more pronounced in areas initially specialized in routine occupations. We test this implication using a measure of geographic computer penetration developed by Doms and Lewis (2006). Based

[42]Michaels (2007) finds that clerical occupations demanded highly educated labor at the start of the twentieth century. But by the 1950s, these were no longer elite occupations. The results in Table 8 likely reflect the fact that the movement of highly-educated labor out of routine occupations was well underway before the computer era.

on private sector survey data on computer inventories, these data measure the number of personal computers per employee at the firm level. Doms and Lewis aggregate this variable to the level of Metropolitan Statistical Areas (MSAs) and purge it of industry by establishment-size effects using a linear regression model.[43] We use the Doms and Lewis 'adjusted computers-per-worker' measure for the years 1990 and 2002, which we match to commuting zones.[44] Following the approach of Doms, Dunne and Troske (1997), we treat the 1990 level of this variable as the 'change' from 1980 to 1990 (thus assuming that the level was close to zero in all areas in 1980). We measure the change in this variable over the subsequent decade using 5/6 of the 1990 to 2002 first-difference.[45]

We estimate models predicting computer adoption (PCs per worker) across commuting zones of the following form:

$$\Delta C_{js\tau} = \alpha + \beta_{1\tau} \times RS_{jst} + \beta_{2\tau}\Delta X_{js\tau} + \gamma_s + e_{js\tau}, \tag{2.5.1}$$

where the dependent variable is the Doms-Lewis measure of computer adoption over time interval τ in commuting zone j in state s, RS_{jst} is the start of period routine task index, and $X_{sj\tau}$ is a vector of contemporaneous controls. The first two columns of Table 11 present separate, by-decade OLS regressions of commuting zone computer adoption during the 1980s and 1990s on the RS measure, state dummies and a constant. The RS has substantial predictive power for computer adoption in both decades (with t-ratios well over 9). The implied difference in computer adoption between the 80th and 20th percentile commuting zone is larger than one full standard deviation of the computer adoption measure in each decade.

Subsequent columns of Table 11 probe the robustness of this relationship by regressing the stacked decadal changes in computer adoption on initial RS and the full set of contemporaneous labor force and demographic change variables used earlier. Surprisingly, all of these covariates are significant predictors of computer adoption in this time period. The RS measure is nevertheless highly robust to their inclusion; with these variables added, its magnitude drops by less than a third and the t-ratio remains above nine. Thus, even accounting for contemporaneous changes in key labor market and demographic variables, it is apparent that commuting zones that were initially specialized in routine occupations adopted computer technology at a differentially rapid rate over the subsequent two decades—presumably to substitute physical capital for human capital in performing routine tasks.

[43] The variable is not adjusted for the educational or occupational composition of MSAs.

[44] Approximately 50 of the 722 commuting zones do not have corresponding computer adoption data and so are dropped from the analysis.

[45] The level of the PC-per-worker measure is not readily interpretable because it is 'residualized,' as above. The cross commuting zone standard deviation of this variable is 0.048 for the 1980 to 1990 change and 0.053 for the 1990 to 2000 change.

Table 11: Routine Task Intensity and Computer Adoption 1980-2000.
Dependent Variable: 'Adjusted PCs per Employee' (Doms and Lewis 2006)

	1980 - 1990		1990 - 2000		1980 - 2000			
	(1)		(2)		(3)		(4)	
Share of Routine Occs.$_{-1}$	0.711	**	0.529	**	0.662	**	0.451	**
	(0.051)		(0.058)		(0.035)		(0.051)	
Δ College/Non-college pop							0.139	**
							(0.014)	
Δ Immigr/Non-college pop							0.257	**
							(0.034)	
Δ Manufact/empl							0.187	**
							(0.046)	
Δ Unempl rate							0.296	*
							(0.129)	
Δ Female empl/pop							0.195	*
							(0.090)	
Δ Age 65+/pop							0.410	**
							(0.146)	
1990-2000 dummy					0.025	**	0.054	**
					(0.007)		(0.006)	
Constant	-0.293	**	-0.190	**	-0.267	**	-0.261	**
	(0.015)		(0.017)		(0.010)		(0.016)	
State dummies	Yes		Yes		Yes		Yes	
R^2	0.653		0.375		0.455		0.535	
N	675		660		1335		1335	

The Doms-Lewis measure of computer adoption reflects the number of personal computers per employee, controlling for 950 industry/establishment interactions. Data for computer adoption in commuting zones is available to us for the years 1990 and 2002; we assume zero computers per worker in 1980 and use 5/6 of the change in computer adoption between 1990 and 2002 as our measure for computer adoption during the 1990s. ~ p ≤ 0.10, * p ≤ 0.05, ** p ≤ 0.01.

2.5.3 Wage Inequality

We finally explore the relationship between task specialization and wage inequality. The analysis first examines the relationship between a commuting zones's routine share and the evolution of aggregate wage inequality—in particular, earnings polarization—as measured by the 90/50 and 50/10 log wage ratios. We next turn to microdata to provide a tightly controlled analysis of changes in wage structure between occupations within commuting zones, holding constant all observable determinants of earnings.

Aggregate Wage Structure

We estimate stacked first-difference regressions for changes in wage inequality within commuting zones, as measured by the 90/10, 90/50 or 50/10 log weekly wage ratio for full-time, full-year workers. Following the format of earlier equations, all models include the start-of-period RS, and a full set of state and time dummies, with alternate specifications containing the full set of labor market and demographic controls used above.

These estimates in Table 12 reveal a striking pattern: commuting zones with a greater routine employment share in 1980 saw a large, differential *polarization* of earnings in the subsequent 25 years. In particular, upper-tail (90/50) inequality rose and lower-tail (50/10) inequality fell in high-RS regions during the 1980-2005 period relative to earlier trends. These relationships are economically large.[46] They are either substantially smaller or of opposite sign in the prior three decades. Thus, the wage polarization seen in economy-wide data for this period is replicated in commuting zones experiencing rapid displacement of routine work.

Evidence From Microdata

Do these patterns of aggregate wage structure change affecting commuting zones specialized in routine employment primarily reflect *compositional* shifts in worker characteristics and occupational characteristics—or, instead, changes in the wage paid to given worker characteristics within a geographic area? To develop a more precise answer to this question, we next turn to microdata on earnings.

Pooling microdata on real log hourly wages from the 1980 Census and the 2005 American Community Survey, we run a set of standard log wage equations augmented with time dummies, commuting-zone dummies, a full set of person-level covariates interacted with time dummies, and an interaction between the start-of-period routine employment share and the 2005 dummy. These models are estimated separately for each of the six major occupation categories discussed in the Introduction (ranging from Professional to Service, see Table 1). In particular, we estimate by OLS:

$$\begin{aligned} \ln w_{ijkt} &= \alpha_k + \beta_{1k} RS_{j,t-1} + \beta_{2k} \left\{ RS_{j,t-1} \times I\left[t \geq 1980\right] \right\} \\ &\quad + X'_{ijkt}\beta_{kt} + \delta_{tk} + \gamma_{jk} + e_{ijkt}, \end{aligned} \tag{2.5.2}$$

where i denotes workers, j denotes commuting zones, t denotes times (1980, 2005) and k denotes occupation. To account for the fact that the main predictive variable, RS, varies only on the commuting zone level and, moreover, that wage levels are not independent

[46]To benchmark magnitudes, note that the predicted differential rise (fall) in the 90/50 (50/10) wage differential in the 75th relative to the 25th percentile commuting zone (ranked by 1980 RS) is 3.0 (-4.3) log points over 1980 through 2005. The contemporaneous weighted mean within-CZ rise in 90/50 (50/10) inequality in this period is 12.2 (7.8) log points.

Table 12. Routine Employment Share and Changes in Wage Inequality, 1950 - 2005.
Dependent Variable: 10 × Annual Change in Wage Inequality Measure

	P90/50		P50/10		P90/10	
	A. Avg Changes (10 × Annual Chg)					
Years 1980-2005	0.049		0.031		0.080	
	(0.059)		(0.072)		(0.092)	
Years 1950-1980	0.039		-0.003		0.036	
	(0.053)		(0.090)		(0.107)	
	B. Regression Analysis					
Share of Routine Occs.$_{-1}$ × 1980-05	0.299	**	-0.426	**	-0.127	
	(0.063)		(0.115)		(0.121)	
Share of Routine Occs.$_{-1}$	0.075	*	0.615	**	0.691	**
	(0.028)		(0.068)		(0.078)	
State dummies	Yes		Yes		Yes	
R^2	0.182		0.277		0.309	

N=3610 (5 time periods x 722 commuting zones). Robust standard errors in parentheses are clustered on state. Models are weighted by start of period commuting zone share of national population. All models include an intercept, and state and period dummies. ~ p ≤ 0.10, * p ≤ 0.05, ** p ≤ 0.01.

among workers in nearby locations, the standard errors of these estimates are clustered at the state by year level. Results are displayed in Table 13.

The first two columns of Table 13 show that commuting zones with a higher routine employment share in 1980 saw large, real wage increases among workers in the two occupation groups with highest education levels between 1980 and 2005. A 10 percentage point larger routine share in 1980 predicts 6.2 log points greater wage growth in managerial and professional occupations and 9.7 log points greater wage growth in technical, sales and administrative occupations (both for males) over these two decades. Effects for females are somewhat larger in professional occupations and smaller in administrative occupations.

Columns 3 and 4 estimate analogous wage models for workers in production and operative occupations—many of them corresponding to middle-skill ('routine') occupations in our conceptual framework. Opposite to the pattern for highly-skilled occupations, a higher routine share of employment in 1980 predicts significant real wage declines in these occupations: a 10 percentage point routine share in 1980 predicts 3.8 to 6.1 log point

Table 13. Routine Employment Share and Wage Changes by Major Occupation Groups, 1980-2005.
Dependent Variable: Log Real Hourly Wage.
Microdata Estimates using Pooled 1980/2005 Census and ACS Samples

	Manager / Prof'nl (1)		Tech / Sales / Admin (2)		Produc-tion (3)		Opera-tives (4)		Service Occs (5)	Service vs. Production, Operatives (6)	
							A. Males				
							C'zone dummies, w/o Person-Level Controls				
Share of Routine Occs.$_{-1}$ x 2005	0.620	**	0.965	**	-0.380	*	-0.605	**	-0.251	0.303	~
	(0.141)		(0.200)		(0.171)		(0.167)		(0.193)	(0.170)	
							C'zone dummies, with Person-Level Controls				
Share of Routine Occs.$_{-1}$ x 2005	0.646	**	0.551	**	0.117		-0.303		0.076	0.332	*
	(0.143)		(0.182)		(0.185)		(0.188)		(0.205)	(0.163)	
n	987,065		858,177		1,038,589		1,304,826		528,027	2,871,442	
							B. Females				
							C'zone dummies, w/o Person-Level Controls				
Share of Routine Occs.$_{-1}$ x 2005	1.004	**	0.805	**	0.009		-0.612	**	-0.138	0.761	**
	(0.138)		(0.140)		(0.301)		(0.197)		(0.178)	(0.228)	
							C'zone dummies, with Person-Level Controls				
Share of Routine Occs.$_{-1}$ x 2005	0.961	**	0.772	**	0.215		-0.146		0.089	0.376	*
	(0.140)		(0.127)		(0.254)		(0.185)		(0.161)	(0.158)	
n	926,169		1,831,136		94,843		528,532		840,657	1,461,565	

Note: Column (6) pools production workers, operatives, and service workers; it reports the coefficient of an interaction term between share of routine occupations in 1980 and a dummy for service workers. Robust standard errors in parentheses are clustered on commuting zones. Models are weighted by a worker's share in total labor supply in a given year. Each cell corresonds to a separate OLS regression. All models include an intercept, and a time dummy for the second period, and commuting zone dummies. Models with person-level controls also include nine dummies for years of education, a quartic in potential experience, dummies for married, non-white and foreign-born, and interactions of all individual level controls with the time dummy. Hourly wages are defined as yearly wage and salary income diveded by the product of weeks worked times usual weekly hours. ~ $p \leq 0.10$, * $p \leq 0.05$, ** $p \leq 0.01$.

declines in wages.[47]

Column 5 presents wage estimates for workers in service occupations.[48] Distinct from other low-education occupations (i.e., production workers and operatives), the relationship between initial routine task share and service wages is small in magnitude and not significant. When wages in service occupations are directly compared to those in produc-

[47]While the precision of the point estimate for wages of females in production occupations is low, the table also makes evident that there are only 10 percent as many females as males in production occupations, whereas there are 40 percent as many females as males in operative occupations.

[48]We do not include wage estimates for farm occupations since a large share of farm labor is undocumented and hence we expect that earnings self-reports in the Census are unreliable.

tive and operative positions in column 6, relative wage growth in service occupations is found to be greater in routine-occupation intensive commuting zones.

The second row of each panel of Table 13 re-estimates these models, augmented with a full set of person-level demographic controls including nine dummies for years of education, a quartic in potential experience, and dummies for married, non-white and foreign-born. These covariates are further interacted with time dummies to allow their slopes to differ by period. The pattern of results is only modestly affected by the inclusion of these additional variables. Estimates for high-skilled occupations are essentially unaffected. Estimates for middle-skilled occupations become less negative or even positive, indicating that part of the negative wage relationship is due to adverse changes in skill composition in these occupations in initially routine-intensive commuting zones. Finally, the estimates for service occupation wages become positive for both males and females when these controls are added, suggesting that compositional shifts may mask *rising* real wages in these occupations.[49]

Reinforcing the earlier results for 90/50 and 50/10 wage inequality, these microdata estimates confirm that commuting zones that were previously specialized in routine jobs saw a distinct pattern of polarizing wage growth among occupations over the subsequent 25 years, with strongly rising wages in high-skill occupations, declining wages in moderately-skilled occupations, and stable wages in low-skill service occupations. Thus, the data clearly support the prediction that displacement of routine tasks within commuting zones is accompanied by growth in both service employment and service wages. What makes this finding particularly compelling is that service occupations are the only low-skill job category that appear to benefit from this process.

2.6 Conclusions

While the past 25 years have seen declining or stagnating real (and relative) earnings and employment of less educated workers, employment in low-skill service occupations presents a striking exception. Between 1980 and 2005, the share of hours worked in service occupations among those with high school or lower education rose by more than 50 percent. Simultaneously, real hourly wages in service occupations increased by 20 log points, considerably exceeding wage growth in other low-skill occupations. In fact, the upward twist of the lower-tail of the U.S. earnings and job distributions that took form during the 1990s is substantially accounted for by rising employment and wages in service occupations.

We offer a hypothesis for the rising demand for service work based on changes in

[49] At a minimum, these results make it appear unlikely that the rising relative wages of service occupations relative to other low-education occupations seen in Table 11 is driven by selection of relatively skilled workers into service jobs.

task specialization induced in part by technical change. Our conceptual framework builds from the observation that the primary job tasks of service occupations are difficult to either automate or outsource since these tasks require interpersonal and environmental adaptability as well as direct physical proximity. Our conceptual model shows that if the demand for the outputs of service occupations does not admit close substitutes, the substitution of information technology for routine tasks used in goods production may, in the long run, lead to rising wages and employment in service occupations.

Motivated by the observation that workers in service occupations must collocate with demanders of their services, we study the determinants of employment and wages in services during 1950 through 2005 in 722 consistently defined commuting zones covering all of US mainland employment. The analysis contrasts the period 1980 to 2005 during which a rapid adoption of information technology took place with a previous period from 1950 to 1980. We use an empirical approach built on the theoretical model, which predicts that, if commuting zones differ initially in the share of employment in routine-intensive occupations, markets with higher routine shares will see larger increases in service occupation employment and greater polarization of earnings between high and middle-skill workers as time advances. If goods and services are sufficiently complementary, the model further implies that wages in service occupations will rise along with service employment.

We explore these predictions using a simple measure of initial specialization in routine-task-intensive employment based on the occupational structure of commuting zones at the start of the sample period. This measure proves strikingly predictive of the changes in task and wage structure implied by the model: reallocation of labor activity from routine tasks; employment growth in low-skilled service occupations; differential adoption of information technology; and polarization of earnings growth. Thus, the changes in task structure that we document accompany growth in wages at the tails of the distribution but not elsewhere.

These findings reveal a process of employment and wage polarization within regional labor markets that parallels the polarization of employment seen at the aggregate level in the US, UK and West Germany. In net, our results suggest an important role for changes in labor specialization—potentially spurred by displacement of routine task activities—as a driver of rising employment and wages in service occupations, and of polarization of employment and wage growth more generally.

Appendix Tables

Appendix Table 1. Ranking of Occupations by RTI Score (Lowest to Highest) for Service Occupations

I. Top Quintile of Ranking (Least Routine Intensive)		II. Second to Fourth Quintile of Ranking	
2	Fire Fighting, Prevention and Inspection	69	Superv. of Landscaping, Gardening, Groundskeeping
5	Public Transportation Attendants and Inspectors	80	Ushers
6	Police and Detectives, Public Service	110	Animal Caretakers, except Farm
9	Crossing Guards	134	Child Care Workers
13	Waiters and Waitresses	142	Guards and Police, except Public Service
16	Housekeepers, Maids, Butlers, and Cleaners	144	Supervisors of Guards
18	Sherrifs, Bailiffs, and Correctional Institution Officers	158	Laundry and Dry Cleaning Workers
24	Baggage Porters, Bellhops and Concierges	161	Supervisors of Food Preparation and Service
29	Recreation and Fitness Workers	206	Bartenders
35	Gardeners and Groundskeepers	239	Hairdressers and Cosmetologists
39	Recreation Facility Attendants		
41	Health and Nursing Aides	III. Bottom Quintile of Ranking (Most Routine Intensive)	
42	Pest Control Occupations		
49	Guides	250	Cooks
50	Supervisors of Cleaning and Building Service	287	Dental Assistants
54	Janitors	292	Barbers
56	Food Preparation Workers	298	Motion Picture Projectionists

Notes: The table indicates the RTI rank of detailled service occupations. It excludes residual occupations groups ("miscellaneous"/"other"/"not elsewhere classified").

Appendix Table 2. Levels and Changes of Share of Employment in Routine Occupations, Overall and by Education Group, 1980-2005

	Share of Employment in Routine Occs				Change 1980-2005
	1980	1990	2000	2005	
All Workers	0.332	0.329	0.330	0.315	-0.017
	(0.05)	(0.03)	(0.03)	(0.03)	
College Graduates	0.323	0.324	0.323	0.311	-0.012
	(0.04)	(0.04)	(0.04)	(0.04)	
Some College	0.377	0.363	0.363	0.349	-0.029
	(0.04)	(0.03)	(0.02)	(0.03)	
High School Graduates	0.367	0.338	0.329	0.311	-0.056
	(0.05)	(0.04)	(0.03)	(0.03)	
High School Dropouts	0.221	0.230	0.241	0.220	-0.001
	(0.04)	(0.04)	(0.04)	(0.05)	

n = 722 Commuting Zones in each decade, weighted by start of period commuting zone share of national population. Routine occupations are defined the occupations with largest routine task / manual task ratios that account for one third of overall employment in 1980.

Appendix Table 3. Cross-Sectional Correlates of the Routine Employment Share in 1980.

	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)	
College/Non-college pop 1980			0.176	**											0.067	**
			(0.020)												(0.020)	
Immigr/Non-college pop 1980					0.308	**									0.216	**
					(0.087)										(0.068)	
Manufact/empl 1980							-0.005								-0.002	
							(0.069)								(0.042)	
Unemployment rate 1980									-1.468	**					0.225	
									(0.183)						(0.169)	
Female empl/pop 1980											0.648	**			0.471	**
											(0.044)				(0.074)	
Age 65+/pop 1980													-0.909	**	-0.295	**
													(0.313)		(0.099)	
Constant	0.291	**	0.213	**	0.288	**	0.292	**	0.398	**	-0.03		0.394	**	0.043	
	0.000		(0.009)		(0.001)		(0.019)		(0.013)		(0.022)		(0.035)		(0.035)	
R^2	0.395		0.660		0.484		0.395		0.560		0.721		0.531		0.816	
State dummies	Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes	

n=722 Commuting Zones. Robust standard errors in parentheses are clustered on state. Models are weighted by start of period commuting zone share of national population and include an intercept. ~ p ≤ 0.10, * p ≤ 0.05, ** p ≤ 0.01.

Theory Appendix

Here we derive the solution to the model for a case where L_r and K are complements ($\mu < 1$). Note that we have $K(t) \to \infty$, so

$$\lim_{t\to\infty} \left[(1-\lambda)\left(\alpha_r L_r(t)\right)^{\mu} + \lambda\left(\alpha_k K(t)\right)^{\mu}\right]^{\beta/\mu} = (1-\lambda)^{\beta/\mu}\left(\alpha_r L_r^*\right)^{\beta}. \tag{2.6.1}$$

Consequently,

$$\begin{aligned}\lim_{t\to\infty} F\left(1, g\left(L_m(t)\right), t\right) &= \lim_{t\to\infty} \left[(1-\lambda)\left(\alpha_r L_r(t)\right)^{\mu} + \lambda\left(\alpha_k K(t)\right)^{\mu}\right]^{\beta/\mu} - \theta e^{-\delta t} K(t) \quad (2.6.2)\\ &\le \lim_{t\to\infty} \left[(1-\lambda)\left(\alpha_r L_r(t)\right)^{\mu} + \lambda\left(\alpha_k K(t)\right)^{\mu}\right]^{\beta/\mu}\\ &= (1-\lambda)^{\beta/\mu}\left(\alpha_r g\left(L_m^*\right)\right)^{\beta}\end{aligned}$$

Moreover, since $K(t)$ solves Eq. (2.2.17), it does better than an arbitrary choice for the capital function. In particular, it does better than $\tilde{K}(t) = t$. Then, we have

$$\lim_{t\to\infty} F(1, g(L_m(t)), t) \geq \lim_{t\to\infty} \left[(1-\lambda)(\alpha_r L_r(t))^\mu + \lambda\left(\alpha_k \tilde{K}(t)\right)^\mu\right]^{\beta/\mu} - \theta e^{-\delta t}\tilde{K}(t) \tag{2.6.3}$$

$$= \lim_{t\to\infty} \left[(1-\lambda)(\alpha_r L_r(t))^\mu + \lambda\left(\alpha_k \tilde{K}(t)\right)^\mu\right]^{\beta/\mu}$$

$$= (1-\lambda)^{\beta/\mu}(\alpha_r g(L_m^*))^\beta.$$

Combining Eqs. (2.6.2) and (2.6.3), we have

$$\lim_{t\to\infty} F(1, g(L_m(t)), t) = (1-\lambda)^{\beta/\mu}(\alpha_r g(L_m^*))^\beta. \tag{2.6.4}$$

In words, since L_r and K are gross complements and K grows, in the limit $L_r(t) = g(L_m(t))$ becomes the bottleneck and determines the production.

Next consider

$$\frac{dF(1, g(L_m(t)), t)}{dL_r(t)} = \beta(1-\lambda)\alpha_r^\mu L_r(t)^{\mu-1}\left[(1-\lambda)(\alpha_r L_r(t))^\mu + \lambda(\alpha_k K(t))^\mu\right]^{(\beta-\mu)/\mu} \tag{2.6.5}$$

Since $K(t) \to \infty$, taking the limit of this expression yields

$$\lim_{t\to\infty} \frac{dF(1, g(L_m(t)), t)}{dL_r(t)} = \beta(1-\lambda)^{\beta/\mu}\alpha_r^\beta g(L_m^*)^{\beta-1}. \tag{2.6.6}$$

Taking the limit of Eq. (2.2.24) and plugging in Eqs. (2.6.4) and (2.6.6), we have

$$\left[(1-\lambda)^{\beta/\mu}(\alpha_r g(L_m^*))^\beta\right]^{1/\sigma_c} = -L_m^{*\,1/\sigma_c}\log(1-L_m^*)\,\beta(1-\lambda)^{\beta/\mu}\alpha_r^\beta g(L_m^*)^{\beta-1}. \tag{2.6.7}$$

The equilibrium level of L_m^* in the limit is the solution to the previous equation, which will be in the interval $(0,1)$.

Moreover, in this case we have

$$p_s \to p_s^*,\ w_m \to w_m^*,\ w_r \to w_r^*,\ w_a \to w_a^*,\ \eta \to \eta^*,$$

i.e. all variables converge to a finite constant. Intuitively, in this case machines and routine labor are gross complements so technological progress is not sufficient to increase output beyond a finite limit (since routine labor becomes the bottleneck). Consequently, the price of services and hence the wage for the manual labor also remain constant. The wage for routine labor remains constant since the routine labor is the bottleneck so there is still value to routine tasks. The abstract wage is also constant since the abstract workers receive a constant share of output, which is constant.

In this case, $w_a(t)/w_m(t)$ ratio also goes to a constant w_a^*/w_m^* regardless of σ_c, in contrast with the conjecture. We summarize our results in the following proposition.

Proposition 3 *When* $\mu < 0$, $\lim_{t\to\infty} L_m(t) = L_m^*$ *where* $L_m^* \in (0,1)$ *is a solution to Eq.* (2.6.7). *In the limit, unskilled labor works in both manual and routine tasks and the wages limit to finite levels*

$$w_m \to w_m^*,\ w_r \to w_r^*,\ w_a \to w_a^*.$$

Chapter 3

This Job is 'Getting Old:' Measuring Changes in Job Opportunities using Occupational Age Structure

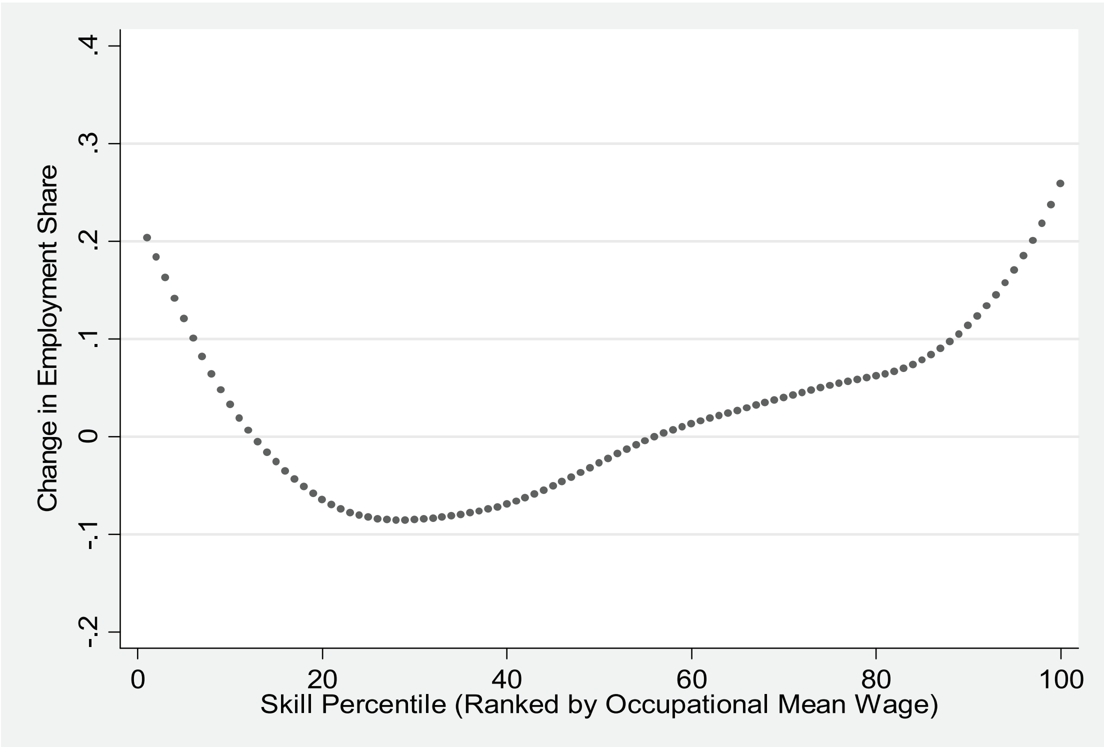
3.1 Introduction

One of the most remarkable developments in the U.S. labor market of the past two and a half decades has been the rapid, simultaneous growth of employment in both the highest- and lowest-skilled jobs. This phenomenon is depicted in Figure 1, which plots changes in the share of aggregate hours worked at each percentile of the occupational skill distribution over the period 1980 through 2005. These skill percentiles are constructed by ranking occupations according to their mean hourly wages in 1980 and grouping them into 100 bins, each comprising one percent of 1980 employment.[1] The pronounced U-shape of Figure 1 underscores that employment growth over this 25 year period has been disproportionate in the top and bottom of the occupational skill distribution. Occupations that were in the lowest and highest deciles of the 1980 distribution grew in relative size by 10 to 25 percent between 1980 and 2005, while occupations in the 2nd through 6th deciles contracted.[2] This hollowing out, or 'polarization,' of the occupational employment distribution is not unique to the United States. Using harmonized European Union Labour Force Survey data, Goos, Manning and Salomons (2008) find that in 14 of 16 European countries for which data are available, high- and low-paying occupations expanded relative

[1] All analyses in the paper use data from the 1980, 1990 and 2000 Census IPUMS and the 2005 American Community Survey (ACS). Samples are limited to workers of age 16 through 64 years in the prior year, and all calculations are weighted by labor supply, equal to the product of the Census sampling weight, weeks worked in the prior year, and usual weekly hours. We group occupations into a balanced panel of 330 harmonized Census Occupation categories encompassing all of U.S. employment over 1980 through 2005.

[2] The series in Figure 1 is smoothed with a locally weighted regression using a bandwidth of 0.8. Results are extremely similar if we use the 2000 Census IPUMS in place of the 2005 ACS.

Figure 1. Smoothed Changes in Employment Share
by Occupational Skill Percentile, 1980 - 2005



to middle-wage occupations in the 1990s and 2000s.

A leading explanation for the hollowing out of the occupation distribution in industrial countries is that non-neutral technical change, augmented by offshoring, is eroding demand for middle-skilled 'routine' cognitive and manual activities, such as bookkeeping, clerical work and repetitive production tasks (Acemoglu, 1999; Autor, Levy and Murnane, 2003, 'ALM' hereafter; Autor, Katz and Kearney, 2006; Blinder, 2007; Goos and Manning, 2007; Autor and Dorn, 2008.[3] Because the core job tasks of these occupations follow precise, well-understood procedures, they are increasingly codified in computer software and performed by machines or, alternatively, offshored over computer networks to foreign worksites. This displacement of routine job tasks raises relative demand for non-routine tasks in which workers hold a comparative advantage over current technology, in particular 'abstract' tasks requiring problem-solving, creativity, or complex interpersonal interactions (e.g., attorneys, scientists, managers), and 'manual' tasks requiring, variously, situational adaptability, visual and language recognition, and in–person interactions (e.g., janitors and cleaners, home health aides, beauticians, construction laborers, security personnel, and motor vehicle operators). Notably, these two categories of non-routine tasks lie at opposite ends of the skill distribution: abstract tasks are the core

[3]See also Manning (2004), and Mazzolari and Ragusa (2008) for an alternative hypothesis attributing the growth of low-skill employment to marketization of household production.

activity of professional specialty and technical occupations while manual tasks are most intensive in personal service, transportation, construction, and operative occupations. Thus, displacement of occupations intensive in routine tasks and growth of occupations intensive in non-routine tasks may give rise to the U-shaped pattern of job growth visible in Figure 1.

An important, unstudied question raised by this pattern of non-neutral occupational change is: where do the routine workers go? In particular, as middle-skill routine occupations decline, which age and skill groups move upward in the occupational distribution towards high-skill, non-routine jobs, and conversely which groups gravitate downwards towards the lower tail of non-routine occupations? Analyzing this process of occupational change offers insights into the shifting opportunity set faced by workers at different age and education levels.

Our analysis relies on a simple and, to the best of our knowledge, novel approach for measuring how changing job opportunities affect worker re-allocation across occupations. The underlying idea of this approach is that because workers develop occupation-specific human capital as they gain work experience, skill specificity makes the costs of occupational mobility higher for older than younger workers Neal (1999). When an occupation declines, therefore, older workers will face an incentive not to exit the occupation while younger workers will face an incentive not to enter. Moreover, firms may react to changing demands for occupations by hiring young workers into growing occupations and curtailing such hiring into contracting jobs. These suppositions imply that occupations will 'get old' as their employment declines—that is, the mean age of an occupation's workforce will rise.

The plan of the paper is as follows. We first offer a simple 'proof of concept' to demonstrate the tight empirical link between declines in an occupation's employment and increases in the mean age of its workforce. The balance of the paper then applies this tool to the study of local labor markets to assess how shifts in occupational structure have affected the job composition of young and old workers at different education levels between 1980 and 2005. In particular, we exploit pre-existing differences in occupational specialization across local labor markets to identify areas subject to differing degrees of 'hollowing out' of employment. We use this variation to measure how, in response to contracting routine employment, workers of differing ages and education levels are reallocated towards the tails of the occupation distribution.

3.2 Are Middle-Skill Jobs Getting Old?

We first document the robust relationship between changes in occupational size and shifts in the age distribution of the occupation's workforce. Table 1 reports simple bivari-

Table 1. Predicting Changes in the Age Structure of Occupations 1980 - 2005 using Changes in Occupation Size and Initial Routine Task-Intensity

	Δ Mean Age (1)		Δ Share of Workers in Age Bracket: Young 16-29 (2)		Prime 30-54 (3)		Older 55-64 (4)	
					A. OLS model 1			
Δ Occ's Share of Total Emp (% pts)	-0.78	**	0.027	**	-0.020	**	-0.007	~
	(0.18)		(0.006)		(0.005)		(0.003)	
					B. OLS model 2			
Occ's Routine Task Intensity in	0.55	**	-0.015	**	0.004		0.011	**
	(0.11)		(0.003)		(0.003)		(0.002)	
					C. OLS model 3			
Occ's Routine Task Intensity in	0.50	**	-0.013	**	0.003		0.010	**
	(0.10)		(0.003)		(0.003)		(0.002)	
Δ Occ's Share of Total Emp (% pts)	-0.66	**	0.023	**	-0.019	**	-0.004	
	(0.18)		(0.006)		(0.005)		(0.003)	
					D. Descriptive statistics			
Mean	3.25		-0.128		0.120		0.008	
SD	(1.99)		(0.065)		(0.057)		(0.036)	

N=330 harmonized occupations. Each column of panels A, B, and C corresponds to a separate OLS regression of the outcome variable at the top of the column on tabulated control variables and a constant. Standard errors are in parentheses. Models are weighted by occupational shares in total hours worked in 1980. ~ $p \leq 0.10$, * $p \leq 0.05$, ** $p \leq 0.01$.

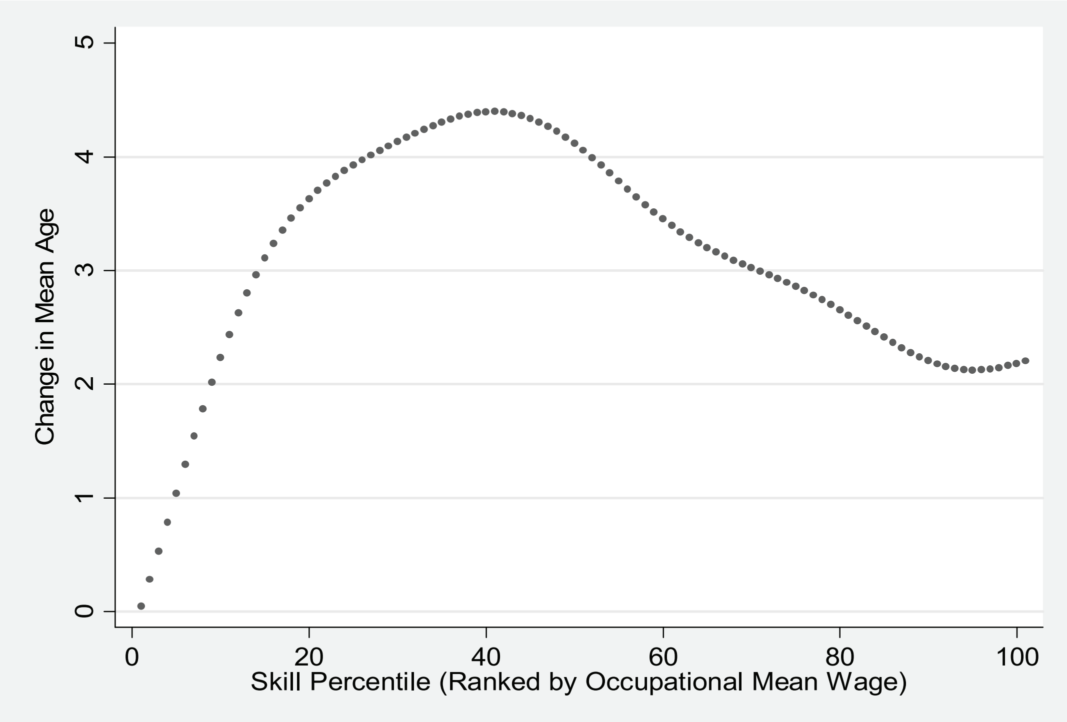
ate regressions of the form:

$$\Delta Y_j = \alpha + \beta_1 \Delta E_j + \varepsilon_j, \tag{3.2.1}$$

where Y_j is the mean age of workers in occupation j or the share of workers in that occupation who fall into a given age bracket, E is the share of an occupation in total employment in a given year, and the Δ operator denotes the change in a variable over the time interval 1980 to 2005.

The average age of the working population rose by 3.3 years during 1980 through 2005, reflecting the aging of the baby boom cohorts. Occupations that contracted over this period aged substantially faster than average. Column (1) of the first panel shows that occupations that contracted by 1 percentage point as a share of aggregate employment between 1980 and 2005 gained in age by an additional 0.78 years relative to the mean.

Figure 2. Smoothed Changes in Mean Worker Age
by Occupational Skill Percentile, 1980 - 2005



Columns (2) through (4) show that, as hypothesized, age increases in contracting occupations are driven by a falling employment share of young workers and rising employment shares of prime age and older workers.

Figure 2 plots smoothed changes in the mean age of workers by occupational skill percentile between 1980 and 2005. This figure shows a distinct inverted U-shape that is a near mirror-image of changes in occupational employment shares depicted in Figure 1. Occupations in the bottom and top two deciles of the skill distribution aged by roughly 2 years between 1980 and 2005, which is substantially *below* the overall average of 3.3 years. By contrast, occupations in the second through sixth skill deciles aged disproportionately rapidly, gaining approximately 4 years on average. Thus, over the last twenty-five years, middle-skill jobs have gotten old.

3.3 Are Routine Task-Intensive Jobs Getting Old?

If routine tasks are indeed being supplanted by technology and offshoring, then employment declines should be concentrated in occupations that are specialized in such tasks. The year 1980 is a particularly apt starting point for gauging the effects of workplace computerization. National Income and Product Accounts data show that the share of computer hardware and software in all U.S. private nonresidential capital investment hovered at approximately 4 percent from 1970 to 1978, and then rose steeply at approximately three-quarters of a percentage point *per year* through the year 2000.[4] Thus, occupations concentrated in routine tasks would be predicted by our hypothesis to experience sharp contractions from 1980 going forward.

To assess this hypothesis using occupational age structure as above, we draw on occupation level data assembled by ALM, who merge job data on task requirements—manual, routine and abstract—from the fourth edition of the US Department of Labor's *Dictionary of Occupational Titles* (U. S. Department of Labor, Employment and Training Administration, 1977) to their corresponding Census occupation classifications. For each occupation j, we form an index of routine task-intensity, RTI:

$$RTI_j = \ln\left(\hat{R}_{j,1980}/\hat{M}_{j,1980}\right), \tag{3.3.1}$$

where $\hat{R}$ and $\hat{M}$ are, respectively, the intensity of routine and manual task input in each occupation in 1980, measured on a 0 to 10 scale. This measure is rising in the relative importance of routine tasks within an occupation and falling in the relative importance of manual tasks. Since RTI does not have a cardinal scale, we standardize it with a mean of zero and an employment weighted, cross-occupation standard deviation of unity in 1980.

This simple measure appears to capture well the job categories that motivate our conceptual framework. Among the 10 most routine task-intensive occupations in our sample of 330, 6 are clerical and accounting occupations and several others represent repetitive physical motion activities. Among the 10 least routine task intensive occupations, 4 are in-person service occupations, while the remainder involve driving motor vehicles.[5]

To test the link between routine task-intensity and changes in age structure, we estimate a variant of equation (3.2.1) in which the RTI measure is included as a predictor of changes in occupational age structure. The second and third panels of Table 1 show that this variable is highly significant in all specifications. Occupations that in 1980 were one standard deviation above the mean of routine-intensity, gain 0.6 years of age relative to the mean over the next twenty-five years. This age gain is driven by declining relative employment of young workers in routine task-intensive occupations, and by rising relative employment of older workers, particularly those ages 55 to 64. The third panel of Table 1 shows that the predictive relationship between routine-intensity and occupation aging is quite robust to controlling for contemporaneous changes in occupations' employment shares—though of course the employment shares of routine task-intensive occupations fall significantly in this period.

Thus, like middle-skill occupations, routine task-intensive occupations are getting old.

[4] Authors' calculations using NIPA data (U.S. Department of Commerce, 2002).

[5] Additional details on the Routine Task Intensity measure are found Autor and Dorn (2008), who develop this measure using the ALM data.

This finding is not entirely surprising, of course; middle-skill occupations are also disproportionately routine task-intensive.

3.4 Where do the Routine Workers Go?

3.4.1 Analysis of Local Labor Markets

We now exploit the robust predictive relationship between occupational decline and aging to study how the decline of routine occupations affects the opportunity set of workers at different age and skill levels. Specifically, we ask which *non-routine* jobs absorb young and older workers as routine task-intensive occupations are displaced.

For this analysis, we shift the unit of observation from changes in age structure within occupations to changes in the age composition of employment within local labor markets, following an approach developed by Smith (2008). Based on the results above, we anticipate that local labor markets that were specialized in routine task-intensive occupations at the start of the sample period should have experienced a differential contraction of middle-skill jobs over the subsequent 25 years. We use this cross-market variation in (expected) occupational declines to analyze the effect of the thinning of the ranks of middle-skill occupations on the occupational distribution of young and old workers.

As a time-consistent measure of local labor markets, we implement the concept of Commuting Zones ('CZs'), developed by Tolbert and Sizer (1996), who used confidential commuting data from the 1990 Census to identify clusters of counties—i.e., Commuting Zones—that exhibit strong commuting ties within clusters but weak commuting across clusters. Our analysis uses 722 CZs that cover the entire mainland of the US, including both metropolitan and rural areas.[6]

To measure cross-market variation in employment in routine task-intensive occupations, we apply a simple binary approach to distinguish 'routine' and 'non-routine' occupations. We classify as routine occupations those that fall in the top-third of the employment-weighted distribution of the RTI measure in 1980. Using this classification, we then assign to each commuting zone k a routine employment share measure (RSH_{kt}) equal to the fraction of CZ employment at the start of a decade that falls in routine task-intensive occupations. The mean of this measure in 1980 is, by construction, equal to 0.33. The population weighted 80/20 percentile difference in routine employment share is 10 percentage points (specifically, $RSH^{P20} = 0.27$ and $RSH^{P80} = 0.37$).

[6]Commuting zones have two advantages over other geographic units typically used for analysis of local labor markets: they are based primarily on economic geography rather than incidental factors such as minimum population or state boundaries; and they cover the entire U.S. In addition, it is possible to use Census Public Use Micro Areas (PUMAs) to consistently match Census geography to CZs for the full period of our analysis (see Autor and Dorn, 2008 for details).

Putting these pieces together, we estimate in Table 2 a set of OLS stacked first-difference models for CZ level changes in occupational employment by age and education:

$$\Delta Y_{ak\tau} = \alpha_\tau + \beta_2 RSH_{kt} + \mathbf{X}'_{kt}\beta_3 + \omega_{ak\tau}, \tag{3.4.1}$$

where Y is an outcome measure for age-education group a in commuting zone k over the 5 or 10 year time interval τ, RSH is the routine employment share in the CZ at the start of the time interval, α is a vector of time dummies, and $\mathbf{X}$ is a vector of start-of-period control variables, including state dummies and measures of the initial age, education, and employment structure of the commuting zone.

3.4.2 Employment in Routine Occupations

Estimates of equation (3.4.1) produce a number of striking results. Panel A of Table 2 shows that, as predicted, CZ's that were initially specialized in routine task-intensive occupations saw substantial declines in the share of workers employed in these occupations between 1980 and 2005. These declines are evident at all age levels, but they are uniformly larger for younger than older workers. Interestingly, the decline in routine employment is greater for non-college workers (high school or lower education) than for college workers (at least one year of college). A potential explanation for this pattern is that less educated workers in routine task-intensive occupations perform a disproportionate share of the routine tasks, and thus are differentially subject to displacement.

3.4.3 Employment in Non-Routine Occupations

Which occupations absorb workers from these different age brackets as routine task-intensive jobs in a Commuting Zone contract? To form a simple accounting, we use occupational wage data from 1980 to evenly divide the two-thirds of employment classified as non-routine into two occupation clusters containing equal shares of 1980 employment, one cluster containing low-wage occupations and the other high-wage occupations. Notably, these occupational clusters roughly correspond to the two non-routine task categories defined above (i.e., abstract and manual). The high-skill non-routine cluster is largely composed of professional specialty and technical occupations, with mean log hourly wages that are 40 percent above the routine occupation mean. The low-skill non-routine group is largely composed of low-education service, labor, and operative occupations, with mean log hourly wages 20 percent below the routine occupation mean.

Panels B and C of Table 2 show that relative declines in routine occupation employment within CZs are primarily offset by relative employment gains in *low-skill* non-routine occupations—jobs that are significantly less skill-intensive and lower-paying than the routine occupations that are displaced. Among the three age brackets we consider, only young

Table 2. Predicting Changes in the Allocation of Age Groups across Occupations using Initial Commuting Zone Employment Shares in Routine-Intensive Occupations, 1980-2005.

Coefficient on [Share of Routine Occs$_{-1}$] for skill group below	Δ Occ's Share of Age Bracket Young 16-29		Prime 30-54		Older 55-64	
	A. Routine-Intensive Occs					
All workers	-0.31	**	-0.21	**	-0.25	**
	(0.02)		(0.01)		(0.03)	
College workers	-0.18	**	-0.11	**	-0.12	**
	(0.03)		(0.01)		(0.03)	
Non-college workers	-0.46	**	-0.28	**	-0.23	**
	(0.04)		(0.03)		(0.05)	
	B. High-Skill Non-Routine Occs					
All workers	0.10	**	-0.01		-0.06	*
	(0.02)		(0.02)		(0.03)	
College workers	0.15	**	0.03		-0.03	
	(0.03)		(0.02)		(0.03)	
Non-college workers	0.04		-0.09	**	-0.19	**
	(0.03)		(0.02)		(0.04)	
	C. Low-Skill Non-Routine Occs					
All workers	0.21	**	0.22	**	0.31	**
	(0.03)		(0.02)		(0.04)	
College workers	0.03		0.08	**	0.15	**
	(0.03)		(0.02)		(0.03)	
Non-college workers	0.42	**	0.37	**	0.42	**
	(0.04)		(0.03)		(0.05)	

N=2166 (3 time periods x 722 commuting zones). Robust standard errors in parentheses are clustered on state. Each cell corresponds to a separate stacked first difference model. Models are weighted by start of period commuting zone share of national population and contain a constant and two time dummies. College workers are those with at least some college education. Occupations are classified as routine task-intensive if they fall in the top third of the employment-weighted distribution of the routine-intensity measure in 1980. Nonroutine occupations are classified as high-skill if they fall in the top half of the employment-weighted distribution of mean wages in non-routine occupations in 1980, and are classified as low-skill otherwise. ~ $p \le 0.10$, * $p \le 0.05$, ** $p \le 0.01$.

workers ages 16 to 29 gain in employment in high-skill non-routine occupations. Both the prime age and older age groups gain employment in low-skill non-routine occupations. Moreover, even among the young, employment gains in high-skill non-routine occupations are less than half as large as gains in low-skill non-routine occupations. Clearly, the hollowing out of employment in initially routine task-intensive local labor markets primarily generates a movement of employment into low-skill, non-routine jobs.

3.4.4 Outcomes by Education Groups

When we drill down on these occupational shifts by education group, however, it is immediately apparent that declines in routine employment have decidedly non-neutral impacts across education groups and across age groups within an education level. The second row of Table 2 shows that college workers in initially routine task-intensive labor markets gain employment in both high- and low-skill non-routine jobs. But the gains in high-skill non-routine employment are concentrated among the young and almost entirely absent among the old. Thus, the bulk of the differential decline in routine employment among young college workers in these labor markets is offset by gains in high-skill non-routine employment. Among prime-age and older workers, however, offsetting employment gains are found mostly in low-skill non-routine jobs. Thus, it appears that the opportunity for high-education workers to reallocate upward depends greatly on age.

The final row of Table 2 portrays an even less encouraging picture for non-college workers. For this skill group, the entire differential decline in routine employment in routine task-intensive CZ's is absorbed by increased low-skill, non-routine employment. Moreover, while young non-college workers roughly hold their ground in high-skill, non-routine occupations, prime age and older non-college workers differentially lose employment in high-skill, non-routine jobs and gain in employment in low-skill, non-routine jobs.

In summary, the occupational structure of *college* workers in routine task-intensive labor markets is hollowing out, with movement of workers towards both tails. The occupational structure of *non-college* workers, however, is shifting uniformly leftward towards lower-paying, non-routine jobs. This leftward shift is most pronounced for older age groups.

These patterns are quite robust. Although the models in Table 2 include only time dummies and implicitly, Commuting Zone effects (recall that these are first-difference models), the qualitative pattern of results is little changed when the model is augmented with state fixed effects and detailed controls for the initial human capital, demographic, and industrial structure of commuting zones.[7]

[7] A supplementary table is available from the authors.

3.5 Conclusion

Aggregate employment has shifted over the last twenty-five years against middle-skill, routine task-intensive work and towards the tails of the occupational skill distribution. Occupations at both tails are appropriately labeled as non-routine, but they differ greatly in skill and pay. The right-hand tail of the distribution encompasses high-skill, non-routine occupations that typically require capabilities in problem-solving, abstract reasoning, and decision-making. The left-hand tail encompasses low-skill non-routine occupations that demand basic human adaptability but little in the way of formal training. The contraction of routine occupations is therefore likely to have different impacts on workers across age and skill groups, depending on their ability to move upward towards high-skill, non-routine jobs or, conversely, downwards towards low-skill, non-routine jobs.

By comparing local labor markets that, due to initial differences in concentration in routine task-intensive activities in 1980, are subject to differing degrees of occupational change over the subsequent twenty-five years, we find that contractions of routine employment within local markets disproportionately raise the share of workers employed in low-skill, non-routine jobs. In fact, only the youngest category of workers exhibits both downward *and* upward occupational reallocation; for other age groups, movement is entirely downward. Highly-educated workers are clearly better prepared to adapt to changing occupational opportunities, and thus it is to be expected that college-educated workers are reallocating upwards as well as downwards. But the degree of upward reallocation is strongly negatively correlated with age: while young college workers are gaining employment in high-skill, non-routine occupations, older college workers are increasingly found in low-skill, non-routine work.

These secular shifts of age and education groups across occupational categories provide only a preliminary sense of what may potentially be learned from changing occupational age structures. We focus here on changes in occupational structure *within* age groups, but do not address how changes in aggregate job composition shape the progression of birth cohorts across occupations as they age. Our ongoing work suggests that the simple tools used above hold some promise for exploring these questions.

Chapter 4

Price and Prejudice: The Interaction between Preferences and Incentives in the Dynamics of Racial Segregation

4.1 Introduction

Racial segregation is a salient feature of urban neighborhoods in the United States. A high degree of racial segregation in a city is associated with worse outcomes for young blacks such as lower education, income, and employment (e.g., Cutler and Glaeser, 1997; Ananat, 2007). A frequently suggested cause of segregation is the racial preference of whites for living among white neighbors. In a seminal paper, Schelling (1971) demonstrates that substantial segregation can result even from weak racial prejudice. After the white residents with lowest tolerance for minorities leave a neighborhood, the minority share increases and induces the departure of less prejudiced whites, thereby causing a sequence of white flight.

The dynamic of neighborhood tipping may also be influenced by the interplay of two features absent from the Schelling model: prices and expectations. The tipping process can depress house values in a neighborhood when the willingness of whites to pay for housing falls in response to an increasing minority share. As a result, forward-looking homeowners may not only want to move out of a neighborhood because of an expected increase in minority population but also, and perhaps as important, to avoid an associated decrease in house values.

This paper proposes an augmented tipping model in which white and minority renters and homeowners interact spatially, through the neighborhood minority share, and financially, through house prices and rents. Whites' utility falls when the neighborhood minority share becomes large and the reduced demand of whites for housing translates into lower house prices and rents. When current neighborhood residents can more accurately predict the probability of a shock that will raise the minority share and lower prices, incumbent homeowners, on receiving a signal that the shock probability is high, have a financial incentive to sell their houses to outside buyers. In contrast, white renters face only the risk of an increasing minority share — not the additional asset value risk — and hence may prefer to stay in the neighborhood rather than incurring a moving cost for departing. Accordingly, the model predicts that the decline of the white population in tipping neighborhoods will be more pronounced when neighborhoods have a large homeownership rate.

Such financial incentives for homeowners to depart from tipping neighborhoods can have important consequences for neighborhood socioeconomic composition. In the Schelling model, the order of whites' departure from a tipping neighborhood is determined solely by individual levels of racial prejudice, which several surveys show is most pronounced among whites with low educational attainment and income (Farley, Fielding and Krysan, 1997; Charles and Guryan, 2008). Thus, the departure of whites with the lowest minority tolerance could increase the income and educational levels of tipping neighborhoods.[1] If, however, homeowners leave ahead of renters, average wealth and educational attainment will fall because homeowners tend to be wealthier and better educated than renters.

Building on recent work by Card, Mas and Rothstein (2008a), the empirical analysis tests these predictions using 1970-2000 data from a large panel of urban neighborhoods (census tracts) covering over 100 Metropolitan Statistical Areas (MSAs). It uses regression discontinuity models to assess whether the magnitude of discontinuous changes in neighborhood composition at empirically estimated minority share tipping points varies with the neighborhoods' initial homeownership rates. The results confirm that neighborhoods with high ownership rates experience substantially larger discontinuous drops in white population and larger declines in house values. The reduction in white population is primarily due to a fall in white homeownership and tipping neighborhoods experience a decrease in income and education levels. These findings support the proposition that pecuniary incentives for homeowners may exacerbate the tipping process.

The analysis also considers several alternative explanations for strong tipping effects in neighborhoods with high ownership rates. Even when the relatively wealthier homeowners are not more prejudiced than renters, they may be more inclined to leave a tipping neighborhood because they can better afford to relocate. Additionally, owners are more likely to have children and parents may be particularly concerned about changes in neighborhood composition and public goods that could affect their offspring. Finally, white residents who are sensitive to an increase in minority population may be concentrated in

[1] Bayer, Fang and McMillan (2005) document the existence of neighborhoods that have mostly black residents together with very high shares of college graduates.

certain geographic areas of cities that have larger homeownership rates, such as suburbs. However, robustness tests show that none of these variations in population composition or neighborhood location can explain the larger tipping effects in owner-dominated neighborhoods.

The paper is organized as follows. Section 2 reviews related literature on neighborhood segregation and tipping. Section 3 outlines a theoretical model of neighborhood tipping that shows different moving incentives for homeowners and renters when neighborhood residents anticipate a possible increase of minority population. Section 4 describes the data and econometric approach for the empirical analysis, after which section 5 presents the regression discontinuity estimates for the changes in owner- and renter-occupied housing at empirical tipping points. Sections 6 and 7 then consider changes in white population and in house values, and section 8 presents results for income and education levels. Section 9 concludes the paper.

4.2 Related Literature

Urban neighborhoods in the United States are characterized by pronounced segregation between whites and blacks, and between natives and recent immigrant groups (Cutler, Glaeser and Vigdor, 1999, 2008). The correlation between such residential segregation and adverse economic outcomes for minorities has been long recognized (Kain, 1968). A large body of literature analyzes the causality of this relationship between neighborhood quality and economic outcomes; examples include work by Cutler and Glaeser (1997) and Ananat (2007) who exploit cross-city variation in segregation, and experimental studies such as Oreopoulos (2003), Jacob (2004) or Kling, Liebman and Katz (2007) which evaluate neighborhood relocation programs. While evidence from this literature is mixed, the creation of racially and economically diverse urban neighborhoods has nonetheless become an important goal of urban public policy.

A different stream of the literature analyzes the determinants and dynamics of segregation at the neighborhood level. Influential work by Schelling (1969, 1971, 1978) proposes that even weak racial preferences can lead to strong residential segregation. In the Schelling model, all white residents have individual tolerance thresholds for the minority share in their neighborhoods, a preference structure that can lead to rapid segregation of a mixed-race neighborhood when the minority share increases beyond a critical tipping point. That is, once the minority share exceeds the threshold of the most prejudiced white, this person departs. As a consequence, the minority share increases further, and a cascade of white flight from the neighborhood begins that eventually leads to an all-minority equilibrium.

While the Schelling model does not integrate prices, many subsequent models derive tipping behavior in frameworks with explicit housing markets (e.g., Miyao, 1978, 1979; Coulson and Bond, 1990; Benabou, 1993; Becker and Murphy, 2000; Card, Mas and Rothstein, 2008a). For instance, the tipping model proposed by Card et al. (2008a) allows that house prices decline when a neighborhood tips and whites' demand for housing in the neighborhood declines: When a shock increases the minority share of a neighborhood, whites' valuation for houses falls and minority agents who were previously outbid can move into the neighborhood. However, in this and all the aforementioned models, neighborhood residents are myopic and fail to anticipate a possible change in house values. As in the Schelling model, the role of expectations in tipping dynamics is therefore not assessed.

However, a model by Frankel and Pauzner (2002) shows that tipping can occur even when agents have rational expectations. These authors analyze an initially white neighborhood that can tip to an all-black equilibrium when the neighborhood becomes more attractive for blacks because of either a deterministic upward trend in black valuation of the neighborhood or small stochastic shocks to that valuation. When whites expect the neighborhood to tip in the future and moving opportunities for whites occur at random times, white residents will leave before black valuation exceeds their own. Expectations of tipping can therefore be self-fulfilling and anticipate the timing of tipping. However, because their model assumes a large and homogeneous pool of potential neighborhood residents, house prices do not change discontinuously when a neighborhood tips. In contrast, the theoretical model developed here (see section 3) will combine forward-looking behavior of agents and falling house prices in the tipping process.

Only recently has large-scale empirical evidence on neighborhood tipping been available. Card et al. (2008a), who analyze the racial dynamics of neighborhoods based on a large panel of census tracts, show significant discontinuous declines in white neighborhood population at empirically estimated city-specific tipping points.[2] Their study also finds small but mostly insignificant declines in house values in tipping neighborhoods.[3] Related work by Saiz and Wachter (2006) provides evidence for a relative devaluation of houses in neighborhoods that experience inflows of immigrants. The empirical analysis conducted here will build on the empirical framework of Card et al. (2008a) to study the impact of neighborhood homeownership rates on tipping behavior.

Although, to the best of my knowledge, no previous theoretical or empirical literature examines the role of homeownership in the dynamics of neighborhood tipping, several papers address the effect of ownership on social capital or neighborhood amenities (e.g., Sampson, Raudenbush and Earls, 1997; DiPasquale and Glaeser, 1999; Hoff and Sen, 2005). This literature argues that homeownership is beneficial for neighborhoods because

[2]A separate study by the same authors does not find evidence of a reverse tipping where a large number of minorities leave a neighborhood once the minority share falls below a tipping point (Card, Mas and Rothstein, 2008b).

[3]Another analysis of the same data finds little evidence for tipping effects, partly because it does not allow for city-specific locations of tipping points (Easterly, 2005).

homeowners are more likely than renters to engage in community activities and contribute to neighborhood public goods. In contrast to renters, homeowners benefit not only from the consumption value of public goods but also from their capitalization in houses values. However, the same exposure of homeowners to changes in house prices can generate an incentive to leave a neighborhood when incumbent residents anticipate a possible increase in the neighborhood minority share that would depress house prices.

4.3 Theoretical Framework

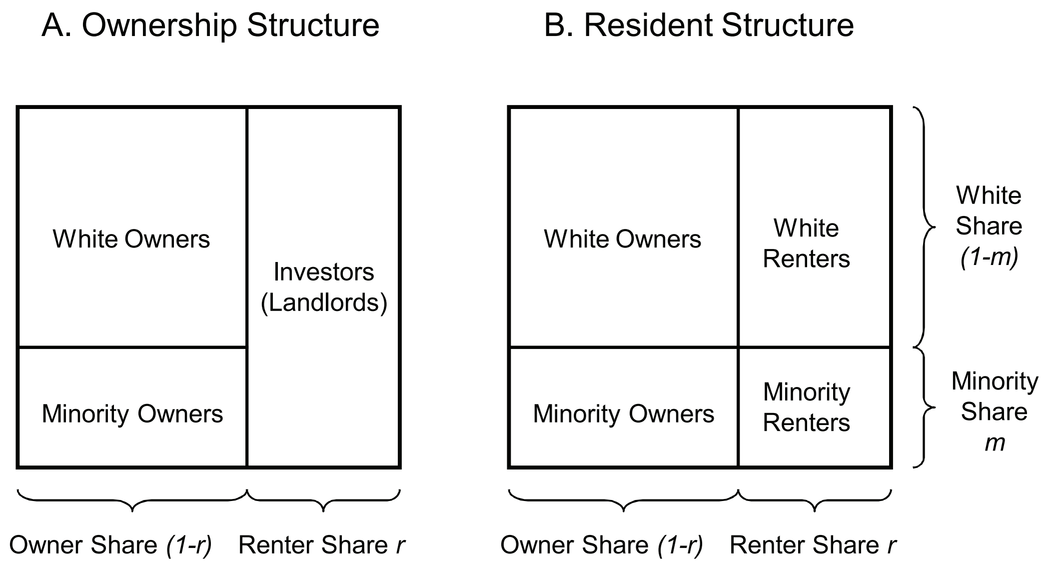
The theoretical model proposed here emphasizes that identical expectations about the probability of future neighborhood tipping can translate into different behavior by homeowners and renters. Specifically, it predicts that among otherwise identical neighborhoods, those with a higher homeowner share will experience a larger discontinuous decrease in white population at a critical value of the minority share at which white owners start leaving the neighborhood.

The key observation underlying this theoretical model is that homeowners and home renters face differential incentives to flee a neighborhood at risk of tipping. This prediction is obtained from one period of transactions in a housing market in which neighborhood incumbents have an informational advantage over potential new residents. In particular, incumbents have more precise information about the probability of an exogenous shock that would raise the share of minority residents in the neighborhood. If the minority share exceeds a certain threshold, whites' taste for the neighborhood and house prices decrease. Thus, whereas both white homeowners and renters dislike a high minority share, homeowners are also affected by falling house prices. If neighborhood incumbents anticipate a high likelihood of a shock that would push the minority share over the critical threshold, white homeowners have an additional financial incentive to leave the neighborhood, whereas white renters may decide to stay to avoid a moving cost.

Overall, the model predicts that neighborhoods with high homeownership rates will experience particularly large reductions in white population and house prices once the minority share exceeds a critical tipping point. White owner-occupied housing falls relative to renter-occupied housing, and the share of wealthy residents decreases because of a correlation between homeownership and wealth.

4.3.1 Neighborhood Structure and Agents

Consider a neighborhood with a fixed supply of N identical houses. There are five types of agents: whites and minorities, who can either have a high or low taste for the neighborhood, and investors. *Whites* and *minorities* can live in the neighborhood as home owners or renters. *Investors* own all rental houses. Investors are profit maximizers



who live outside the neighborhood as absentee landlords; they do not affect neighborhood racial composition. All agents are risk-neutral and can own one house at most.

Figure 1 illustrates a possible neighborhood ownership and resident structure. A share r of all houses are owned by investors while the remaining $(1-r)$ houses are owned by white and minority owner-occupiers. The resident structure of the neighborhood is composed of white and minority homeowners and white and minority renters. The model will later allow for separate renter shares r_w and r_m for whites and minorities. The initial minority share of the neighborhood is denoted by m_0.

The model focusses on the moving decisions of agents during one period. If the racial composition of residents who leave the neighborhood differs from the racial makeup of new residents, the neighborhood minority share will change. At the end of the period, all homeowners receive the residual value of their houses. This residual value is equal to the discounted value of an infinite stream of the equilibrium rent that is to be expected from an additional round of housing market transactions in a hypothetical second period.

Two crucial differences between white and minority agents can give rise to neighborhood tipping with falling house prices. Fundamentally, racial segregation occurs because of white agents' preference for white neighbors. Hence, whites' monetary *utility* u_t^w from living in the neighborhood during one period is negative if the neighborhood minority share m_t exceeds the critical threshold $\bar{m} > m_0$.[4] An agent's utility also depends on personal tastes for the neighborhood. Hence, the consumption utility $u_t^{w_H}$ for whites with

[4]The homogeneity in white tolerance of minority residents implies that the sequence of white agents' departure from the neighborhood in a case of white flight must be determined by a factor other than differences in prejudice.

a high taste for the neighborhood is

$$u_t^{w_H} = \begin{cases} v_H & \text{if } m_t \leq \bar{m} \\ -v_H & \text{if } m_t > \bar{m} \end{cases} \tag{4.3.1}$$

Minority residents' utility u_t^m from living in the neighborhood also depends on their tastes but not on racial composition.[5] The residential consumption utility $u_t^{w_H}$ for minorities with a high taste for the neighborhood is

$$u_t^{m_H} = v_H \quad \forall m_t \tag{4.3.2}$$

The utility of agents with a low taste for the neighborhood (L-types) is smaller by a constant k than the utility of agents of the same race who have a high taste for the neighborhood (H-types):

$$u_t^{i_L} = u_t^{i_H} - k \quad \text{where } i = w, m \tag{4.3.3}$$

Because L-types have a strictly lower taste and willingness to pay for the neighborhood given the initial minority share $m_0 < \bar{m}$, the initial population consists entirely of white and minority H-types.

This preference structure implies that absent expectations, mixed-race neighborhoods will be stable for minority shares below $\bar{m}$, but whites will have an incentive to leave once the minority share increases beyond $\bar{m}$. However, when agents are forward looking and anticipate that the minority share can be affected by shocks, whites may start to leave a neighborhood at a critical minority share tipping point that is smaller than $\bar{m}$.

The second difference between whites and minorities is a limitation on the number of potential minority residents with a high taste for the neighborhood so that the neighborhood cannot fill with only H-type minorities:

$$n_{m_H} < N \tag{4.3.4}$$

In contrast, the numbers of H-type whites, L-type whites, L-type minorities, and investors are large and each exceed $2N$.

The distinction between agents with high and low taste for the neighborhood is important when vacancies in the neighborhood fill with new residents. As long as both H-type whites and H-type minorities among potential new residents have equal willingness to pay for housing, there are enough new residents with high taste to fill any number of vacancies

[5]The model could also allow that the residential utility of minorities slightly increases with a higher minority share. Survey data on residential preferences of whites and blacks however suggests that blacks have a weaker own-race preference than whites (Farley, Fielding and Krysan, 1997).

in the neighborhood while L-types are outbid. However, when the neighborhood minority share exceeds $\bar{m}$ and whites are no longer willing to live there, the neighborhood can only fill when house prices and rents fall to the level that minority agents with low taste for the neighborhood are willing to pay.

4.3.2 Resident Turnover and Information

The model covers one period of transactions in the housing market. At the start of each period, a share λ of residents learn that they have to move because of exogenous individual reasons, such as an attractive job offer that makes moving a strictly dominating strategy over staying. The probability of these separations is assumed to be equal across whites, minorities, and investors.

The $(1 - \lambda)$ agents that are not forced to leave have the option of moving voluntarily. However, every agent that moves out of a house must leave the neighborhood and incurs a moving cost c that satisfies

$$c < v_H \tag{4.3.5}$$

Agents cannot change their tenure status within the neighborhood.

When both white and minority H-types have the same willingness to pay for the neighborhood, a proportionate sample of white and minority H-types from outside will fill neighborhood vacancies. However, it is uncertain whether the minority share among these potential new residents will be equal to or larger than that of the neighborhood. The neighborhood's incumbent residents can observe one of two possible signals regarding new residents' racial composition: s_0, which implies that the new resident minority share will be equal to the neighborhood's initial minority share, and s_H, which implies a probability θ of "minority shock," a disproportionate share of minorities among new residents. The probability that incumbents receive the signal s_H is $\gamma < 1$. Based on these signals, the minority share m_1^{new} among new residents will be

$$m_1^{new} \mid s_0 = m_0 \qquad \text{with probability } 1 \tag{4.3.6a}$$

$$m_1^{new} \mid s_H = \begin{cases} m_0 & \text{with probability } 1 - \pi \\ 2m_0 & \text{with probability } \pi \end{cases} \tag{4.3.6b}$$

The model does not rely on a specific underlying type of signal. The incumbent residents who observe events in the neighborhood on a day-to-day basis may for instance observe that a firm which primarily employs minorities is scouting for a business location in the neighborhood.[6]

[6] Zax and Kain (1996) provide evidence for neighborhood racial change as a consequence of company relocation.

Agents living outside the neighborhood, however, have less information than incumbents. Hence, although outsiders know that the incumbents observe a signal s_H with probability γ, they do not themselves receive the signal.[7] Moreover, outsiders have only imprecise information about the initial neighborhood composition: they only observe whether the minority share at the beginning of a period exceeds $\bar{m}$.

Nonetheless, because new residents moving into the neighborhood anticipate the possibility that the neighborhood minority share could increase beyond $\bar{m}$ in the same period, outsiders' expectation for the probability of their new neighborhood having a minority share above $\bar{m}$ is given by

$$E[Prob(m_1 > \bar{m})] \mid no\ signal = \gamma'\pi' \tag{4.3.7}$$

where γ' is the expected probability that a given vacancy has occurred in a neighborhood where incumbents observed signal s_H, and π' is the expected probability that the minority share will increase beyond $\bar{m}$ given signal s_H (see the theoretical appendix for a discussion of some properties of γ' and π'). Although agents outside the neighborhood take into account the possibility of a rising minority share, they receive no signal about the likelihood of such a change in a particular neighborhood. Therefore, potential new residents assume that the probability of a minority share above $\bar{m}$ and the associated reduction in whites' consumption utility is equally large across all potential locations in their choice set.

All agents move simultaneously and agents cannot observe the volume or composition of other agents who depart or move in at the same time. Hence, current residents can only observe the neighborhood minority share once all transactions are complete. Moreover, although agents can observe equilibrium prices, the presence of large groups of agents with identical valuations for houses implies that a change in the volume of sales does not need to result in a changing equilibrium price that would give away incumbents' signal (the section on price determination discusses this property in more detail).[8]

4.3.3 Game Sequence

The following game sequence covers one period of transactions in the housing market.

1. At the start of the period, a share λ of all agents learns that they must sell their

[7] The model would retain the same basic dynamics if the signal were observed by a large share of incumbent residents and by a small proportion of outsiders. The assumption that the signal is observed only by incumbent residents but not by investors who are absentee landlords is not crucial for the tipping dynamic because the non-discriminating investors do not affect the racial composition of the neighborhood.

[8] A plausible alternative setup is to treat the share λ of agents who have to move for exogenous reasons as a random variable. Outsiders could then use observable fluctuations in market volumes and prices to update their prior of γ' but they could not immediately invert the signal from market outcomes.

homes or move out of rental houses because of exogenous individual reasons.

2. Incumbent neighborhood residents receive either the signal s_0 or s_H about the racial composition of potential new residents with high valuation who may move into the neighborhood.

3. Nature randomly determines the racial composition of potential new residents with high taste for the neighborhood according to the probabilities indicated by the signal.

4. Nature randomly determines a share $(1 - r)$ of potential new residents who joins investors in bidding for homeownership and a share r who will bid for rental housing.[9] The fact that whites may be disproportionately selected to bid for homeownership implies that their homeownership rate can exceed that of minorities.[10]

5. The market for house ownership clears. When the number of investors and potential owner-occupiers who are willing to pay exactly the equilibrium price exceeds the number of available houses, a share r of all houses sold are bought by investors and the remaining houses are acquired by owner-occupiers. The racial composition of new owner-occupiers is then determined according to the proportion of whites and minorities in the pool of potential owner-occupiers with equal willingness to pay.

6. The market for rental houses clears. When many white and minority agents are willing to pay exactly the equilibrium rent, the racial composition of new renters is determined according to the proportion of whites and minorities in the pool of potential renters with equal willingness to pay.

7. All departures from the neighborhood and all entries into the neighborhood take place simultaneously.

8. Neighborhood residents observe the resulting neighborhood minority share, m_1. The consumption values that residents obtain from living in the neighborhood are determined according to equations (4.3.1) to (4.3.3). Like all payments for housing, the payoff for the consumption utility accrues at the beginning of the period.

9. At the end of the period, all homeowners are paid the residual value of their houses, which is computed as the discounted value of an infinite stream of the equilibrium rent to be expected from an additional round of housing market transactions.

[9] The separate bidding for homeownership and rental housing allows for a sequential clearing of the markets for homeownership and rental units. Although this assumption can readily be relaxed in the basic version of the model, its relevance in a more general setup is discussed in the subsequent section.

[10] Charles and Hurst (2002) provide evidence for racial discrimination in mortgage lending that could result in a lower probability for minorities to be among bidders for homeownership.

4.3.4 Price Determination

The derivation of prices begins with the computation of the residual house value that homeowners obtain at the end of the period. The residual value, as noted above, is based on the equilibrium rent expected from an additional round of housing market transactions. This rent will be equal to the marginal agent's willingness to pay for rental housing after observing the neighborhood minority share m_1.

When all agents inside and outside the neighborhood observe $m_1 \leq \bar{m}$, both white and minority H-types expect a consumption utility of v_H based on this minority share. If the alternative of living outside the neighborhood yields a utility of zero, H-types of both racial groups are willing to pay a rent of up to v_H, and, because of the large number of white H-types, v_H will be the expected equilibrium rent if $m_1 \leq \bar{m}$.

In contrast, when agents observe $m_1 > \bar{m}$, all whites expect a negative utility from living in the neighborhood. Because whites' disutility from staying in the neighborhood during an additional period is smaller than the moving cost, all whites will leave. However, since there are too few potential minority residents with high valuation to occupy all N houses, the neighborhood must fill with L-type minorities whose residential utility is $v_H - k$. Thus, a neighborhood with minority share $m_1 > \bar{m}$ will shift to all-minority equilibrium with a lower rent level of $v_H - k$.

Accordingly, the residual house price P^{end} that agents obtain at the end of the period depends on whether the minority share m_1 exceeds $\bar{m}$:

$$P^{end} = \begin{cases} P_H \equiv \frac{v_H}{1-\beta} & \text{if } m_1 \leq \bar{m} \\ P_L \equiv \frac{v_H - k}{1-\beta} & \text{if } m_1 > \bar{m} \end{cases} \tag{4.3.8}$$

where β is a discount factor that is equal for all agents.

At the beginning of the period, potential new residents must determine their willingness to pay for rental housing or for homeownership in the neighborhood. The derivation of their willingness to pay invokes two standard nonarbitrage conditions used in many models of location and tenure choice: First, at equilibrium prices, marginal agents are indifferent to whether to locate inside or outside the neighborhood. Second, since homeowners and renters must be indifferent relative to the same outside option, the expected payoffs from being a renter or a homeowner must equate.[11]

The residential utility from living outside the neighborhood in a location with a minority share below $\bar{m}$ is zero, and the likelihood for a minority share above $\bar{m}$ and the associated reduction in whites' payoff is expected to be equal across all potential locations. New residents are thus indifferent between locating inside or outside the neighborhood when the rent in the neighborhood equals their consumption utility. Hence, given an initial

[11] Glaeser and Gyourko (2007) provide a critical assessment of arbitrage conditions in housing markets.

minority share $m_0 < \bar{m}$, both white and minority H-types from outside the neighborhood are willing to pay an equilibrium rent of

$$p = v_H \tag{4.3.9}$$

Likewise, the price that potential new residents are willing to pay for homeownership makes them indifferent to the alternative of living outside the neighborhood, an option that in turn yields the same expected payoff as being a renter in the neighborhood at the equilibrium rent p. New residents' willingness to pay for houses at the beginning of the period is equal to the equilibrium rent p that they would pay as a renter plus the expected residual value of a house obtained at the end of the period. As noted previously, agents from outside the neighborhood expect a probability $\gamma'\pi'$ that their neighborhood will have a minority share above $\bar{m}$ after transactions in the housing market and therefore, according to equation (4.3.8), a residual house value of only P_L instead of P_H.

Investors from outside the neighborhood who are potential new owners of rental houses have the same information as potential new owner-occupiers and therefore expect the same probabilities for the two possible residual house values. Hence, investor willingness to pay for houses is the sum of the expected rental income and the expected residual house value, which equates to owner-occupiers' willingness to pay. The equilibrium house price is thus

$$P \equiv p + \beta[(1 - \gamma'\pi')P_H + \gamma'\pi' P_L] \tag{4.3.10}$$

Note that while the effect of $m_1 > \bar{m}$ on whites' consumption utility is equal for owners and renters, the impact of a high minority share on residual house values only affects homeowners. Therefore, a larger expected probability $\gamma'\pi'$ for a minority share above $\bar{m}$ lowers house prices relative to rents and appropriately compensates home owners for the asset value risk.[12] Absent moving costs, incumbent owner-occupiers will always find it attractive to sell at price P when they obtain the signal s_H and expect that the probability of $m_1 > \bar{m}$ is larger than $\gamma'\pi'$.

It is noteworthy that the equilibrium house price P does not change when a larger fraction of home owners decides to sell upon observing the signal s_H. The presence of groups of agents with equal tastes and equal willingness to pay implies that the demand curve for housing is a step function that is flat over certain intervals of transaction volumes. Given the initial minority share $m_0 < \bar{m}$, there are more than N H-type white and minority agents who are not initial neighborhood residents and who have a high willingness to pay for the neighborhood. The willingness to pay of the marginal new resident is therefore equal irrespective of the share of initial neighborhood residents that chooses to leave. In

[12] Sinai and Souleles (2005) make the related observation that the spread between house prices and rents increases in the volatility of rents if agents are risk-averse.

this setup, the number of departing residents does not affect the equilibrium price and outsiders cannot infer incumbents' signal when they observe that price. Conversely, when agents observe a minority share above $\bar{m}$, the number of agents with a high willingness to pay for housing drops sharply as whites no longer have a positive valuation for living in the neighborhood and H-type minorities cannot fill the neighborhood alone.

It is noteworthy that the sequential clearing of the markets for homeownership and for rental houses can generate a locally flat demand curve for homeownership even if potential residents' tastes for the neighborhood were drawn from a continuum. Idiosyncratic tastes for the neighborhood differentiate the willingness to pay among potential owner-occupiers but not among investors who only have a financial interest in the neighborhood but not an individual taste for living there. Investors' valuation for housing, which is based on expected rental income and expected residual house value, is thus the same for the large number of investors. Hence, unless investors are fully outbid by potential owner-occupiers, the ownership market will always clear at investors' valuation. As a result, there are no price changes that would allow outsiders to infer the signal and eliminate the arbitrage opportunity for incumbent homeowners.[13]

4.3.5 Moving Decisions

The moving decisions of incumbent residents depend on moving costs, prices and rents, and on the expected minority share m_1, which determines consumption utility and the residual values of houses at the end of the period. Moving decisions are also influenced by agents' beliefs about other residents' moving behavior.

Consider first the case where residents determine their moving decisions without taking into account moving decisions of other residents. If residents observe the signal s_0, agents will anticipate no change in the neighborhood minority share because vacancies will be filled with new residents whose racial composition is equal to the initial minority share m_0. All current residents will therefore expect a consumption utility v_H and an end-of-period house value of P_H, and none will have the incentive to move and voluntarily incur the moving cost c.

However, if neighborhood agents receive the signal s_H, which implies a probability π that the minority share among new residents will be biased toward minorities, the moving pattern will differ. When a minority shock occurs and the λ vacancies created by residents forced to move are filled with a group of new residents that includes a share $2m_0$ of minorities, then the neighborhood minority share will increase to

$$m_1 = (1 + \lambda)m_0 \tag{4.3.11}$$

[13] A setup with heterogeneous idiosyncratic tastes and sequential market clearing might however allow the equilibrium rent to fall when a larger number of incumbent residents moves on observing s_H. A falling rent would decrease white renters' incentive to leave.

Proposition 1. *If agents in the neighborhood observe the signal s_H and do not take into account the moving decisions of other residents, there will be a discontinuous change in neighborhood composition at a critical tipping point $m_0^* \equiv \frac{1}{1+\lambda}\bar{m}$. If $m_0 \leq m_0^*$, no residents will move voluntarily. If $m_0 > m_0^*$,*

$$\textit{white owners leave if} \quad c < (\pi - \gamma'\pi')[2v_H + \beta(P_H - P_L)] \tag{4.3.12a}$$

$$\textit{white renters leave if} \quad c < (\pi - \gamma'\pi')[2v_H + \quad 0 \quad] \tag{4.3.12b}$$

$$\textit{minority owners leave if} \quad c < (\pi - \gamma'\pi')[\ 0 \ + \beta(P_H - P_L)] \tag{4.3.12c}$$

while minority renters always stay. White home owners have unambiguously the strongest incentive to leave the neighborhood.

Proof. See Appendix. ■

Intuitively, two factors provide incentives for agents to leave the neighborhood on observing the signal s_H. First, all white residents in a neighborhood with minority share $m_0 > m^*$ have an incentive to leave because their residential utility would fall from v_H to $-v_H$ if the racial composition of new residents were biased toward minorities and the minority share increased above $\bar{m}$. Second, all owners have an incentive to sell because the value of their houses would fall if the minority share increased beyond $\bar{m}$. Thus, even though the equilibrium house price P incorporates prospective buyers' expectation that the residual house value will be P_L with probability $\gamma'\pi'$, incumbents who observe the signal s_H expect a higher probability $\pi > \gamma'\pi'$ of a low residual price and therefore have a financial incentive to sell and leave the neighborhood.

White home owners have the unambiguously strongest incentive to move. If the minority share increases above $\bar{m}$, they suffer both a drop in house values and a reduction of residential consumption utility. By contrast, only one of the two moving incentives applies to either white renters or minority owners while none applies to minority renters.

The following analysis will make two assumptions on the structure of moving incentives. First, suppose that the probability of a minority shock is small enough relative to moving costs that equation (4.3.12a) but not equations (4.3.12b) and (4.3.12c) are fulfilled. The probability of a shock is then small enough relative to moving costs that only white owners choose to leave after observing the signal s_H while white renters and minority residents stay. Second, the analysis will assume $2v_H > \beta(P_H - P_L)$ which implies that white renters have a larger moving incentive than minority owners. This assumption yields a clear ranking of moving incentives by resident group which is useful for the analysis of moving decisions of strategically optimizing agents that will be discussed below.

Under these assumptions, proposition 1 implies that a comparison between neighborhoods with minority shares around m^* shows a discontinuity in net white population change: Neighborhoods with a minority share of $m_0 \leq m^*$ have a stable white population unless a minority shock occurs. In contrast, the mere anticipation of a possible shock is sufficient to lower white population in neighborhoods with a minority share above $m_0 > m^*$ where white owners always depart on observing the signal s_H.[14]

The moving choices of proposition 1 thus yield two important testable predictions: At the tipping point m^*, (i) the population of white homeowners should decrease relative to the population of white renters,[15] and (ii) the magnitude of the tipping discontinuity in white population should be larger in neighborhoods having a large initial homeowner share among whites (or equivalently, a low renter share).

Indeed, in a neighborhood with a large share of home owners, white owner departure may be sufficient to raise the minority share above $\bar{m}$ even when no minority shock occurs. The departure of white owners will then have immediate effects on white renters whose consumption utility falls, and on minority owners who experience a drop in residual house values. When residents behave strategically, their moving decisions will take into account these effects of others' moving decisions on their own payoff.

Proposition 2. *If agents expect that no other residents will leave upon observing s_H, unless staying is strictly dominated by another strategy, there will be a discontinuous change in neighborhood composition at a critical tipping point m_0^*. The probability of $m_1 > \bar{m}$ and falling house values will increase in the initial homeowner share among whites* $(1 - r_w)$*:*

$$Prob(m_1 > \bar{m}) = \begin{cases} 0 & \text{if } m_0 \leq m_0^* & (4.3.13a) \\ \pi & \text{if } m_0 > m_0^* \text{ and } (1 - r_w) \leq (1 - r_w^*) & (4.3.13b) \\ 1 & \text{if } m_0 > m_0^* \text{ and } (1 - r_w) > (1 - r_w^*) & (4.3.13c) \end{cases}$$

where $1 - r_w^* \equiv \frac{\bar{m} - m_0}{(1-\lambda)(1-m_0)m_0}$ *and by assumption,* $1 - r_w^* < r_m$*. If* $m_0 > m_0^*$*, all white owners will leave in case* (4.3.13b) *whereas all white owners, white renters, and minority owners will leave in case* (4.3.13c).

Proof. See Appendix. ■

Proposition 2 shows that neighborhood incumbents will never leave a neighborhood with an initial minority share of $m_0 \leq m_0^*$ because an increase in the minority share

[14]It should be noted that the model could readily accommodate heterogeneity in moving costs. The response to the signal s_H could then be limited to the departure of white owners with low moving costs, thus creating smaller population movements that may be empirically more plausible.

[15]This prediction is so far based on the assumption that some of the houses that are vacated by owner-occupiers will be bought by investors and used as rental houses. It can however also result from heterogeneity in wealth (see following section).

above $\bar{m}$ can be averted when all neighborhood incumbents stay.[16] Conversely, when the initial minority share is above m^*, white owners will always depart. In neighborhoods with large white homeownership, the expectation of a possible increase in minority share beyond $\bar{m}$ can become self-fulfilling: As white owners leave a neighborhood with an owner share above $(1 - r_w^*)$, the minority share rises beyond $\bar{m}$. White renters anticipate this increase in minority share and therefore leave as well to avoid a negative consumption utility. Furthermore, minority owners have an incentive to depart in order to avoid a reduction in house values: Since a large ownership rate weakly increases the probability that the minority share will exceed $\bar{m}$, it by consequence also raises the likelihood that the (residual) price of houses will fall at the tipping point m^*.

An additional third testable prediction of the model is therefore that around the tipping point m^*, (iii) neighborhoods with a large owner share (or small renter share) should be more likely to experience a drop in house prices. The empirical analysis will test this prediction along with the earlier predictions that white owners should be particularly likely to depart at the tipping discontinuity and that neighborhoods with high owner shares and low renter shares should hence experience larger decreases in white population.

4.3.6 Extension: Heterogeneity in Wealth

The model provides a richer set of predictions when it allows for heterogeneity in agents' wealth. This extension of the model assumes the presence of rich and poor agents who have identical tastes for the neighborhood. Rich agents can afford to either buy or rent a house while poor agents are credit constrained and can only rent. The indifference of the rich agents to either owning or renting at equilibrium prices implies that the same share of rich agents can be consistent with different values of the renter share.

Suppose that the share of rich agents among new potential residents equals the proportion of incumbent residents that are wealthy. Thus, if a proportionate sample of incumbents were replaced with new residents, the share of wealthy agents would not change. However, proposition 2 suggests that among both racial groups, owners are more likely than renters to depart when the initial neighborhood minority share exceeds the tipping point m^*. Consequently, the composition of departing residents is biased toward wealthy agents, and the average income level in both racial groups falls when a proportionate sample of rich and poor new residents moves in. Moreover, a reduction in the homeownership rate is necessary when there are fewer new wealthy residents than owner-occupiers who depart. The model thus provides a fourth prediction that (iv) there will

[16]While the expectation structure of proposition 2 is attractive in that it avoids excessive movement of residents, it is theoretically conceivable that incumbents instead expect that other residents will always panic and leave when observing the signal s_H unless leaving is a dominated strategy. This alternative expectation structure is discussed in the theoretical appendix. Its main property is that the moving behavior of whites owners and renters does not differ.

be a discontinuous fall in neighborhood income level at the tipping point m^* when home owners are more likely than renters to depart on observing the signal s_H. Importantly, such a reduction in wealthy population will not be driven by greater racial prejudice of wealthy whites but by a financial incentive for home owners to leave the neighborhood ahead of a possible decrease in house values.

4.4 Empirical Framework

The theoretical model emphasizes that identical expectations about the probability of future neighborhood tipping can translate into different behavior by homeowners and renters. Specifically, it predicts that when the minority share of a neighborhood exceeds a critical tipping point, white homeowners are more likely to depart than white renters. The objective of the empirical analysis is therefore to test whether tipping neighborhoods lose more homeowners than renters, and whether the magnitude of discontinuous changes in neighborhood racial composition and house prices at empirically estimated tipping points varies according to neighborhoods' initial proportions of homeowners and renters. In addition, the analysis tests the predictions that neighborhoods with higher homeownership rates and lower renter shares experience a reduction in house values, and that income levels fall discontinuously at the tipping point. Based on the well-known correlation between income and education, it also tests for a decrease in education levels in tipping neighborhoods. The analysis builds on an empirical identification of tipping points in recent work by Card et al. (2008a).

4.4.1 Data

The neighborhood data is taken from the Neighborhood Change Database (NCDB), which includes census tract data from the 1970, 1980, 1990, and 2000 decennial censuses. These tracts are areas with a population of about 3,000-4,000 individuals and represent neighborhoods with a relatively homogeneous demography and housing stock. The NCDB provides a panel of tracts based on Census 2000 tract boundaries onto which earlier data have been mapped.[17] This analysis focuses on changes in tract characteristics over the three 10-year periods 1970-1980, 1980-1990, and 1990-2000.

The "cities" included in the sample are metropolitan statistical areas (MSAs) that encompass both central city and suburban neighborhoods. The sampling process, which follows Card et al. (2008a), creates separate samples for each 10-year period and excludes tracts having very few residents at the beginning of a decade and MSAs with less than 100

[17] The mapping to boundaries at the end of the sample period is not optimal because of the possible endogeneity of later boundaries. Results for a subsample of tracts without boundary changes in the 1990-2000 period are, however, similar to those reported for the full sample.

Table 1. Average Characteristics of All Tracts and Tracts +/-2% Around Tipping Point, 1970-1990

	A. All Tracts			B. Tracts +/-2% Around Tipping Point		
	1970	1980	1990	1970	1980	1990
	(1)	(2)	(3)	(1)	(2)	(3)
No. of MSAs in Sample	104	113	114	104	113	114
No. of Tracts in Sample	35,725	39,283	40,187	9,306	5,502	5,047
No. of Tracts in 1/3 Sample	11,906	13,091	13,402	3,087	1,834	1,684
Total Population	3333.7	3507.8	3940.8	3233.2	3507.8	3940.8
	(2052.8)	(1763.2)	(1805.2)	(2010.0)	(1763.2)	(1805.2)
Minority Share	0.163	0.235	0.291	0.049	0.082	0.107
	(0.245)	(0.293)	(0.310)	(0.048)	(0.065)	(0.080)
Renter Share	0.347	0.364	0.383	0.295	0.313	0.307
	(0.242)	(0.250)	(0.246)	(0.200)	(0.212)	(0.188)
Renter Share, Whites			0.309			0.293
			(0.207)			(0.177)
Renter Share, Minorities			0.567			0.467
			(0.261)			(0.269)
Log Avg Family Income	10.727	10.804	10.923	10.794	10.880	11.054
	(0.323)	(0.355)	(0.443)	(0.271)	(0.276)	(0.346)
Sh. Families with Income >45'000$, Whites			0.535			0.544
			(0.443)			(0.174)
Sh. Families with Income >45'000$, Minorities			0.380			0.511
			(0.206)			(0.234)
Share of College Graduates	0.123	0.179	0.223	0.140	0.202	0.260
	(0.108)	(0.134)	(0.161)	(0.114)	(0.129)	(0.153)
Share of College Graduates, Whites			0.259			0.262
			(0.166)			(0.156)
Share of College Graduates, Minorities			0.150			0.259
			(0.141)			(0.182)

See data appendix for sample selection. The descriptive statistics include only the 1/3 of the tract sample that has not been used for the estimation of tipping point locations. Data on renter share, income, and education by race is only available for 1990. The race-specific statistics for whites or minorities are weighted by the number of white or minority residents in a tract.

developed tracts (see the data appendix for details of the sample selection). Panel A of Table 1 presents summary statistics for the tract samples from each 10-year period. The sample size grows with each decade from about 36,000 tracts in 104 MSAs for 1970-1980 to about 40,000 tracts in 114 MSAs for 1990-2000.

Throughout the analysis, the term "white population" refers to non-Hispanic whites, while "minority population" refers predominantly to black or Hispanic residents but also includes Native Americans, Asians, Pacific Islanders, and other nonwhite races. The tract tabulations of the decennial censuses, which are integrated into the NCDB data, only separate population counts for non-Hispanic whites as of 1980. Following Card et al.

(2008a), non-Hispanic white shares for 1970 have been imputed by first regressing the non-Hispanic white tract population share in 1980 on the share of whites, blacks, and Hispanics in the same period, and then multiplying the resulting coefficients with the tract share of whites, blacks, and Hispanics in 1970. As Panel A, table 1 shows, the minority share in the sample tracts increased considerably over time, from 16% in 1970 to 29% in 1990.

The NCDB data is most comprehensive for 1990 and 2000 when homeownership, income, and education variables were tabulated separately by race. Prior to 1990, the census either did not report these cross-tabulations, or values were suppressed for many tracts in the sample; the analysis of these outcomes by race therefore focuses on the 1990-2000 period. Likewise, education and income variables for non-Hispanic whites had to be imputed from separate counts for whites and Hispanics (again, the data appendix provides a detailed description of variable definitions and imputations).

4.4.2 Identification of Potential Tipping Points

The empirical analysis uses regression specifications that draw on the literature of regression discontinuity models to identify decadal changes in population composition and housing variables around city-level tipping points. As a prerequisite for this analysis, the location of potential tipping points must be empirically estimated from the data. This study uses the tipping point values that were identified and generously shared by David Card, Alexandre Mas, and Jesse Rothstein.

Card et al. (2008a) report two methods for identifying tipping points which both yield similar tipping point estimates. This present study focuses on the simpler technique, which builds on the methodology for identifying structural breaks. The structural break approach regresses decadal changes in white tract population expressed as a share of a tract's initial population on a constant and a dummy for an initial minority share above m^*, where $m^* \in (0, 50)$. The regression is run separately for each city in each decade using only tracts with initial minority shares of up to 60 percent. The candidate tipping point for each city-decade cell is the value of m^* that maximizes the R^2 of the regression.

While the structural break approach identifies a possible tipping value for every city and decade, the sign and magnitude of a change in white population around this value is not a priori known and has to be estimated in a subsequent, separate step of the analysis that is described in section 4.4 below. That same step of the analysis will also address changes in other outcome variables such as in owner- and renter-occupied housing, house prices, and income and education levels. Nevertheless, if the same data were used to identify tipping point values and estimate the magnitude of discontinuities around these values, the estimated magnitude of changes would be upward biased in absolute value and standard hypothesis tests would too often reject the null hypothesis of zero discontinuity. The analysis therefore builds on a split-sample technique that identifies tipping points using only a randomly selected subsample of 2/3 of all tracts and then estimates the magnitude of tipping discontinuities based on the remaining 1/3. The independent selection of the two subsamples allows the use of conventional hypothesis tests to evaluate the magnitude and statistical significance of tipping discontinuities (Card et al., 2008a).

Based on the subsample of 2/3 of all tracts, the resulting average estimates for tipping points across all cities are minority shares of 9 percent in 1970, 12 percent in 1980, and 14 percent in 1990. The estimated tipping point values are in line with patterns of cross-time, cross-city variation in racial attitudes of the white population. The increase in tipping point values over time is consistent with a growing white tolerance for living with minority neighbors, and Card et al. (2008a) also show that cities whose white residents report more favorable attitudes toward minorities in the General Social Survey tend to have higher tipping points.

Panel B of Table 1 reports descriptive statistics for tracts with a minority share within two percentage points of the estimated tipping points. Tracts near the tipping point have slightly higher income and education levels and a smaller renter share than average tracts in the sample. The population of tracts near the tipping point also differs notably from the average population in terms of education levels by race which are reported in the 1990 data. Whereas the share of residents with college degrees is equal for whites and minorities in neighborhoods near the tipping point, in the overall sample, minorities have considerably lower educational attainment. A similar observation applies for income levels. While average family income by race is not reported in the NCDB, bracketed income data allows to compute the share of families with an income above 45,000 US dollars in 1999 value. That share is similar for white and minority families in neighborhoods near the tipping point (54 vs. 51 percent) while there is a considerably larger racial income gap in the overall sample (54 vs. 38 percent). This observation is noteworthy because it implies that the departure of a random sample of whites from a neighborhood near the tipping point would hardly change the overall income and education levels of the neighborhood.[18] Despite relatively similar education and income levels, minority residents in neighborhoods near the tipping point are more likely than white residents to be renters. These relatively lower homeownership rates for minorities or immigrants are well-documented in the literature (e.g., Collins and Margo, 2001; Coulson, 1999) and may partly result from racial discrimination in the mortgage application process (Charles and

[18] An analysis of census house values by Bayer, Ferreira and McMillan (2007) suggests that homeowners not only value own-race neighbors but also neighbors with similar educational attainment and income levels. If whites had considerably more education than minorities, such preferences would exacerbate the departure of educated whites from a neighborhood once a random group of whites have left. However, because the educational attainment of whites and minorities is similar, it is unclear a priori whether such a dynamic will unfold.

Hurst, 2002).

4.4.3 Distribution and Correlates of the Renter Share

The empirical analysis tests the model's prediction that neighborhoods with a large proportion of homeowners and hence a small share of renters will experience larger discontinuous decreases in white population at the tipping point. The renter share is defined as the share of renter-occupied houses among all occupied housing units. It is by construction equal to one minus the owner share.

Panel A of Table 2 shows percentiles of the distribution of the renter share for neighborhoods with minority shares within two percentage points of the city-level tipping point. There is considerable variation in the neighborhood renter structure, which should facilitate the identification of heterogeneous treatment effects. In each year, the 10th percentile of the renter share falls at about 9 percent renter-occupied housing, while the 90th percentile of the distribution falls near 60 percent.

Panel B of table 2 shows correlates of the renter share. It reports the coefficients from separate regressions of the tract renter share on the indicated tract characteristic, controlling for MSA fixed effects. Tracts with a higher renter share tend to be populated by families with lower average income and fewer children and are more frequently located in central cities than in suburbs. They also have more neighboring tracts with a minority share that exceeds the city-level tipping point.[19] The correlation of renter share with family income and share of families with children raises the possibility that heterogeneity of tipping effects by renter share may be driven by the differential tastes and behavior of rich and poor agents or of families with and without children. The variation in renter share by location within a city could also lead to spurious results if tastes of residents differ by location. The empirical analysis will explicitly account for these factors.

A further noteworthy characteristic of renter share is its extremely strong correlation with renter share among white residents which is not surprising given that neighborhoods near the tipping point are predominantly white. Hence, the identifying variation in renter share used throughout the analysis is plausibly driven primarily by variation in white ownership configurations. Unfortunately, these are not available for the full sample period.

4.4.4 Estimation of Tipping Effects

The magnitude of discontinuities in population and housing outcomes around the potential tipping points is estimated using a regression setup commonly used in regression discontinuity analysis. All estimations use the randomly selected 1/3 of tracts not previously used to identify the locations of the candidate tipping points.

[19]Tract neighbors are defined as tracts whose central point falls within a 3-mile buffer zone around a given tract's boundary.

Table 2. Distribution and Correlates of Renter Share for Tracts with a Minority Share within +/-2% of Tipping Point

	1970 (1)		1980 (2)		1990 (3)	
	A. Distribution of Renter Share					
10th Percentile	0.092		0.091		0.092	
20th Percentile	0.135		0.133		0.141	
50th Percentile	0.236		0.264		0.276	
80th Percentile	0.440		0.487		0.462	
90th Percentile	0.605		0.618		0.571	
	B. Correlates of Renter Share					
Log Avg. Family Income	-0.240	***	-0.176	***	-0.324	***
	(0.032)		(0.021)		(0.010)	
Share of Families with Children under Age 18	-0.835	***	-0.464	***	0.326	***
	(0.062)		(0.061)		(0.055)	
Central City Location	0.129	***	0.109	***	0.215	***
	(0.014)		(0.013)		(0.012)	
Share of Neighboring Tracts beyond Tipping Point	0.165	***	0.06	***	0.215	***
	(0.023)		(0.016)		(0.012)	
Renter Share Whites					1.009	***
					(0.003)	
Renter Share Minorities					0.502	***
					(0.022)	

Each data point in panel B reports the coefficient of a separate regression of renter share on the specified variable and MSA fixed effects. Tract neighbors are defined to have their central point within a 3 mile buffer zone around a given tract's boundary. * p ≤ 0.10, ** p ≤ 0.05, *** p ≤ 0.01.

The basic regression analysis uses a model of the form

$$\Delta y_{ic\tau} = \alpha + p(d_{ict})\beta_1 + \beta_2 1[d_{ict} > 0] + \beta_4 r_{ict} + X_{ict}\beta_5 + \gamma_c + \epsilon_{ic\tau} \tag{4.4.1}$$

where τ stands for the decades 1970-1980, 1980-1990, and 1990-2000, and t is the year at the start of the respective decades (i.e., 1970, 1980, and 1990). The decadal change of an outcome variable, $\Delta y_{ic\tau}$, is regressed on a quartic polynomial $p(d_{ict})$ in the deviation of a tract's minority share from the city-level tipping point, a dummy variable $1[d_{ict} > 0]$ for a tract minority share above the tipping point, a set of city fixed effects γ_c, the renter share r_{ict}, and a vector X_{ict} of neighborhood-level control variables, all measured at the beginning of a decade. Inclusion of city fixed effects implies that all models analyze within-city variation. Reported standard errors are clustered by MSA.

A second regression setup of the form

$$\Delta y_{ic\tau} = \alpha + p(d_{ict})\beta_1 + \beta_2 1[d_{ict} > 0] + \beta_3 1[d_{ict} > 0] * r_{ict} + \beta_4 r_{ict} + X_{ict}\beta_5 + \gamma_c + \epsilon_{ic\tau} \tag{4.4.2}$$

interacts the dummy variable for the tipping discontinuity with the renter share and thus allows for heterogeneity of the treatment effect according to a tract's initial share of renters. The coefficient β_2 for the tipping dummy provides the predicted tipping effect for an all-homeownership neighborhood while the coefficient β_3 for the interaction term yields the predicted difference in tipping effects for neighborhoods with higher renter shares and lower homeownership rates.

4.5 Tipping Discontinuity for Homeowners and Renters

The key prediction of the theoretical model is that homeowners have a stronger incentive than renters to leave a neighborhood in anticipation of immanent racial tipping and falling house prices. The model predicts that the selective departure of homeowners will result in a reduction of owner-occupied housing relative to renter-occupied housing when part of the previously owner-occupied housing units fill with renters. Alternatively, when the assumption of a constant number of occupied houses in the neighborhood is relaxed, a neighborhood may experience a relative reduction in population growth as homeowners depart.

Panel A of Table 3 presents decade-specific estimates of equation (4.4.1) for the change in the number of owner-occupied and renter-occupied housing units at the candidate tipping points. All models include a quartic polynomial in the deviation of tract minority share from the city-specific tipping point, as well as MSA fixed effects. The regressions also control for a tract's renter share at the beginning of a decade, and for five additional tract characteristics in the base year: the log of the mean family income, the share of families with children under 18, the unemployment rate, and the fraction of vacant and single-unit housing units in a tract.

In each of the three decades, owner-occupied housing decreases significantly by 6-8 log points at the tipping discontinuity. In contrast, the change in renter-occupied housing does not display a discontinuous change. These results for owner- and renter-occupied housing imply a reduction in the growth rate of the housing stock around the tipping point. The table also reports the results of Wald tests for the equality of the coefficients for owning and renting derived from fully interacted regressions that combine the separate regressions for owner- and renter-occupied housing. For all three decades, the null hypothesis of an equally large decrease in owner- and renter-occupied houses can be rejected at least marginally.

The lower panel of table 3 reports separate estimates of the change in owner- and renter-occupied housing for whites and minorities which can be computed based on the race-specific data for 1990-2000. The results for whites indicate a significant drop in white homeownership while there is no apparent decrease in the number of houses occupied by

Table 3. Change in Owner-Occupied and Renter-Occupied Housing Units. Dependent Variables: Log Change in Owner-Occ or Renter-Occ Housing Units

	I. All Races, 1970-2000					
	A. 1970-1980		B. 1980-1990		C. 1990-2000	
	Own	Rent	Own	Rent	Own	Rent
Beyond Tipping Point	-0.073 ***	-0.019	-0.080 ***	-0.026	-0.056 ***	0.015
	(0.025)	(0.024)	(0.023)	(0.025)	(0.013)	(0.017)
Wald Test Own=Rent	F(1,103)=3.19 * p=0.08		F(1,112)=4.07 ** p=0.05		F(1,113)=15.35 *** p=0.00	
n	11,806	11,581	12,931	13,042	13,229	13,333
R^2	0.34	0.29	0.23	0.34	0.14	0.07

	II. By Race, 1990-2000			
	A. Whites		B. Minorities	
	Own	Rent	Own	Rent
Beyond Tipping Point	-0.056 ***	-0.005	-0.038	-0.021
	(0.016)	(0.022)	(0.027)	(0.032)
Wald Test Own=Rent	F(1,113)=5.74 ** p=0.02		F(1,113)=0.20 p=0.66	
n	12,622	12,687	12,580	11,568
R^2	0.29	0.13	0.20	0.18

All models control for MSA fixed effects, a quartic polynomial for the difference between a tract's minority share and the estimated MSA tipping point at the beginning of a decade. They also control for renter share, log average family income, share of families with children under age 18, unemployment rate, share of vacancies, and share of single-unit homes in a tract at the beginning of a decade. Robust standard errors in parentheses are clustered by MSA. * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

white renters. The point estimates for minorities suggest a modest decrease in owner-occupied relative to renter-occupied housing, but the imprecise estimate for the decrease in minority homeownership is not significantly different either from zero or from the change in minority-occupied rental units.

These results are consistent with the prediction that neighborhoods primarily lose white homeowners at the tipping point. While white owner-occupied housing discontinuously falls, there is no evidence for a reduction in white renter-occupied housing.

4.6 Discontinuity in Racial Composition

The magnitude of racial tipping — the rapid reduction in white relative to minority population in neighborhoods whose minority share exceeds a critical tipping point — should depend on a neighborhood's tenure structure. If the departure of white residents from tipping neighborhoods is mostly confined to white homeowners, then neighborhoods with higher homeownership rates and smaller renter shares should experience larger outflows in white population. This section presents estimates for the change of white, minority, and overall population at empirical tipping points, allowing for heterogeneity of tipping effects by initial renter share of a neighborhood. The analysis also evaluates a series of alternative hypotheses that provide potential explanations for a relationship between neighborhood tenure structure and the magnitude of the reduction in white population at the tipping discontinuity.

4.6.1 Baseline Results and Robustness Tests

The first four columns of Table 4 present estimates for the decadal change in white population as a fraction of initial tract population, alternatively excluding and including a vector of tract-level demography and housing controls. The estimated coefficients in columns (1) and (3) corroborate that the growth rate of the white population is discontinuous around the estimated tipping points; a finding documented by Card et al. (2008a). The regressions that include demographic and housing controls, reported in column (3), show statistically significant discontinuous declines in the growth rate of white population by -10, -11, and -9 percentage points for the three decades 1970-1980, 1980-1990, and 1990-2000.

While the models reported in column (1) and (3) show mean tipping effects across all tracts, those shown in columns (2) and (4) are based on the regression equation (4.4.2) that allows for variation in the magnitude of the tipping discontinuity by including an interaction term between the tipping dummy and the renter share of a tract in the base year. The positive and highly significant coefficients of these interaction terms confirm that the discontinuous decrease in white population at the candidate tipping points is smaller for neighborhoods with a high renter share. Conversely, neighborhoods with a large proportion of homeowners experience a larger decline in white population.[20]

The coefficient of the tipping dummy in the interacted specification reports the predicted tipping effect for neighborhoods with a homeowner share of 100%. In each decade, the coefficient estimates imply that this predicted discontinuity in white population growth for pure-ownership neighborhoods is roughly twice as large as the main effect for the whole sample. To facilitate the interpretation of the results in column (4), Table 5 reports the predicted tipping effects for tracts at different percentiles of the renter share distribution. The white population is predicted to fall by about 15-18 percentage points in the owner-dominated neighborhoods at the 10th percentile of the renter share distribution. There is only a modest decrease of 0-4 percentage points for tracts at the 90th percentile of the

[20]All models in table 4 control for a fourth-order polynomial in the difference between a tract's initial minority share and the city-level tipping point. Appendix Table A1 shows that the estimated discontinuities in white population are very similar for models that include polynomials of lower or higher order. A model with separate polynomials on both sides of the tipping point estimates a larger mean tipping effect in the 1970s.

Table 4. Decadal Change in Overall, White, and Minority Population. Dependent Variable: Change of Indicated Population in Percentage Points of Initial Total Tract Population

	White Non-Hispanics				Minorities		Total Population	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				A. 1970-1980				
Beyond Tipping Point	-12.28 ***	-24.72 ***	-9.92 ***	-17.51 ***	-0.53	1.17	-10.45 ***	-16.34 **
	(3.79)	(5.69)	(3.47)	(5.35)	(1.37)	(2.54)	(3.72)	(5.93)
Beyond TP x Renter Share		39.12 ***		23.58 **		-5.28		18.30 *
		(8.92)		(9.64)		(4.94)		(10.97)
R^2	0.20	0.21	0.27	0.27	0.23	0.23	0.25	0.25
				B. 1980-1990				
Beyond Tipping Point	-12.45 ***	-25.76 ***	-11.25 ***	-19.92 ***	0.60	2.34	-10.65 ***	-17.58 **
	(3.85)	(5.98)	(3.44)	(5.54)	(1.06)	(1.93)	(3.94)	(6.29)
Beyond TP x Renter Share		40.78 ***		26.48 ***		-5.29		21.19 **
		(8.06)		(7.72)		(3.77)		(8.83)
R^2	0.23	0.24	0.32	0.32	0.27	0.27	0.32	0.32
				C. 1990-2000				
Beyond Tipping Point	-9.81 ***	-21.86 ***	-8.52 ***	-19.37 ***	1.70 **	5.72 ***	-6.83 ***	-13.64 **
	(1.83)	(2.70)	(1.76)	(2.61)	(0.79)	(1.22)	(2.18)	(3.18)
Beyond TP x Renter Share		38.64 ***		34.55 ***		-12.84 ***		21.71 **
		(4.55)		(4.57)		(2.41)		(5.36)
R^2	0.12	0.13	0.16	0.17	0.19	0.19	0.15	0.15
4th Polynomial Minority Sh	y	y	y	y	y	y	y	y
Renter Share	y	y	y	y	y	y	y	y
MSA Fixed Effects	y	y	y	y	y	y	y	y
Demogr/Housing Controls	n	n	y	y	y	y	y	y

N=11,886 in 1970-1980; N=13,077 in 1980-1990; N=13,378 in 1990-2000. Regressions include only tracts that have not been used to compute candidate tipping points. All models control for initial renter share, MSA fixed effects, and a quartic polynomial for the difference between a tract's minority share and the estimated MSA tipping point at the beginning of a decade. Models in columns 3-8 also control for log average family income, share of families with children under age 18, unemployment rate, share of vacancies, and share of single-unit homes in a tract at the beginning of a decade. Robust standard errors in parentheses are clustered by MSA. * p ≤ 0.10, ** p ≤ 0.05, *** p ≤ 0.01.

renter share. These results confirm that the magnitude of the tipping discontinuity varies strongly with the initial renter share: neighborhoods with high homeownership rates experience substantially larger drops in white population around candidate tipping points.

The remainder of table 4 estimates discontinuities in the growth rate of the minority and the overall population. Columns (5) and (6) show that the discontinuous increase in the growth rate of the minority population is quite small and only statistically significant in 1990-2000. The growth in minority population is somewhat larger in neighborhoods with large ownership rates that lose more white residents. The strong decrease in white population and the modest increase in minority population combine to produce a discontinuous decrease in overall population around candidate tipping points. This decline in population growth is larger in neighborhoods with more homeownership and a smaller renter share. One possible explanation for such a decrease is a drop in the supply of new

Table 5. Predicted Change in White Population at Different Percentiles of the Renter Share Distribution.

	I. 1970-80	II. 1980-90	III. 1990-00
	(1)	(2)	(3)
Percentile of Renter Share			
10th Percentile	-15.34	-17.50	-16.20
Median	-11.95	-12.92	-9.82
90th Percentile	-3.25	-3.56	0.37

Predicted values are obtained by evaluating the regression coefficients from column 4 of table 4 at the 10th, 50th, and 90th percentiles of the renter share distribution, which are reported in table 2.

houses once a neighborhood tips.[21]

The results in Table 6 provide further evidence for the robustness of the discontinuity in white population growth at candidate tipping points. These models control more flexibly for the initial renter share of a tract using a fourth order polynomial. The resulting model is then augmented with controls for 10-year lags in the renter and minority share (lags that can only be determined for the later two decades). A final set of models also controls for 10-year lags in the demographic and housing control variables. Although the addition of this rich set of control variables attenuates the estimated coefficients, all models continue to find a significant decrease in the growth rate of white population at candidate tipping points, one that is larger for neighborhoods with high homeownership rates and small renter shares.

4.6.2 Alternative Hypotheses

The results of Tables 4 to 6 provide robust evidence that neighborhoods with large homeownership rates experience larger drops in white population at the tipping discontinuity. While homeowners may be particularly inclined to depart from a tipping neighborhood in order to avoid a potential decrease in house values, it is also conceivable that the differential behavior of owners and renters is driven by differences in tastes. Notably, the departure of owners may be driven by Tiebout (1956) sorting if a change in neighborhood racial composition affects public goods of a neighborhood that are primarily consumed by owners. These concerns are particularly justified in view of the systematic relationship between renter share and various measures of population composition and tract location, as previously reported in table 2. This section will therefore address a variety of alternative explanations for the selective departure of homeowners from tipping neighborhoods.

[21] Card et al. (2008a) show that outflows of white population are approximately offset by inflows of minority population in tracts where a scarcity of undeveloped land restricts potential growth in the housing stock.

Table 6. Decadal Change in White Population - Robustness Tests. Dependent Variable: Change of White Population in Percentage Points of Initial Total Tract Population

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				A. 1970-1980				
Beyond Tipping Point	-9.92 ***	-17.51 ***	-9.72 ***	-14.54 ***				
	(3.47)	(5.35)	(3.44)	(5.27)	n/a		n/a	
Beyond TP x Renter Share		23.58 **		15.08				
		(9.64)		(9.49)				
				B. 1980-1990				
Beyond Tipping Point	-11.25 ***	-19.92 ***	-10.62 ***	-16.45 ***	-10.32 ***	-16.12 ***	-6.88 ***	-11.17 *
	(3.44)	(5.54)	(3.44)	(5.78)	(3.63)	(6.12)	(2.58)	(4.31)
Beyond TP x Renter Share		26.48 ***		17.71 **		17.42 **		12.50 *
		(7.72)		(8.57)		(8.78)		(6.69)
				C. 1990-2000				
Beyond Tipping Point	-8.52 ***	-19.37 ***	-7.93 ***	-15.96 ***	-8.12 ***	-15.45 ***	-7.42 ***	-13.64 **
	(1.76)	(2.61)	(1.73)	(2.77)	(1.72)	(2.65)	(1.34)	(2.18)
Beyond TP x Renter Share		34.55 ***		25.35 ***		23.15 ***		19.68 **
		(4.57)		(5.63)		(4.94)		(3.93)
4th Polynomial Renter Sh	n	n	y	y	y	y	y	y
Lagged Minority, Renter Sh -10y	n	n	n	n	y	y	y	y
Lagged Controls -10y	n	n	n	n	n	n	y	y

All models control for initial renter share, MSA fixed effects, a quartic polynomial for the difference between tract minority share and the estimated MSA tipping point, as well as the control variables log average family income, share of families with children under age 18, unemployment rate, share of vacancies, and share of single-unit homes, all measured at the beginning of a decade Models in columns 3-8 also control for a quartic polynomial in the difference between tract renter share and MSA mean. Models in column 5-8 sequentially add ten-year lags of tract minority share and renter share, and ten-year lags of log average family income, share of families with children under age 18, unemployment rate, share of vacancies, and share of single-unit homes. Robust standard errors in parentheses are clustered by MSA. * p ≤ 0.10, ** p ≤ 0.05, *** p ≤ 0.01.

Consider first the hypothesis that home owners are more likely to leave a tipping neighborhood because of their superior ability to afford housing in an alternative location. Indeed, in a study of suburbanization during 1960-1980, Boustan (2007a) identifies a price premium for houses in largely white suburban neighborhoods compared to houses of similar quality in racially mixed inner-city neighborhoods. Even if rich whites are not more racially prejudiced than poorer whites, they may nevertheless be more likely to leave a tipping neighborhood when living in a predominantly white neighborhood is a normal good. It is therefore conceivable that the observed stronger tipping effects in owner-dominated neighborhoods are driven by the higher income levels of such neighborhoods rather than homeownership.

As a first test to evaluate this alternative hypothesis, Table 7 explores whether the discontinuity in white population growth at candidate tipping points remains larger for neighborhoods with high homeownership rates when the sample is stratified into subsamples of tracts that are relatively homogeneous in income levels. Specifically, it groups all neighborhoods into five subsamples of equal size according to their average family income. Columns (1) and (2) report estimated coefficients for neighborhoods whose mean family income puts them into the top 20% of the sample, while each subsequent pair of

Table 7. Decadal Change in White Population by City-Level Quintiles of Tract Mean Average Family Income. Dependent Variable: Decadal Change of White Population in Percentage Points of Initial Total Tract Population

	Tracts sorted by Mean Average Family Income (Rich to Poor)									
	1st Quintile of MSA		2nd Quint. of MSA		3rd Quintile of MSA		4th Quintile of MSA		5th Quintile of MSA	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					A. 1970-1980					
Beyond Tipping Point	-1.5	-5.6	-12.5	-7.6	-13.2 **	-27.6 ***	-5.2	-13.0	-4.8	-34.2 *
	(7.4)	(9.4)	(8.7)	(8.9)	(6.2)	(9.9)	(5.8)	(10.0)	(4.6)	(10.6)
Beyond TP x Renter Share		15.0		-18.2		44.6 **		20.4		64.5 *
		(19.0)		(17.7)		(21.0)		(15.5)		(18.7)
					B. 1980-1990					
Beyond Tipping Point	-2.3	-7.5	-16.5 **	-22.8 **	-14.0 ***	-20.2 ***	-4.0	-7.9	-6.9	-24.2 *
	(9.6)	(13.2)	(6.8)	(9.7)	(4.1)	(5.3)	(4.1)	(7.0)	(4.3)	(7.9)
Beyond TP x Renter Share		20.5		22.2 *		18.5		9.7		36.0 *
		(20.0)		(13.1)		(12.5)		(10.6)		(11.4)
					C. 1990-2000					
Beyond Tipping Point	-10.5 **	-21.4 ***	-9.0 **	-16.7 ***	-5.4 **	-13.1 ***	-6.4 **	-17.6 ***	-1.2	-17.4 *
	(4.2)	(6.3)	(4.0)	(5.4)	(2.7)	(3.2)	(2.5)	(4.4)	(3.3)	(5.5)
Beyond TP x Renter Share		41.4 ***		25.3 **		23.4 ***		29.3 ***		34.6 *
		(13.1)		(12.4)		(6.0)		(7.9)		(10.4)

All models use the baseline specification with demographic and housing controls as described in the table notes of table 4. Robust standard errors in parentheses are clustered by MSA. * p ≤ 0.10, ** p ≤ 0.05, *** p ≤ 0.01.

columns refers to subsamples with lower incomes. Strikingly, the results in table 7 show the point estimates for mean tipping effects to be negative in all 15 subsamples, whereas the interaction term of the tipping point dummy and the renter share is positive in all subsamples with the exception of the second income quintile in 1970-1980. The majority of the estimated coefficients, including all estimates for 1990-2000, are also statistically significant despite the smaller sample size in these regressions. Conversely, there is no evidence for particularly strong tipping effects among the tracts in the top income quintile which together with the second and third quintile accounts for a disproportionate share of neighborhoods near the tipping point. These results provide evidence against the hypothesis that the larger tipping effects in ownership-dominated neighborhoods are due to higher income levels of such neighborhoods.

Another possible explanation for larger tipping effects in neighborhoods with high homeownership rates is that residents with greater sensitivity to racial change might be particularly concentrated among homeowners. These residents' aversion to racial change may stem not only from racial prejudice but also from a preference for public goods and neighborhood amenities that may be affected by a higher minority share. In particular, it is conceivable that parents worry about their children's heightened exposure to changes in the neighborhood: School-age children may be be affected by a rising minority share both through a change in the composition of their classmates and possibly through changes in

Table 8. Decadal Change in White Population - Alternative Hypotheses. Dependent Variable: Decadal Change of White Population in Percentage Points of Initial Total Tract Population

	A. 1970-1980			B. 1980-1990			C. 1990-2000		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Beyond Tipping Point	-17.5 ***	-20.6 ***	-11.8 **	-19.9 ***	-20.6 ***	-15.0 ***	-19.4 ***	-22.2 ***	-14.2 **
	(5.3)	(5.7)	(4.7)	(5.5)	(5.8)	(5.2)	(2.6)	(2.6)	(2.6)
Beyond TP x Renter Share	23.6 **	27.3 **	21.1 **	26.5 ***	21.0 **	27.9 ***	34.5 ***	35.9 ***	31.6 **
	(9.6)	(12.1)	(9.6)	(7.7)	(9.2)	(8.5)	(4.6)	(4.4)	(4.8)
Beyond TP x Ln Avg. Family Income		34.0 ***			10.4 **			9.0 ***	
		(6.5)			(4.1)			(2.6)	
Beyond TP x Share Families with Children		-29.0			-70.0 ***			-66.1 ***	
		(17.6)			(14.0)			(12.1)	
Beyond TP x Central City Location			-3.7			-6.4			1.7
			(4.2)			(4.0)			(2.4)
Beyond TP x Share of Tipped Neighbor Tracts			44.9 ***			15.4 **			-3.0
			(6.7)			(6.3)			(2.8)

All models use the baseline specification with demographic and housing controls as described in the table notes of table 4. Models in columns 2-3, 5-6, and 8-9 include interaction terms between estimated MSA tipping points and log average family income, share of families with childeren below age 18, central city location, and share of neighboring tracts with a minority share above the tipping point, all measured in deviations from MSA means; as well as main effects for the interacted variables. Tract neighbors are defined to have their central point within a 3 mile buffer zone around a given tract's boundary. Robust standard errors in parentheses are clustered by MSA. * p ≤ 0.10, ** p ≤ 0.05, *** p ≤ 0.01.

the quality of neighborhood amenities such as schools or public parks.[22] As a consequence, families with children may be particularly inclined to leave a tipping neighborhood.

The hypotheses that the large tipping discontinuities for white population in owner-dominated neighborhoods are due to higher income levels or a larger share of families with children is assessed in columns (2), (5), and (8) of Table 8 where the baseline model specifications are augmented with interaction terms between the tipping dummy and (i) the tract log average family income and (ii) the share of families in a tract with children under 18, both measured as a deviation from MSA means. The estimates show that the result for the renter share interaction term is robust to the inclusion of these additional variables.

While the discontinuous drop in white population is larger in neighborhoods with higher homeownership rates, it is actually somewhat smaller in neighborhoods with higher income levels. The latter result seems consistent with survey evidence that wealthier whites express a greater willingness to live in racially mixed neighborhoods (Farley et al., 1997). The positive correlation of wealth and homeownership, however, implies that despite higher tolerance for minorities, wealthy whites might still be the first to leave a tipping neighborhood to avoid decreasing house values.

Neighborhoods with a larger share of families with children indeed show larger tipping

[22]A large body of research examines the effects of racial segregation across schools and of racial composition of peers on student achievement. See, for example, Guryan (2004), Angrist and Lang (2004), Card and Rothstein (2007), and Vigdor and Ludwig (2007).

discontinuities in white population. This finding is consistent with the notion that families with children might react more sensitively to an increase in minority share. The results are also in line with recent work by Baum-Snow and Lutz (2008) who find evidence for an outflow of white population from central city school districts after court-ordered desegregation in the 1970s. The effect of the share of families with children on the magnitude of the predicted tipping discontinuity is sizable. Based on the coefficients for 1990-2000, a one standard deviation increase in the share of families with children, measured for tracts near the tipping point, predicts an additional decrease in white population of 7 percentage points at the tipping point. An equally large drop in white population results from a one standard deviation decrease in the renter share. In contrast, a one standard deviation increase in log average family income mitigates the drop in white population by 3 percentage points.

A third alternative explanation for the relationship between tenure structure and tipping effects is based on the observation that neighborhoods with a higher renter share differ from owner-dominated neighborhoods not only with regard to resident composition but also in terms of spatial location within a city. In particular, as table 2 shows, central city neighborhoods and neighborhoods in the proximity of tracts with a minority share above the tipping point tend to have larger renter shares. If whites in different areas of a city vary in terms of their racial preferences or their taste for public goods that are affected by the racial composition of a neighborhood, then spatial variation in homeownership can proxy for variation in tastes. Specifically, the suburbs and areas far from previously tipped neighborhoods might attract white residents who are more sensitive to racial composition and who locate away from central cities and areas close to tipped neighborhoods because these areas may be more likely to attract minority residents.[23] Such a spatial sorting by tastes could then predict stronger tipping effects in suburbs and areas far from tipped neighborhoods which both tend to have relatively high homeownership rates.

To evaluate this hypothesis, columns (3), (6), and (9) of table 8 again report the estimations for the baseline model for change in white population augmented with interaction terms between the tipping point dummy and (i) central city location and (ii) share of tract neighbors with minority shares above the tipping point, measured in deviations from MSA means. Tract neighbors are defined as tracts whose central point is within a three-mile buffer zone around a given tract's boundary. Consistent with earlier findings (Card et al., 2008a), the results indicate no systematic difference in the magnitude of the tipping discontinuity between central-city and suburban neighborhoods. For the first two decades, the decrease in white population at the estimated tipping point is larger in

[23]The rapid suburbanization in the post-war era has been extensively analyzed; see, for example, Margo (1992), Bajari and Kahn (2005), Baum-Snow (2007), and Boustan (2007b). A process of circular ghetto expansion in which an existing ghetto spreads to adjacent neighborhoods is described by Moebius and Rosenblat (2001).

neighborhoods with few tipped neighbors. This relationship no longer appears, however, in the 1990-2000 decade. Conversely, the magnitude of the tipping discontinuity continues to be larger in neighborhoods with large homeownership in all three decades.

In summary, the results in tables 7 and 8 suggest that the larger tipping discontinuities in neighborhoods with high homeownership rates cannot be explained by a number of apparent correlates of the renter share. While the possibility that tastes differ between homeowners and renters cannot be conclusively ruled out, the evidence suggests that variation in neighborhood income level, the share of families with children, and neighborhood location are not responsible for the strong tipping effects in owner-dominated neighborhoods.

4.7 Discontinuity in House Prices and Rents

The results so far establish that tipping neighborhoods tend to lose white homeowners rather than renters, and that tipping effects are larger in magnitude in owner-dominated neighborhoods even when controlling for various measures of neighborhood composition and location that are related to owner and renter shares. The theory outlined in section 3 predicts that the selective departure of homeowners may be driven by the risk of a drop in house prices when the neighborhood minority share exceeds the tipping point. Although the data does not allow to directly assess whether departing homeowners expected house prices to fall, the analysis is able to test the model prediction that declines in house prices should be more pronounced in the neighborhoods with large homeownership that experience stronger declines in white population.

Table 9 presents the results for the change in the log average house values and log average rents at the tipping discontinuity.[24] Columns (1) and (3) of table 10 present the estimates from the regression discontinuity models that use the same specification as columns (1) and (3) in table 4. In line with results by Card et al. (2008a), the estimated discontinuous declines in house values of 1-2 log points are small and mostly insignificant. Columns (2) and (4) show the results from augmented models that include an interaction term between the tipping dummy and the renter share. The outcomes for 1980-1990 and 1990-2000 indicate that the magnitude of price drops at candidate tipping points varies significantly with neighborhood renter share. Neighborhoods with high ownership rates do not only experience a greater decline in white population, as previously shown, but also a significantly larger reduction in house values. According to the coefficients in column (2), ownership-dominated neighborhoods at the 10th percentile of the renter share

[24]House values and rents in the census are self-reported and may be affected by misreporting and long adjustment lags (Bayer et al., 2007), as well as by noisy measurement due to tabulation at intervals. These characteristics may make it more difficult to observe substantial price changes as neighborhoods tip.

Table 9. Decadal Change in Average House Values and Rents. Dependent Variable: Change of Log Average House Value or Rent

	I. House Values				II. Rents	
	(1)	(2)	(3)	(4)	(5)	(6)
				A. 1970-1980		
Beyond Tipping Point	-0.008	-0.027	-0.004	-0.015		
	(0.013)	(0.022)	(0.013)	(0.019)	n/a	
Beyond TP x Renter Share		0.059		0.034		
		(0.053)		(0.045)		
n	9,752	9,752	9,752	9,752		
R^2	0.39	0.39	0.40	0.40		
				B. 1980-1990		
Beyond Tipping Point	-0.024 **	-0.059 ***	-0.023 **	-0.065 ***	-0.023 **	-0.054 ***
	(0.010)	(0.015)	(0.010)	(0.015)	(0.011)	(0.014)
Beyond TP x Renter Share		0.111 ***		0.130 ***		0.092 ***
		(0.032)		(0.033)		(0.026)
n	12,221	12,221	12,220	12,220	12,273	12,273
R^2	0.63	0.63	0.63	0.64	0.20	0.20
				C. 1990-2000		
Beyond Tipping Point	-0.016	-0.039 **	-0.012	-0.031 *	-0.008	-0.005
	(0.010)	(0.017)	(0.009)	(0.016)	(0.007)	(0.011)
Beyond TP x Renter Share		0.074 **		0.060 *		-0.012
		(0.036)		(0.035)		(0.019)
n	13,024	13,024	13,024	13,024	13,309	13,309
R^2	0.49	0.49	0.50	0.50	0.15	0.15
Renter Share	y	y	y	y	y	y
Demogr/Housing Controls	n	n	y	y	y	y

All models control for initial renter share, MSA fixed effects, and a quartic polynomial for the difference between a tract's minority share and the estimated MSA tipping point at the beginning of a decade. Models in columns 3-6 also control for log average family income, share of families with children under age 18, unemployment rate, share of vacancies, and share of single-unit homes in a tract at the beginning of a decade. Robust standard errors in parentheses are clustered by MSA. Data on average rents is missing for most tracts in 1970. * p ≤ 0.10, ** p ≤ 0.05, *** p ≤ 0.01.

distribution experienced a decrease in house value of -4.9 and -3.2 log points in 1980-1990 and 1990-2000, respectively. Conversely, the expected price change in renter-dominated neighborhoods at the 90th percentile of the renter share is zero for both decades. The coefficients for 1970-1980 show a qualitatively similar pattern, but they are smaller and less precisely estimated than those for later decades.

The theoretical model predicts that once whites' taste for the neighborhood declines, rents should fall along with house prices. Indeed, the results for rents in columns (5) and (6) show a discontinuous decline of rents for the period 1980-1990, one that is equal in magnitude to the fall in house values. However, the estimates for 1990-2000 provide no such evidence for a significant discontinuity in the average rent. The data for 1970-1980 can unfortunately provide little further evidence for rental change at the tipping discontinuity because rents in 1970 are only reported for a small subsample of tracts.

In summary, although the responses of house prices and rents to neighborhood tipping tend to be relatively modest and imprecisely measured, the results provide support for the hypothesis that the magnitude of price declines increases with the neighborhood homeownership rate. Most particularly, a higher ownership rate and lower renter share predict both larger discontinuous declines in white and overall population at candidate tipping points and larger decreases in house values.

4.8 Discontinuities in Income and Education

The results of the previous sections are consistent with the theoretical model's prediction that tipping neighborhoods primarily lose homeowners who might decide to leave in order to avoid falling house prices. Due to the correlation of homeownership and wealth, whites with higher incomes might be particularly likely to depart from a neighborhood that is beyond the tipping point even though survey evidence firmly suggests that wealthier whites are less racially prejudiced than their poorer peers.[25] An extension of the model that allows heterogeneity in wealth therefore predicts that the income level of a neighborhood should fall at the tipping discontinuity. Moreover, the positive correlation between education and income implies that the reduction in average income might be associated with a decrease in average educational attainment. The following two subsections consider the change in income and education levels at the tipping discontinuity.

4.8.1 Income

The first panel of Table 10 provides estimates for the change in the log average family income of a tract at the tipping discontinuity. The figures reported in columns (1), (3), and (5) indicate that the mean tipping effect is a -2.2, -2.8, and -1.3 log point decrease in average family income in 1970-1980, 1980-1990, and 1990-2000, respectively. The change in population composition required to generate such decreases in average income can be illustrated using the mean tipping discontinuity in population of -11, -11, and -7 percentage points given for the three decades in column (5), table 4. If the only population change at the tipping point were the departure of the indicated proportions of residents, the departing residents would need an income about 20 to 25 percent above the income of the remaining residents to generate the observed decreases in average income levels.

The results reported in columns (2), (4), and (6) of table 10 further indicate that the

[25] There is both theoretical and empirical work that identifies income and wealth as critical determinants of homeownership; see, for example, Henderson and Ioannides (1983), Jones (1990), and Fu (1991).

Table 10. Change in Income Levels. Dependent Variables: Change Log Avg Family Income; Change Share of Families Above Income Threshold

	I. Change Log Avg. Family Income, All Races					
	A. 1970-1980		B. 1980-1990		C. 1990-2000	
	(1)	(2)	(3)	(4)	(5)	(6)
Beyond Tipping Point	-0.022	-0.058 ***	-0.028 ***	-0.040 ***	-0.013 **	-0.032 ***
	(0.016)	(0.015)	(0.007)	(0.011)	(0.006)	(0.012)
Beyond TP x Renter Share		0.113 ***		0.036		0.061 *
		(0.038)		(0.030)		(0.031)
n	11,863	11,863	13,052	13,052	13,357	13,357
R^2	0.16	0.16	0.32	0.32	0.15	0.15

	II. Share of Families with Annual Income >45K$			
	A. Whites, 1990-2000		B. Minorities, 1990-2000	
Beyond Tipping Point	-0.009 **	-0.010 *	-0.003	-0.008
	(0.004)	(0.005)	(0.007)	(0.011)
Beyond TP x Renter Share		0.005		0.016
		(0.010)		(0.023)
n	13,155	13,155	13,109	13,109
R^2	0.07	0.07	0.04	0.04

All models use the baseline specification with demographic and housing controls as described in the table notes of table 4. Robust standard errors in parentheses are clustered by MSA. * p ≤ 0.10, ** p ≤ 0.05, *** p ≤ 0.01.

drop in income level is larger for neighborhoods with large homeownership. For instance, the predicted reduction in average income for neighborhoods at the 10th percentile of the renter share is between -3 and -5 log points for each of the three decades. It should also be noted that the estimated decreases in average income are commensurate to the declines in house values.

Even though the NCDB does not report the variable for average family income separately by race, it does provide the number of families in broad income brackets for 1990 and 2000. Because transforming these interval data into a continuous variable is challenging given the sensitivity of assumptions about family distribution within income brackets, the analysis of income change by race focuses on the change in the share of white and minority families with an income above 45,000 dollars. Specifically, it exploits the fact that the real value of 45,000 dollars in 1999 is almost precisely equal to the value of 35,000 dollars in 1989.[26] Since these two numbers form borders of reported income intervals in the respective years, the data allows to consistently observe the share of white and minority families with incomes over 45,000 dollars. The lower panel of table 10 shows a significant decrease in the share of white residents with an income above 45,000 dollars

[26]Census income data refers to the year preceding the survey.

at the tipping discontinuity. The point estimates for minorities also suggest an income reduction, but they are less precisely estimated and not statistically significant.

Overall, these results provide evidence that tipping decreases the average income of a neighborhood. Importantly, the reduction in income is not only due to a change in racial composition but it can separately be observed among white residents, and to a weaker degree among minorities. Moreover, it is worth noting that this finding of a decline in income levels at the tipping discontinuity need not be at odds with the smaller discontinuities in white population for wealthier neighborhoods (see table 8). Whereas wealthier whites may be more tolerant of minority neighbors than less affluent whites, the wealthy are also more likely to be homeowners and may therefore have a greater financial incentive to leave a neighborhood.

4.8.2 Education

To add to the evidence for a discontinuous reduction in neighborhood quality at the tipping point, Table 11 provides results for the change in the share of neighborhood residents age 25 or higher who have a college degree.[27] The columns (1), (3), and (5) in the first panel of the table indicate a significant tipping discontinuity in the share of college graduates in both 1980-1990 and 1990-2000 but not in 1970-1980. In each of the three decades, neighborhoods with smaller initial renter shares and more homeowners experience a larger reduction in the share of residents with college degrees; in neighborhoods without renters, the predicted reduction in the share of college graduates is about one percentage point.

The similarity of education levels for the white and minority populations in neighborhoods near the tipping point that was shown in table 1 implies that a discontinuity in education would not result if the tipping effect were limited to the departure of a random sample of white residents. Indeed, the second panel of table 11 shows a discontinuous reduction in the share of college graduates within racial groups that is particularly pronounced for whites and again weaker for minorities.

Taken together, the results for income and education levels indicate a decline in neighborhood quality along both these dimensions when neighborhoods tip. The reduction in income and education is particularly apparent among whites and it is thus consistent with the notion that tipping neighborhoods primarily lose white homeowners, a group of residents that tends to be wealthier and better educated that renters.

[27] Benabou (1993) shows that the local concentration of residents with high educational attainment can strongly affect the residents of a neighborhood when there are local complementarities in human capital investments.

Table 11. Change in Educational Levels. Dependent Variable: Change Share of College Graduates among Residents Age 25+

I. All Races, 1970-2000

	A. 1970-1980		B. 1980-1990		C. 1990-2000	
	(1)	(2)	(3)	(4)	(5)	(6)
Beyond Tipping Point	0.002	-0.006	-0.006 **	-0.015 ***	-0.006 **	-0.012 ***
	(0.003)	(0.004)	(0.003)	(0.004)	(0.002)	(0.004)
Beyond TP x Renter Share		0.026 **		0.027 ***		0.018 *
		(0.010)		(0.010)		(0.009)
n	11,886	11,886	13,066	13,066	13,371	13,371
R^2	0.14	0.14	0.15	0.15	0.14	0.14

II. By Race, 1990-2000

	A. Whites		B. Minorities	
Beyond Tipping Point	-0.009 ***	-0.014 ***	0.000	-0.004
	(0.003)	(0.004)	(0.007)	(0.009)
Beyond TP x Renter Share		0.016		0.014
		(0.012)		(0.016)
n	13,246	13,246	13,274	13,274
R^2	0.04	0.04	0.03	0.03

All models use the baseline specification with demographic and housing controls as described in the table notes of table 4. Robust standard errors in parentheses are clustered by MSA. * p ≤ 0.10, ** p ≤ 0.05, *** p ≤ 0.01.

4.9 Conclusions

White preferences for neighborhoods with small minority shares can give rise to a tipping dynamic of racial segregation by which whites begin leaving a neighborhood once its minority share exceeds a critical tipping point. This paper proposes that if a neighborhood tips and white demand for housing falls in response to a growing minority population, homeowners may fear devaluation of their houses. Even though white owners and white renters both want to avoid an increase in the neighborhood minority share, homeowners have the additional financial incentive to leave a neighborhood ahead of immanent tipping to avoid a decrease in wealth. To illustrate that this market mechanism can exacerbate the tipping process, this paper develops a tipping model with asymmetric information and flexible prices which predicts that discontinuous population changes in tipping neighborhoods will be larger for neighborhoods with higher homeownership rates.

Financial incentives in the tipping process can also have important implications for the sequence of white residents' departure from segregating neighborhoods. In the classic model by Schelling (1971), a tipping neighborhood first loses the white residents with lowest racial tolerance, the most prejudiced of whom, survey evidence suggests, are typically persons of low education and income levels. Nonetheless, homeowners' financial incentives may contribute to a reversal of this sequence of white resident departure; that is, homeowners who on average are relatively wealthy and highly educated may leave the neighborhood ahead of poorer and less educated renters due not to greater racial prejudice but to avoid a loss in asset value.

This prediction of stronger tipping effects in neighborhoods with high homeownership rates is tested empirically using tipping values estimated by Card et al. (2008a). The results confirm that in every decade from 1970 to 2000, once the minority share exceeds the tipping point, neighborhoods with a high homeownership share experience considerably larger drops in white neighborhood population than those with more renters. Neighborhoods with a large homeownership rate also experience larger decreases in home values at the tipping discontinuity.

The effects of neighborhood tipping go beyond a change in racial composition: when a neighborhood tips, both income and education levels fall in the overall population, among white residents, and possibly among minority residents. Moreover, despite no evidence that higher incomes in a neighborhood exacerbate tipping, these results are consistent with the model's prediction that tipping neighborhoods will lose relatively wealthy and well-educated homeowners, who are more likely to leave than renters due not to greater racial prejudice but because of owners' financial incentive to avoid falling house prices. Indeed, tipping neighborhoods experience a sharp drop in white owner-occupied housing only, with no change observable in white renter-occupied housing.

This analysis thus provides evidence that tipping, usually considered a nonmarket interaction, may be augmented by market forces. The relative magnitude of these market and nonmarket channels and the complementarities between them warrants further study.

Appendix Table

Appendix Table 1. Decadal Change in White Population - Alternative Specification of Minority Share Polynomial. Dependent Variable: Change of White Population in Percentage Points of Initial Total Tract Population

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					A. 1970-1980					
Beyond Tipping Point	-17.8 ***	-22.4 ***	-10.1 ***	-15.9 ***	-9.9 ***	-17.5 ***	-11.8 ***	-18.8 ***	-19.5 ***	-25.8 *
	(2.5)	(4.8)	(3.3)	(5.1)	(3.5)	(5.3)	(3.5)	(5.2)	(4.4)	(6.0)
Beyond TP x Renter Share		13.1		17.7 *		23.6 **		21.7 **		20.7 *
		(9.3)		(9.3)		(9.6)		(9.7)		(10.0)
					B. 1980-1990					
Beyond Tipping Point	-15.0 ***	-22.5 ***	-11.2 ***	-18.5 ***	-11.3 ***	-19.9 ***	-12.2 ***	-21.1 ***	-16.2 ***	-25.3 *
	(2.0)	(4.1)	(3.1)	(5.2)	(3.4)	(5.5)	(3.9)	(6.0)	(5.1)	(6.8)
Beyond TP x Renter Share		22.2 ***		22.4 ***		26.5 ***		27.0 ***		29.0 *
		(7.8)		(7.7)		(7.7)		(7.8)		(7.7)
					C. 1990-2000					
Beyond Tipping Point	-12.2 ***	-23.0 ***	-9.9 ***	-20.0 ***	-8.5 ***	-19.4 ***	-8.1 ***	-19.3 ***	-7.8 ***	-19.0 *
	(1.3)	(2.3)	(1.7)	(2.6)	(1.8)	(2.6)	(1.8)	(2.6)	(2.7)	(3.1)
Beyond TP x Renter Share		33.5 ***		32.8 ***		34.5 ***		35.0 ***		36.2 *
		(4.6)		(4.7)		(4.6)		(4.5)		(4.5)
Minority Share Polynomial (Difference from TP)										
none	x	x								
2nd order			x	x						
4th order					x	x			x	x
6th order							x	x		
separate polyn on both sides of TP									x	x

All models use the baseline specification with demographic and housing controls as described in the table notes of table 4 but vary the order of the polynomial for the difference between a tract's minority share and the estimated MSA tipping point. The models of columns (9) and (10) include separate 4th order polynomials on both sides of the tipping point. * p ≤ 0.10, ** p ≤ 0.05, *** p ≤ 0.01.

Data Appendix

The empirical analysis is based on tract-level tabulations of decennial census data mapped onto Census 2000 tract boundaries in the Neighborhood Change Database (NCDB). The sample includes tracts from every Metropolitan Statistical Area (MSA, based on 1999 definition) containing at least 100 tracts with two exclusions: tracts whose decadal population growth rate exceeds the MSA mean by more than five standard deviations and tracts whose decadal growth in white population exceedes 500 percent of the baseline population.

The analysis of tipping discontinuities uses the estimated structural break tipping points identified by Card et al. (2008a). I thank David Card, Alexandre Mas, and Jesse Rothstein for generously sharing these data. All regressions include only the randomly selected 1/3 of the tract sample not used for the estimation of tipping points.

Throughout the analysis, the term “white population” refers to white non-Hispanics, while “minority population” comprises all blacks, Hispanics, native Americans, Asians, Pacific Islanders, and members of other nonwhite races. However, because census tabulations do not report separate population counts for white and nonwhite Hispanics in 1970, in accordance with Card et al. (2008a), this study draws on the 1980 data to estimate a regression of the white non-Hispanic population share in a tract on the white, black, and Hispanic shares, and then employs the coefficient estimates from this regression to predict the white non-Hispanic share of tract populations in 1970. To identify population changes between 1970 and 1980, it performs an analogous imputation for the non-Hispanic white share in 1980.

The best available data on homeowner status, income, and education by race and ethnicity are those for 1990 and 2000. In earlier years, these variables were either not reported or the Census Bureau suppressed values for a large number of tracts. Nonetheless, although homeowner status is reported separately for non-Hispanic whites in 1990 and 2000, income and education data is only tabulated for either all whites or all Hispanics (not for non-Hispanic whites). Hence, the number of white Hispanics in a given income or education cell (e.g., white Hispanic college graduates) is imputed by multiplying the reported share of a tract's white Hispanics by the number of Hispanics that fall into the income or education cell (e.g., share of whites among Hispanics x Hispanics that are college graduates). The imputed number of Hispanic whites in a cell is then subtracted from the number of all whites in that cell to obtain the number of non-Hispanic whites with a given income or education level.[28] The reported results for income and education of non-Hispanic whites are very similar to the results for all whites.

All regressions control for the renter share, which is defined as the share of renter-occupied housing units among all occupied housing units. An additional vector of tract population and housing characteristics includes the natural logarithm of mean family income, the share of persons aged 16+ who are in the civilian labor force and are unemployed, the share of families with children under age 18, the fraction of homes that are vacant, and the fraction of homes that are single unit.

Data on family income by race is only reported in broad intervals. Rather than converting these interval data to a continuous measure, this analysis exploits the fact that the real value of \$35,000 in 1989 almost coincides with the real value of \$45,000 in 1999 (income data in the census is based on the year that precedes the census). Because both values are available as interval borders in the respective years, the data thus allow consistent measurement of the share of white or minority families with an income above \$45,000

[28]In the 2000 census, individuals were allowed to indicate multiple races. Whenever tabulations for every possible combination of races are available, the NCDB counts persons who choose combinations of white, Native American, and "other race" as white and all other multiple-race persons as minorities. The income and education data, however, only report a summary count of all multiple-race persons for each income or education cell. To compute the share of whites among multiple race persons for each MSA, this analysis uses microdata from the 5 percent extract of the 2000 census (Ruggles et al., 2004). Income and education counts for multiple-race persons are then assigned to whites and minorities according to these proportions.

by 1999 value. All real dollar amounts are inflated using the Personal Consumption Expenditure price index.

Theory Appendix

Ad Section 3.2. Expectations of Outsiders. According to equation (4.3.7), outsiders expect a probability of $\gamma'\pi'$ that their new neighborhood has a minority share above $\bar{m}$ where γ' is the expected probability that incumbents of the neighborhood observed signal s_H and π' is the expected probability that the minority share increases beyond $\bar{m}$ given signal s_H. In the formation of these expectations, outsiders can take into account that neighborhood incumbents may choose to move upon observing s_H. Consider the standard case where only white home owners leave. A share r of all houses that change ownership are bought by investors and used as rental houses, and the departure of white owners will thus both increase the number of new owner-occupiers and new renters that move into the neighborhood. Suppose the share of new residents who move in given signal s_H is $\lambda' > \lambda$. The expected probability that a vacancy arises in a neighborhood whose incumbents obtained the signal s_H is then

$$\gamma' \equiv \frac{\gamma\lambda'}{(1-\gamma)\lambda + \gamma\lambda'} > \gamma \tag{A.1}$$

where γ is the probability that incumbents observe the signal s_H. Whether or not the minority share in a neighborhood with signal s_H will exceed $\bar{m}$ depends on the exact composition of the neighborhood which is not observed by outside agents. Let π' be outsiders' expectation of the probability that the neighborhood minority share will exceed $\bar{m}$ in case of the signal s_H. The expected probability of moving into a neighborhood whose minority share rises above $\bar{m}$ is then equal to

$$E[Prob(m_1 > \bar{m})] \mid no\ signal = \gamma'\pi' \tag{A.2}$$

Without making further assumptions on outsiders' expectations, it is clear that for a small likelihood of the signal s_H, $\gamma'\pi' \to 0$ as $\gamma \to 0$. In the rare case that incumbent residents do observe the signal s_H, they will anticipate a much larger probability of a minority share above $\bar{m}$ than outsiders who consider such a change very unlikely.

Proof of Proposition 1. If $m_0 = m_0^* \equiv \frac{1}{1+\lambda}\bar{m}$ and λ vacancies are filled with new residents of whom a share $2m_0$ are minorities, then $m_1 = (1+\lambda)m_0 = \bar{m}$. Without taking into account moving decisions of other neighborhood incumbents, residents will expect that $m_1 > \bar{m}$ with probability 0 if $m_0 \leq m_0^*$ and no agent has an incentive to move voluntarily. If however $m_0 > m_0^*$, agents expect that $m_1 > \bar{m}$ with probability π. Whites' consumption utility falls by $2v_H$, from v_H to $-v_H$, if $m_1 > \bar{m}$ instead of $m_1 \leq \bar{m}$. Incumbents do not obtain signals for alternative locations and therefore expect that $m_1 > \bar{m}$ occurs with probability $\gamma'\pi'$ when locating outside the neighborhood. White renters will move if the sum of expected utility outside the neighborhood (net of housing costs) and moving costs exceeds the sum of expected utility inside the neighborhood and rental payment,

$$-\gamma'\pi' 2v_H - c = (1-\pi)v_H - \pi v_H - v_H \tag{A.3}$$

which solves to (4.3.12b). Minority owners will move if the payoff from selling at the equilibrium price P, moving at a cost c and obtaining a utility of zero outside the neighborhood exceeds the expected payoff from staying inside the neighborhood,

$$P - c = v_H + \beta[(1-\gamma'\pi')P_H + \gamma'\pi' P_L] - c > v_H + \beta[(1-\pi)P_H + \pi P_L] \tag{A.4}$$

which solves to (4.3.12c). The moving condition for white owners results from a combination of the arguments for minority owners and white renters. ■

Proof of Proposition 2. If $m_0 \leq m_0^*$, no group of current residents has an incentive to deviate from the strategy of staying because $m_1 \leq \bar{m}$ even when a minority shock occurs. If $m_0 > m_0^*$ and equation (4.3.12a) is fulfilled, moving is a strictly dominating strategy for white owners and all neighborhood incumbents will anticipate their departure. If no minority shock takes place, the voluntary departure of the $(1-\lambda)(1-m_0)(1-r_w)$ white owners who were not forced to leave raises the minority share to

$$m_1 = m_0 + (1-\lambda)(1-m_0)(1-r_w)m_0 \tag{A.5}$$

The resulting minority share is equal to $\bar{m}$ if the initial home owner share among whites is

$$1 - r_w = 1 - r_w^* \equiv \frac{\bar{m} - m_0}{(1-\lambda)(1-m_0)m_0} \tag{A.6}$$

Thus, if the neighborhood has a white owner share $(1-r_w) \leq (1-r_w^*)$, the departure of white owners will not increase m_1 above $\bar{m}$ and therefore $Prob(m_1 > \bar{m}) = \pi$. Neither minority owners nor white renters will deviate from the strategy of staying when equations (4.3.12b) and (4.3.12c) do not hold.

If the neighborhood has a white owner share $(1-r_w) > (1-r_w^*)$, the minority share will always rise beyond $\bar{m}$ if only white owners depart, even when no shock occurs. White renters will therefore also chose leave in order to avoid a disutility from the increased minority share, thus exacerbating the decrease in white population and increase in m_1.

Since P^{end} is lower when $m_1 > \bar{m}$, minority home owners will also depart. Their departure partly mitigates the increase in m_1 that is caused by the departure of white residents; however, the condition

$$1 - r_w^* < r_m \tag{A.7}$$

implies that the departure of minority homeowners cannot push m_1 back below $\bar{m}$. It can then readily be verified that $m_1 > \bar{m}$ when only minority renters stay and the neighborhood fills with new residents of whom a share m_0 are minorities. If condition (A.7) were not fulfilled, then $m_1 \leq \bar{m}$ when all minority owners would leave along with all white residents. Minority owners would then prefer a mixed strategy where their moving decision is randomized so that $\pi < Prob(m_1 > \bar{m}) < 1$ when $1 - r_w > 1 - r_w^*$. Even in that case, however, the probability of exceeding $\bar{m}$ is weakly increasing in $1 - r_w$. ■

Ad Section 3.5. Alternative Expectations of Residents. Suppose agents expect that on observing s_H, all residents leave unless leaving is a dominated strategy. For minority renters, the strategy of leaving is always strictly dominated by staying because the payoff of minority renters does not fall when $m_1 > \bar{m}$. All agents will therefore expect that minority renters stay. However, if no minority shock takes place and the renter share satisfies condition (A.7), then the voluntary departure of all residents but minority renters always raises the minority share to

$$m_1 > \bar{m} \tag{A.8}$$

The combination of minority renters who stay in the neighborhood and a minority share m_0 among new residents adds to a minority share $m_1 > \bar{m}$. Based on the expectation that other residents depart as well, it is therefore rational for all neighborhood incumbents but minority renters to leave. As a consequence, it is possible that agents choose to leave the neighborhood even in case of an initial minority share $m_0 \leq m^*$ where an increase in the minority share above $\bar{m}$ could have been avoided if all residents stayed. Importantly, this alternative expectation structure does not predict a differential moving behavior for white owners and white renters; a prediction that is rejected by the empirical analysis. ■

Thesis Appendix: Data Contributions

A.1 Consistent Panel of Occupations

A.1.1 Census Occupations

The United States Census records the detailed titles of workers' occupations. The publicly available Census data aggregates this occupation information and reports several hundred 3-digit occupation codes. Most of these occupation codes refer to narrow groups of occupations, such as "chemical engineers" or "petroleum engineers" while residual categories such as "engineers, n.e.c. (not elsewhere classified)" contain workers who could not be matched to a more detailed code. The occupational classification system gets redefined for every decennial Census. Changes to the occupation system often reflect the growth and decline of specific occupations. For instance, "loom fixers" formed a detailed occupation code until 1970 but they were later integrated into the broader group of "machinery maintenance occupations" as loom fixers' importance in the labor market declined. By contrast, a detailed occupation code for "speech therapists" was introduced only in 1980. Before, these workers were classified among "therapists and healers, n.e.c.".

Appendix Table 1 lists the number of civilian occupations in each Census occupation system since 1950. The Census Bureau introduced ever more detailed occupation classifications between 1950 and 1980, thus almost doubling the number of reported occupation codes from 268 to 504. Each of the subsequent data sets provides at least 465 occupation codes. The two largest overhauls of the occupation system took place for the 1980 and 2000 Censuses. In these years, the Census Bureau not only added new occupations but also ceased using many detailed occupation categories that had been previously reported.

The occupational classification schemes also sort detailed occupations into broader occupation groups. One of these broader groups of occupations is called "service occupations". These jobs typically involve caring for, serving, or protecting others. Appendix Table 1 shows the number of service occupations in each Census occupation system since 1950. The number of service occupation codes increased over time along with the overall growth in occupation codes. Furthermore, the 2000 Census slightly expanded the set of

Appendix Table 1. Number of Occupations in the 1950-2000 Censuses and 2005 ACS

	Census 1950	Census 1960	Census 1970	Census 1980	Census 1990	Census 2000 (5% Sample)	ACS 2005
No. of occupations	268	295	440	504	502	471	465
No. of service occs	28	32	44	44	46	59	57

Occupation counts include all civilian occupations. The 5% sample of the 2000 Census reports 34 fewer occupation codes than the 1% sample.

occupations that are considered to be "service occupations" by including additional jobs such as "gardeners and groundskeepers" or "movie picture projectionists".

In order to track detailed occupations over time, empirical work has to rely on crosswalks that match occupation codes from different Census years. Meyer and Osborne (2005) provide a carefully constructed crosswalk that matches occupations from the 1960-2000 Censuses to a system of 386 '*occ1990*' occupation codes which are roughly based on the 1990 Census occupation system. Due to the large similarity between occupations in the 1950 and 1960 Censuses, the *occ1990* system can readily be extended back to 1950. Furthermore, it can also be used for the 2005 American Community Survey which uses a slightly aggregated version of the 2000 Census occupation system.

One limitation of the *occ1990* system is that its occupation panel is unbalanced. For instance, there are no observations for the *occ1990* category "economics instructors" in the 2000 Census. The specific fields of college teachers are no longer reported in the 2000 Census and economics instructors therefore have to be included in the broader *occ1990* category "subject instructors, college". This unbalanced panel structure would be problematic for the empirical work in this thesis that studies employment or wage changes within detailed occupations between 1980 and 2005.

A.1.2 *Occ1990dd* Occupation System

We therefore developed a new occupation system with 330 '*occ1990dd*' codes that provides a balanced panel of occupations covering the 1980, 1990, and 2000 Censuses and the 2005 ACS. The balanced panel allows for the analysis of changes at the level of detailed occupations between 1980 and 2005.[29] Most *occ1990dd* codes result from a simple aggregation of *occ1990* occupations, Moreover, the *occ1990dd* system makes a particular effort to provide consistent definitions of the detailed service occupations which are studied in the first essay of this thesis. Appendix Table 2 lists all *occ1990dd* occupations and details the construction of occupation codes whose definitions differ from

[29] We also experimented with an extension of the balanced panel back to the 1950 Census. Due to the smaller number of occupations reported in the earlier Censuses and considerable changes in the occupation system between 1970 and 1980, an extended balanced panel would be much less detailed.

the *occ1990* codes by Meyer and Osborne (2005).

The *occ1990dd* codes can be aggregated into broader occupation groups. Broader groups of occupations can be tracked quite well in every Census since 1950 even if some detailed occupation codes are not reported for every year. For simplicity, we distinguish the six major occupation groups that are reported in the 1990 Census on which the *occ1990* and *occ1990dd* occupation systems are based. The first of these groups is "managerial and professional specialty occupations" which covers most of the best paid occupations in the labor market. "Technical, sales, and administrative support occupations" cover a workforce that is on average better educated than any other occupation group apart from managers and professionals. The remaining four major occupation groups which all have relatively low education levels are "service occupations", "farming, forestry, and fishing occupations", "precision production, craft, and repair occupations", and "operators, fabricators, and laborers".

The designation of "service occupations" closely follows the updated and extended definition that is used in the 2000 Census. The effort to identify a consistent set of service occupations over time should mitigate the concern that the recent growth of employment and wages in service occupations may be influenced by a changing definition of these occupations. Moreover, any spurious changes that may nevertheless be caused by the 1980-2005 occupation crosswalk should be confined to the 1990-2000 period because the only noteworthy changes to the Census occupation system since 1980 took place between the 1990 and 2000 Censuses. The growth of service employment and wages does, however, also prevail in the 1980-1990 and 2000-2005 periods which are nearly unaffected by changes to occupational classification.

A.1.3 Matching Task Data to Occupations

We use job task data from the Dictionary of Occupational Titles (DOT; U. S. Department of Labor, Employment and Training Administration, 1977) to characterize the task content of occupations. The DOT assesses the occupational tasks of more than 12,000 highly detailed occupations. Our DOT data is based on an aggregation of these detailed occupations to the three-digit occupation codes of the 1970 Census.[30] The DOT task scores are available for all but a few occupation codes. Occupations with missing task data are mostly residual categories like "miscellaneous operatives" or "clerical workers, not specified" that may cover workers whose Census forms provided imprecise occupational information. We impute task scores for these occupations by computing, e.g., the task values for "professionals, not elsewhere classified" as the average task score of all professional occupations weighted by each occupation's total work hours in 1970.

The DOT 1977 task values are matched to *occ1990dd* occupation codes in several steps.

[30] See Autor et al. (2003) for a more detailed description of this data.

Appendix Table 2. "Occ1990dd" Occupation System, 1950-2005

Occ1990dd Code	Occupation Groups and Occupation Titles	Census 1950 Codes	Census 1960 Codes	Census 1970 Codes	Balanced Panel 1980-2005: Census 1980 Codes	Census 1990 Codes	Census 2000 (5% Sample) Codes	ACS 2005 Codes
	A. Managerial and Professional Specialty Occupations							
	A.1 Executive, Administrative, and Managerial Occupations							
4	Chief executives, public administrators, and legislators	250	270		3, 4	3, 4	1, 3	1
7	Financial managers			x	x	x	x	x
8	Human resources and labor relations managers				x	x	x	x
13	Managers and specialists in marketing, advert., PR		x	x	x	x	x	x
14	Managers in education and related fields	x	x	x	x	x	x	x
15	Managers of medicine and health occupations			x	x	x	x	x
18	Managers of properties and real estate	x	x	x	x	x	x	x
19	Funeral directors	x	x	x	x	x	x	x
22	Managers and administrators, n.e.c.	260, 290, 357	275, 280, 290	195, 201, 220, 222-224, 245, 246	5, 17, 19	5, 16, 17, 21, 22	2, 10, 11, 14, 22, 30, 31, 33, 34, 36, 40, 42, 43, 60, 72	2, 10, 11, 14, 22, 30, 31, 33, 34, 36, 42, 43, 60, 72
	A.2 Management Related Occupations							
23	Accountants and auditors	x	x	x	x	x	x	x
24	Insurance underwriters				x	x	x	x
25	Other financial specialists	x	x		x	x	x	x
26	Management analysts				x	x	x	x
27	Personnel, HR, training, and labor rel. specialists				x	x	x	x
28	Purchasing agents and buyers of farm products	x	x	x	x	x	x	x
29	Buyers, wholesale and retail trade	x	x	x	x	x	x	x
33	Purchasing managers, agents, and buyers, n.e.c.	x	x	x	x	x	x	x
34	Business and promotion agents				x	x	x	x
35	Construction inspectors			x	x	x	x	x
36	Inspectors and compliance officers, outside	x	x	x	x	x	x	x
37	Management support occupations				x	x	x	x
	A.3 Professional Specialty Occupations							
43	Architects	x	x	x	x	x	x	x
44	Aerospace engineers	x	x	x	x	x	x	x
45	Metallurgical and materials engineers	x	x	x	x	x	x	x
47	Petroleum, mining, and geological engineers	48	91	20	46, 47	46, 47	150, 152	152
48	Chemical engineers	x	x	x	x	x	x	x
53	Civil engineers	x	x	x	x	x	x	x
55	Electrical engineers	x	x	x	x	x	x	x
56	Industrial engineers	x	x	x	x	x	x	x
57	Mechanical engineers	x	x	x	x	x	x	x
59	Engineers and other professionals, n.e.c.	49, 99	93, 195	23, 196	49, 54, 58, 59	49, 54, 58, 59	142, 144, 150, 153	134, 142, 144, 150, 153

Occ1990dd Code	Occupation Groups and Occupation Titles	Census 1950 Codes	Census 1960 Codes	Census 1970 Codes	Census 1980 Codes	Census 1990 Codes	Census 2000 (5% Sample) Codes	ACS 2005 Codes
64	Computer systems analysts and computer scientists			x	x	x	x	x
65	Operations and systems researchers and analysts			x	x	x	x	x
66	Actuaries			x	x	x	x	x
68	Mathematicians and statisticians	67, 83	135, 174	35, 36	67, 68	67, 68	124	124
69	Physicists and astronomists	x	x	x	x	x	x	x
73	Chemists	x	x	x	x	x	x	x
74	Atmospheric and space scientists			x	x	x	x	x
75	Geologists	x	x	x	x	x	x	x
76	Physical scientists, n.e.c.	x	x	x	x	x	x	x
77	Agricultural and food scientists	x	x	x	x	x	x	x
78	Biological scientists	x	x	x	x	x	x	x
79	Foresters and conservation scientists	x	x	x	x	x	x	x
83	Medical scientists				x	x	x	x
84	Physicians	x	x	x	x	x	x	x
85	Dentists	x	x	x	x	x	x	x
86	Veterinarians	x	x	x	x	x	x	x
87	Optometrists	x	x	x	x	x	x	x
88	Podiatrists			x	x	x	x	x
89	Other health and therapy occupations	x	x	x	x	x	x	x
95	Registered nurses	58, 59	150, 151	75, 923	95	95	313	313
96	Pharmacists	x	x	x	x	x	x	x
97	Dieticians and nutritionists	x	x	x	x	x	x	x
98	Respiratory therapists				x	x	x	x
99	Occupational therapists				99	99	315	315
103	Physical therapists				103	103	316	316
104	Speech therapists				x	x	x	x
105	Therapists, n.e.c.	x	x	x	x	x	x	x
106	Physicians' assistants				x	x	x	x
154	Subject instructors, college	12-29	31-60	102-140	113-153	113-153	220	220
155	Kindergarten and earlier school teachers			x	x	x	x	x
156	Primary school teachers	x	x	x	x	x	x	x
157	Secondary school teachers		x	x	x	x	x	x
158	Special education teachers				x	x	x	x
159	Teachers, n.e.c.		184	141, 145	159	159	234, 255	234, 255
163	Vocational and educational counselors			x	x	x	x	x
164	Librarians	x	x	x	x	x	x	x
165	Archivists and curators			x	x	x	x	x
166	Economists, market and survey researchers	x	x	x	x	x	x	x
167	Psychologists	x	x	x	x	x	x	x
169	Social scientists and sociologists, n.e.c.	52, 84	102, 175	24, 26, 92, 94	168, 169	168, 169	186	186
173	Urban and regional planners			x	x	x	x	x
174	Social workers	x	x	x	x	x	x	x
176	Clergy and religious workers	x	x	x	x	x	x	x
177	Welfare service workers			954	467	465	202	202
178	Lawyers and judges	55	105	30, 31	178, 179	178, 179	210, 211	210
183	Writers and authors	x	x	x	x	x	x	x
184	Technical writers				x	x	x	x

Occ1990dd Code	Occupation Groups and Occupation Titles	Census 1950 Codes	Census 1960 Codes	Census 1970 Codes	Census 1980 Codes	Census 1990 Codes	Census 2000 (5% Sample) Codes	ACS 2005 Codes
185	Designers	x	x	x	x	x	x	x
186	Musicians and composers	x	x	x	x	x	x	x
187	Actors, directors, and producers	x	x	x	x	x	x	x
188	Painters, sculptors, craft-artists, and print-makers	x	x	x	x	x	x	x
189	Photographers	x	x	x	x	x	x	x
193	Dancers	x	x	x	x	x	x	x
194	Art/entertainment performers and related occs	x	x	x	x	x	x	x
195	Editors and reporters	x	x	x	x	x	x	x
198	Announcers			x	x	x	x	x
199	Athletes, sports instructors, and officials	x	x	x	x	x	x	x
	B. Technical, Sales, and Administrative Support Occupations							
	B.1 Technicians and Related Support Occupations							
203	Clinical laboratory technologies and technicians	x	x	x	x	x	x	x
204	Dental hygienists			x	x	x	x	x
205	Health record technologists and technicians			x	x	x	x	x
206	Radiologic technologists and technicians			x	x	x	x	x
207	Licensed practical nurses	x	x	x	x	x	x	x
208	Health technologists and technicians, n.e.c.			x	x	x	x	x
214	Engineering technicians		190	154, 155, 162	213-216	213-216	155	155
217	Drafters	x	x	x	x	x	x	x
218	Surveryors, cartographers, mapping scientists/techs	x	x	x	x	x	x	x
223	Biological technicians			x	x	x	x	x
224	Chemical technicians			x	x	x	x	x
225	Other science technicians	x	x	x	x	x	x	x
226	Airplane pilots and navigators	x	x	x	x	x	x	x
227	Air traffic controllers			x	x	x	x	x
228	Broadcast equipment operators	76	164	171	228	228	290	290
229	Computer software developers			x	x	x	x	x
233	Programmers of numerically controlled machine tools			x	x	x	x	x
234	Legal assistants and paralegals				x	x	x	x
235	Technicians, n.e.c.	96	192	165, 173	235	235	196	196
	B.2 Sales Occupations							
243	Sales supervisors and proprietors	x	x		x	x	x	x
253	Insurance sales occupations		x	x	x	x	x	x
254	Real estate sales occupations	x	x	x	x	x	x	x
255	Financial service sales occupations	x	x	x	x	x	x	x
256	Advertising and related sales jobs	x	x	x	x	x	x	x
258	Sales engineers		x	x	x	x	x	x
274	Salespersons, n.e.c.	300, 410, 430, 490	301, 381, 383, 394	261, 280-282, 285, 296	257, 259, 284, 285	257, 259, 284, 285	485, 494, 496	485, 494, 496
275	Retail salespersons and sales clerks			283, 284, 314	263-269, 274, 275	263-269, 274, 275	474-476, 484	474-476, 484
276	Cashiers	x	x	x	x	x	x	x
277	Door-to-door sales, street sales, and news vendors	x	x	x	x	x	x	x
283	Sales demonstrators, promoters, and models	223, 514	382	262	283	283	490	490

Occ1990dd Code	Occupation Groups and Occupation Titles	Census 1950 Codes	Census 1960 Codes	Census 1970 Codes	Census 1980 Codes	Census 1990 Codes	Census 2000 (5% Sample) Codes	ACS 2005 Codes
	B.3 Administrative Support Occupations							
303	Office supervisors			x	x	x	x	x
308	Computer and peripheral equipment operators			x	x	x	x	x
313	Secretaries and stenographers	350	342, 345	370-372, 376	313, 314	313, 314	570	570
315	Typists		360	391	315	315	582	582
316	Interviewers, enumerators, and surveyors			x	x	x	x	x
317	Hotel clerks				x	x	x	x
318	Transportation ticket and reservation agents	x	x	x	x	x	x	x
319	Receptionists and other information clerks		341	364	319, 323	319, 323	540	540
326	Correspondence and order clerks				x	x	x	x
328	Human resources clerks, excl payroll and timekeeping	x	x	x	x	x	x	x
329	Library assistants	x	x	x	x	x	x	x
335	File clerks		x	x	x	x	x	x
336	Records clerks				325, 336	325, 336	520, 542	520, 542
337	Bookkeepers and accounting and auditing clerks	x	x	x	x	x	x	x
338	Payroll and timekeeping clerks		x	x	x	x	x	x
344	Billing clerks and related financial records processing			303, 341, 342	339, 343, 344	339, 343, 344	511	511
346	Mail and paper handlers			x	x	x	x	x
347	Office machine operators, n.e.c.	341	325	344, 355	345, 347	345, 347	590	590
348	Telephone operators	x	x	x	x	x	x	x
349	Other telecom operators	x	x	x	x	x	x	x
354	Postal clerks, exluding mail carriers		x	x	x	x	x	x
355	Mail carriers for postal service	x	x	x	x	x	x	x
356	Mail clerks, outside of post office	x	x		x	x	x	x
357	Messengers	x	x	x	x	x	x	x
359	Dispatchers	x	x	x	x	x	x	x
361	Inspectors, n.e.c.	x	x					
364	Shipping and receiving clerks	x	x	x	x	x	x	x
365	Stock and inventory clerks		x	x	x	x	x	x
366	Meter readers			x	x	x	x	x
368	Weighers, measurers, and checkers			x	x	x	x	x
373	Material recording, sched., prod., plan., expediting cl.			x	x	x	x	x
375	Insurance adjusters, examiners, and investigators	x	x	x	x	x	x	x
376	Customer service reps, invest., adjusters, excl. insur.			x	x	x	x	x
377	Eligibility clerks for government prog., social welfare				x	x	x	x
378	Bill and account collectors	x	x	x	x	x	x	x
379	General office clerks				379	379	586	586
383	Bank tellers	x	x	x	x	x	x	x
384	Proofreaders			x	x	x	x	x
385	Data entry keyers			x	x	x	x	x
386	Statistical clerks			x	x	x	x	x
387	Teacher's aides			382	387	387, 467	254	254
389	Administrative support jobs, n.e.c.	390, 533	370, 450	311, 394, 395, 396	369, 374, 389	369, 374, 389	522, 583, 593	522, 593

Occ1990dd Code	Occupation Groups and Occupation Titles	Census 1950 Codes	Census 1960 Codes	Census 1970 Codes	Census 1980 Codes	Census 1990 Codes	Census 2000 (5% Sample) Codes	ACS 2005 Codes
	C. Service Occupations							
	C.1 Housekeeping and Cleaning Occupations							
405	Housekeepers, maids, butlers, and cleaners	700, 751-753, 764	802, 820, 821, 823, 824, 832	901, 902, 940, 941, 950, 982, 984	405, 407, 449	405, 407, 449	423	423
408	Laundry and dry cleaning workers	643, 710	674, 803	611, 630, 983	403, 747, 748	403, 747, 748	830	830
	C.2 Protective Service Occupations							
415	Supervisors of guards				x	x	x	x
417	Fire fighting, fire prevention, and fire inspection occs	x	x	x	x	x	x	x
418	Police and detectives, public service	773	853	964	6, 414, 418	6, 414, 418	371, 382, 384, 385	371, 382, 384, 385
423	Sheriffs, bailiffs, correctional institution officers	771, 782	852, 854	963, 965	423, 424	423, 424	370, 380	370, 380
425	Crossing guards	785	860	960	425	425	394	394
426	Guards and police, except public service	763	851	962	426	426	391, 392	391, 392
427	Protective service, n.e.c.				x	x	x	x
	C.3 Other Service Occupations							
	Food Preparation and Service Occupations							
433	Supervisors of food preparation and service				433	433	401	401
434	Bartenders	x	x	x	x	x	x	x
435	Waiters and waitresses	x	x	x	x	x	x	x
436	Cooks	754	825	912, 981	404, 436, 437	404, 436, 437	400, 402	400, 402
439	Food preparation workers				439	439	403	403
444	Miscellanious food preparation and service workers	760	830, 835	911, 913, 914, 916	438, 443	438, 443	405, 406, 412-415	405, 406, 412-415
	Healthcare Support Occupations							
445	Dental Assistants	302	303	921	445	445	364	364
447	Health and nursing aides	730	810	922, 925	446, 447	446, 447	360-363, 365, 461	360-363, 365, 461
	Building/Grounds Cleaning/Maintenance Occs							
448	Supervisors of cleaning and building service				x	x	x	x
450	Superv. of landscaping, lawn service, groundskeeping				485	485	421	421
451	Gardeners and groundskeepers	930	964	755	486	486	425	425
453	Janitors	x	x	x	x	x	x	x
455	Pest control occupations				x	x	x	x
	Personal Appearance Occupations							
457	Barbers		814	935	457	457	450	450
458	Hairdressers and cosmetologists	x	x	x	x	x	x	x
	Recreation and Hospitality Occupations							
459	Recreation facility attendants	732	813	932	459	459	430, 440, 443	430, 440, 443
461	Guides				x	x	x	x
462	Ushers	x	x	x	x	x	x	x
464	Baggage porters, bellhops and concierges	x	x	x	x	x	x	x
466	Recreation and fitness workers	77	165	101	175	175	462	462
467	Motion picture projectionists	562	493	505	773	773	441	441

Occ1990dd Code	Occupation Groups and Occupation Titles	Census 1950 Codes	Census 1960 Codes	Census 1970 Codes	Census 1980 Codes	Census 1990 Codes	Census 2000 (5% Sample) Codes	ACS 2005 Codes
	Child Care Workers							
468	Child care workers		x	x	x	x	x	x
	Misc. Personal Care and Service Occupations							
469	Personal service occupations, n.e.c	720, 731, 761, 790	804, 812, 831, 890	933, 943, 945, 976, 986	454, 469	454, 469	446, 465	446, 465
470	Supervisors of personal service jobs, n.e.c				456	456	432	432
471	Public transportation attendants and inspectors			931	465	463	455, 941	455, 941
472	Animal caretakers, except farm			740	487	487	434, 435	434, 435
	D. Farming, Forestry, and Fishing Occupations							
	D.1 Farm Operators and Managers							
473	Farmers (owners and tenants)	100	200	801	473, 474	473, 474	21	21
475	Farm managers	123	222	802, 806, 821	475, 476	475, 476	20	20
	D.2 Other Agricultural and Related Occupations							
479	Farm workers, incl. nursery farming	810-840	901-903, 905	822-824, 846	477, 479, 484	477, 479, 484	434, 605	434, 605
488	Graders and sorters of agricultural products	x	x		x	x	x	x
489	Inspectors of agricultural products				x	x	x	x
496	Timber, logging, and forestry workers	x	x	x	x	x	x	x
498	Fishers, marine life cultivators, hunters, and kindred	910	962	752	483, 498, 499	483, 498, 499	610	610
	E. Precision Production, Craft, and Repair Occupations							
	E.1 Mechanics and Repairers							
503	Supervisors of mechanics and repairers				x	x	x	x
505	Automobile mechanics and repairers	x	x	x	x	x	x	x
507	Bus, truck, and stationary engine mechanics				x	x	x	x
508	Aircraft mechanics	x	x	x	x	x	x	x
509	Small engine repairers				x	x	x	x
514	Auto body repairers			x	x	x	x	x
516	Heavy equipement and farm equipment mechanics			x	x	x	x	x
518	Industrial machinery repairers				x	x	x	x
519	Machinery maintenance occupations	x	x	x	x	x	x	x
523	Repairers of industrial electrical equipment	x	x	x	x	x	x	x
525	Repairers of data processing equipment			x	x	x	x	x
526	Repairers of household appliances and power tools			x	x	x	x	x
527	Telecom and line installers and repairers	x	x	x	x	x	x	x
533	Repairers of electrical equipment, n.e.c.	551	473	484	533, 538	533, 538	703, 711	703, 711
534	Heating, air conditioning, and refrigeration mechanics		x	x	x	x	x	x
535	Precision makers, repairers, and smiths	x	x	x	x	x	x	x
536	Locksmiths and safe repairers				x	x	x	x
539	Repairers of mechanical controls and valves				x	x	x	x
543	Elevator installers and repairers				x	x	x	x
544	Millwrights	x	x	x	x	x	x	x
549	Mechanics and repairers, n.e.c.	553, 554, 605, 614, 615	475, 480, 610, 620, 621	403, 486, 492, 495, 571, 572, 575, 586	547, 549, 864	547, 549	734, 751, 755, 756, 762	734, 751, 755, 756, 762

Occ1990dd Code	Occupation Groups and Occupation Titles	Census 1950 Codes	Census 1960 Codes	Census 1970 Codes	Census 1980 Codes	Census 1990 Codes	Census 2000 (5% Sample) Codes	ACS 2005 Codes
	E.2 Construction Trades							
558	Supervisors of construction work				x	x	x	x
563	Masons, tilers, and carpet installers	x	x	x	x	x	x	x
567	Carpenters	x	x	x	x	x	x	x
573	Drywall installers			x	x	x	x	x
575	Electricians	x	x	x	x	x	x	x
577	Electric power installers and repairers			433	577	577	704, 741	704, 741
579	Painters, construction and maintenance	x	x	x	x	x	x	x
583	Paperhangers	x	x	x	x	x	x	x
584	Plasterers	x	x	x	x	x	x	x
585	Plumbers, pipe fitters, and steamfitters	x	x	x	x	x	x	x
588	Concrete and cement workers	x	x	x	x	x	x	x
589	Glaziers	x	x	x	x	x	x	x
593	Insulation workers	x	x	x	x	x	x	x
594	Paving, surfacing, and tamping equipment operators			x	x	x	x	x
595	Roofers and slaters	x	x	x	x	x	x	x
597	Structural metal workers	x	x	x	x	x	x	x
598	Drillers of earth			x	x	x	x	x
599	Misc. construction and related occupations	611	613	440	596, 599	596, 599	671, 675, 676	671, 675, 676
	E.3 Extractive Occupations							
614	Drillers of oil wells				614	614	680	680
615	Explosives workers	x	x	x	x	x	x	x
616	Miners	x	x	x	x	x	x	x
617	Other mining occupations				617	617	694	694
	E.4 Precision Production Occupations							
628	Production supervisors or foremen	x	x	x	x	x	x	x
634	Tool and die makers and die setters	x	x	x	x	x	x	x
637	Machinists	x	x	x	x	x	x	x
643	Boilermakers	x	x	x	x	x	x	x
644	Precision grinders and fitters				x	x	x	x
645	Patternmakers and model makers	570	502	514	645, 656, 676	645, 656, 676	806	806
649	Engravers	x	x	x	x	x	x	x
653	Other metal and plastic workers	591, 612	525, 614	535, 536, 540	646, 653, 654	646, 653, 654	652, 816	652
657	Cabinetmakers and bench carpeters	x	x	x	x	x	x	x
658	Furniture/wood finishers, other prec. wood workers			443	658, 659	658, 659	851	851
666	Dressmakers, seamstresses, and tailors	590, 633	524, 651	551, 613	666, 667	666, 667	835	835
668	Upholsterers	x	x	x	x	x	x	x
669	Shoemakers, other prec. apparel and fabric workers	525, 645	432, 680, 705	444, 542, 636	669, 674	669, 674	833	833
675	Hand molders and shapers, except jewelers			x	x	x	x	x
677	Optical goods workers	x	x	x	x	x	x	x
678	Dental laboratory and medical applicance technicians			x	x	x	x	x
679	Bookbinders	x	x	x	x	x	x	x
684	Other precision and craft workers	594	545		684	684	822	822

Occ1990dd Code	Occupation Groups and Occupation Titles	Census 1950 Codes	Census 1960 Codes	Census 1970 Codes	Census 1980 Codes	Census 1990 Codes	Census 2000 (5% Sample) Codes	ACS 2005 Codes
686	Butchers and meat cutters	x	x	x	x	x	x	x
687	Bakers	x	x	x	x	x	x	x
688	Batch food makers				x	x	x	x
694	Water and sewage treatment plant operators				x	x	x	x
695	Power plant operators	x	x	x	x	x	x	x
696	Plant and system operators, stationary engineers	x	x	x	x	x	x	x
699	Other plant and system operators				x	x	x	x
	F. Operators, Fabricators, and Laborers							
	F.1 Machine Operators, Assemblers, and Inspectors							
703	Lathe, milling, and turning machine operatives	535	452	454, 652, 653	703-705	703-705	801	801
706	Punching and stamping press operatives			x	x	x	x	x
707	Rollers, roll hands, and finishers of metal	x	x	x	x	x	x	x
708	Drilling and boring machine operators			x	x	x	x	x
709	Grinding, abrading, buffing, and polishing workers	x	x	x	x	x	x	x
713	Forge and hammer operators	x	x	x	x	x	x	x
719	Molders and casting machine operators	x	x	x	x	x	x	x
723	Metal platers			x	x	x	x	x
724	Heat treating equipment operators	x	x	x	x	x	x	x
727	Sawing machine operators and sawyers	x	x	x	x	x	x	x
729	Nail, tacking, shaping and joining mach ops (wood)				728, 729	728, 729	854	854
733	Other woodworking machine operators				726, 733	726, 733	855	855
734	Printing machine operators, n.e.c.	520, 571, 575, 613	423, 503, 512, 615	423, 434, 515, 530, 531	734, 735, 737	734, 735, 737	824	824
736	Typesetters and compositors	x	x	x	x	x	x	x
738	Winding and twisting textile and apparel operatives			x	x	x	x	x
739	Knitters, loopers, and toppers textile operatives	x	x	x	x	x	x	x
743	Textile cutting and dyeing machine operators				743	743	836, 840	840
744	Textile sewing machine operators			x	x	x	x	x
745	Shoemaking machine operators	x	x	x	x	x	x	x
747	Clothing pressing machine operators				x	x	x	x
749	Miscellanious textile machine operators	x	x	x	x	x	x	x
753	Cementing and gluing machne operators				x	x	x	x
754	Packers, fillers, and wrappers		x	x	x	x	x	x
755	Extruding and forming machine operators				755, 758	755, 758	792, 872	792, 872
756	Mixing and blending machine operators	x	x	x	x	x	x	x
757	Separating, filtering, and clarifying machine operators				x	x	x	x
763	Food roasting and baking machine operators				x	x	x	x
764	Washing, cleaning, and pickling machine operators				x	x	x	x
765	Paper folding machine operators				x	x	x	x
766	Furnance, kiln, and oven operators, apart from food	x	x	x	x	x	x	x
769	Slicing, cutting, crushing and grinding machine	555	490	501, 612	768, 769	768, 769	785, 871	785, 871
774	Photographic process workers	x	x	x	x	x	x	x

Occ1990dd Code	Occupation Groups and Occupation Titles	Census 1950 Codes	Census 1960 Codes	Census 1970 Codes	Census 1980 Codes	Census 1990 Codes	Census 2000 (5% Sample) Codes	ACS 2005 Codes
779	Machine operators, n.e.c.	690	775	660, 690-696	673, 693, 714, 715, 717, 725, 759, 777, 779, 798	673, 693, 714, 715, 717, 725, 759, 777, 779, 798	894, 896	894, 896
783	Welders, solderers, and metal cutters	685	721	665, 680	783, 784	783, 784	771-773, 775	771-773, 775
785	Assemblers of electrical equipment		x	x	x	x	x	x
789	Painting and decoration occupations			543	789	789	881	881
799	Production checkers, graders, and sorters in manufacturing		643, 671	452, 610, 624, 625	689, 796, 797, 799	689, 796, 797, 799	874	874
	F.2 Transportation and Material Moving Occupations							
803	Supervisors of motor vehicle transportation				x	x	x	x
804	Truck, delivery, and tractor drivers	632, 683, 960	650, 715, 971, 972	705, 706, 715, 763	804-806, 856	804-806, 856	913, 960	913, 960
808	Bus drivers	x	x	x	x	x	x	x
809	Taxi cab drivers and chauffeurs	682	714	714	809, 814	809, 814	914, 915	911, 914, 915
813	Parking lot attendants			x	x	x	x	x
823	Railroad conductors and yardmasters	203, 631	252, 645	226, 704	823	823	924	924
824	Locomotive operators: engineers and firemen	x	x	x	x	x	x	x
825	Railroad brake, coupler, and switch operators	x	x	x	x	x	x	x
829	Ship crews and marine engineers	240, 673	265, 703	221, 661, 701	497, 828, 829, 833	497, 828, 829, 833	930, 931, 933	930, 931
834	Miscellanious transportation occupations	623	635		834	834	942	942
844	Operating engineers of construction equipment			x	x	x	x	x
848	Crane, derrick, winch, hoist, longshore operators	513	415	424	845, 848, 849	845, 848, 849	951, 956	951, 956
853	Excavating and loading machine operators	x	x		x	x	x	x
859	Stevedores and misc. material moving occupations	660, 940	690, 965	726, 760	843, 859, 876	843, 859, 876	965, 973, 975	965, 973, 975
865	Helpers, constructions				x	x	x	x
866	Helpers, surveyors			x	x	x	x	x
869	Construction laborers			750, 751	869	869	626, 673	626, 673
873	Production helpers				873, 874	873, 874	895	895
875	Garbage and recyclable material collectors			x	x	x	x	x
878	Machine feeders and offbearers				x	x	x	x
885	Garage and service station related occupations	x	x	x	x	x	x	x
887	Vehicle washers and equipment cleaners	x	x	x	x	x	x	x
888	Packers and packagers by hand			x	x	x	x	x
889	Laborers, freight, stock, and material handlers, n.e.c.	970	973, 985	753, 762, 770, 780, 785, 796	877, 883, 889	868, 877, 883, 889	674, 675, 962	674, 675, 962

Occupation codes from the 1950-1990 Censuses, the 2000 Census 5% sample, and the 2005 American Community Survey are indicated whenever the composition of an "occ1990dd" deviates from the "occ1990" code with the same number. The definition of all other occupations is taken from the "occ1990" crosswalk by Meyer and Osborne (2005) and "x" indicates the availability of an occupation code in a given dataset.

We first map the DOT data to the occupation codes of the 1980 Census using a Census file that relates 1970 codes to 1980 codes based on a subsample of occupation responses from the 1980 Census that were coded according to both occupation schemes. The task scores for the 1980 occupation codes can readily be mapped to the nearly identical 1990 codes. For a small number of occupations with missing data, task values are again imputed based on the task scores for similar jobs. Finally, we derive task scores for *occ1990dd* occupations by averaging the task data of all 1990 Census occupations that form a given *occ1990dd* code, weighting each 1990 Census occupation by its contribution to the total work hours of the respective *occ1990dd* code.

Following Autor et al. (2003), we use a simple average of the DOT task variables *DCP* (direction, control, and planning of activities) and *GED-MATH* (quantitative reasoning requirements) to measure "abstract tasks". The average of *STS* (adaptability to work requiring set limits, tolerances, or standards) and *FINGDEX* (finger dexterity) captures "routine tasks" and *EYEHAND* (eye, hand, foot coordination) operationalizes "manual tasks". While abstract tasks are typically concentrated in the most skilled occupations, medium- and lower-skilled jobs tend to combine routine and manual tasks. We use the routine and manual task scores of each *occ1990dd* occupation to compute an index of routine task-intensity, RTI, according to

$$RTI_k = \ln\left(\hat{R}_{k,1980}/\hat{M}_{k,1980}\right) \tag{A.9}$$

where $\hat{R}$ and $\hat{M}$ are, respectively, the intensity of routine and manual task input in for each occupation k, measured on a 0 to 10 scale. Equation (A.9) is undefined for occupations with a zero manual task score and we therefore use the manual score of the 5th percentile of the occupation distribution to derive an RTI score for these jobs.

The RTI value of an occupation provides a rough measurement of the relative importance of routine tasks in that occupation. Furthermore, since computer technology also supplies routine task inputs, RTI can be interpreted as an occupation's potential susceptibility to displacement by automation.

A.2 Commuting Zone Microdata

A.2.1 Local Labor Market Concepts

The analysis of local labor markets is motivated by the notion that employers and workers interact within a space bounded by places of work and places of residence. This local determination of market outcomes can lead to persistent geographic differences in wage and employment levels and trends (Topel, 1986).

Empirical studies of local labor market dynamics require both a geographic delineation of such markets as well as the ability to observe the necessary geographic information in the data source. In practice, the latter constraint often dictates the definition of local markets.

Due to a lack of precise geographical information, empirical studies often consider the fifty states of the United States as local labor markets (e.g., Topel, 1986). This definition has evident limitations. There is no apparent economic motivation why the boundaries of local labor markets should coincide with state boundaries. On the one hand, many states are arguably too large for being considered a single local market. On the other hand, local labor markets should reasonably be allowed to cross state boundaries. In particular, there are many urban areas overlapping with state lines (e.g., New York City/Jersey City, Washington D.C./Arlington, Kansas City MO/Kansas City KS), notably because cities developed on both sides of rivers that serve as state boundaries.

The use of counties as local labor markets (e.g., Gould, Weinberg and Mustard, 2002) provides a considerably more detailed geographic structure than states but raises similar methodological concerns. Again, local markets are restricted to be located within a single state, and while many states are too large to form a local market, counties may often be too small. Furthermore, county-level information is usually not available in publicly accessible microdata.

The most popular concept for local labor markets in recent research are Metropolitan Statistical Areas (MSAs) which are used, e.g., by Card (2001), Beaudry, Doms and Lewis (2006), and Mazzolari and Ragusa (2008). MSAs are defined for statistical purposes by the Office of Management and Budget (Office of Management and Budget, 2000) and consist of clusters of counties that cover a city and its surrounding suburbs. Various sources of microdata such as the Integrated Public Use Microdata Series of the Census (IPUMS) or the Current Population Survey (CPS) report MSA codes and data for these geographic units is thus readily available to researchers. The use of MSAs as a concept for local labor markets has certainly more economic appeal than the use of states. MSAs typically cover areas with reasonably commutable distances, and they can be located both within states or overlapping state boundaries.

The main disadvantage of MSAs is, however, that they only cover areas of the United States with major urban population centers. Moreover, the geographic definition of MSAs changes over time. MSA delineations are newly determined for every decennial Census and they are based on slightly different definition criteria over the years. In practice, the geographic area of MSAs tends to grow along with the sprawl of urban settlements.[31] An additional complication for empirical work is that many MSAs are only partially identified

[31] Jaeger, Loeb, Turner and Bound (1998) propose the use of more detailled geographic units reported in 1970-1990 IPUMS data to identify areas that roughly correspond to MSAs according to their 1970 definitions. This approach reduces, but does not eliminate, the inconsistency in the geographic measurement of MSAs.

in IPUMS data since 1980. As a consequence of these issues, an MSA code for a given city may refer to a different geographic area in every decennial Census data. The changing geography of MSAs is troublesome for research that wants to track shifts in labor force composition of local labor markets over time. The gradual addition of outlying suburbs to MSAs biases observed changes in composition towards the characteristics of suburban residents who differ substantially from urban residents.

A.2.2 Commuting Zones

The first two essays of this thesis pursue an alternative approach for the definition of local labor markets based on the concept of Commuting Zones (CZs). Commuting zones are clusters of counties that are characterized by strong commuting ties within CZs, and weak commuting ties across CZs. They have been defined by Tolbert and Killian (1987) and Tolbert and Sizer (1996) with the explicit goal of creating geographic units that capture the economic notion of local labor markets.

CZs have been computed using county-level residence-to-work commuting data from the 1980 Census and later based on such data from the 1990 Census. The strength of commuting ties between two counties i and j is measured according to

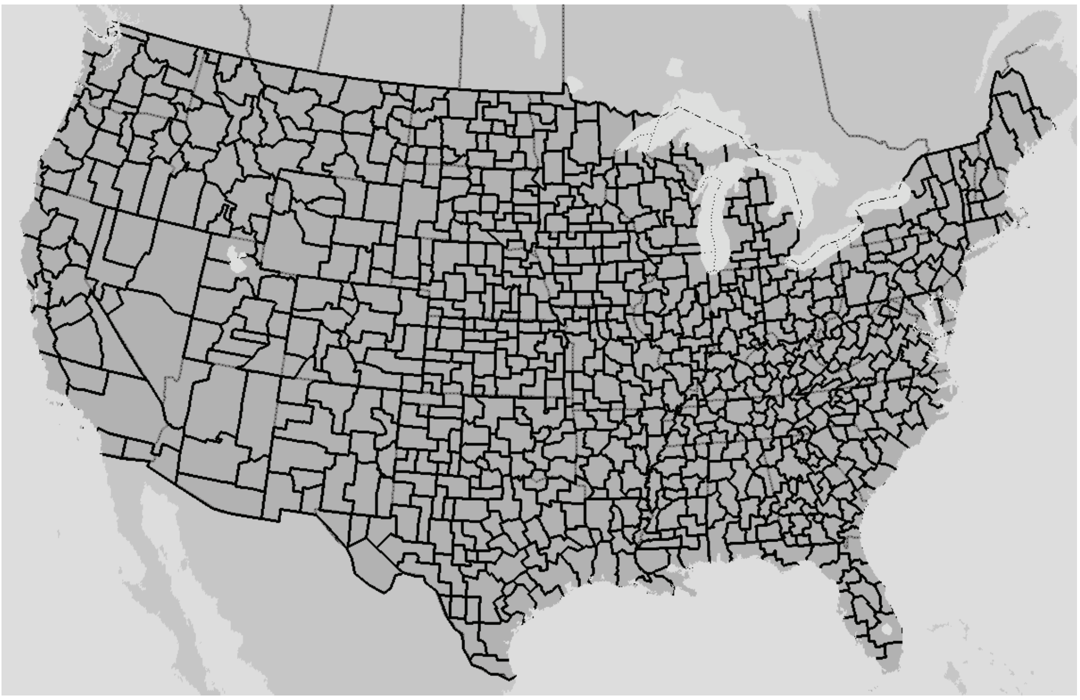
$$T_{ij} = \frac{c_{ij} + c_{ji}}{argmin(r_i, r_j)} \tag{A.10}$$

where r_i is the number of all workers residing in county i and c_{ij} is the number of workers who reside in county i but work in county j. The commuting tie statistic T_{ij} hence divides the flow of workers who commute in either direction between the two counties i and j by the number of workers who live in the smaller of the two counties. By using just the labor force of the smaller county in the denominator, the statistic is better able to identify commuting ties between counties with large size differentials where the total commuter flow might seem negligible relative to the larger county's population even if all residents of the smaller county commuted to the larger one.

CZs are derived using a clustering algorithm for average linkage that starts by grouping the county pair with largest value of T_{ij} and subsequently forms clusters of interrelated counties. CZs are defined such that the average value of T_{ij} for the county pairs in a CZ is above 0.02. By consequence, commuting ties between CZs are so weak that further aggregation would decrease the within-cluster average of T_{ij} below 0.02. Based on commuting patterns from 1980, the clustering procedure yields 764 CZs while 1990 data produces 741 CZs. In each year, the average CZ consists of about four counties. Appendix Figure 1 shows a map of the 722 CZs that result from commuting patterns in the 48 mainland states in 1990.

CZs provide an economically appealing concept of local markets because they reflect

Appendix Figure 1. Geography of Commuting Zones in 1990



the notion that employers and workers in a local market should be located within commutable distances. In contrast to MSAs, the definition of CZs does not rely on the presence of a major city and CZs hence cover the entire area and entire workforce of the United States. Despite these advantages, CZs have hardly been used in empirical economic research, likely because this geographic unit is not reported in publicly accessible microdata. The following section however develops a procedure to match IPUMS data to CZs.

A.2.3 Matching Census Microdata to Commuting Zones

The United States Census records the precise location of every respondent's residence. The publicly available IPUMS Census microdata does, however, suppress this exact geographical information. Data confidentiality laws require that public release microdata must not report geographic units that contain fewer than 100,000 residents.[32] The most detailed geographic units in IPUMS data are thus defined to comprise between 100,000 and 200,000 residents each. These units which were newly delineated for each decennial Census are alternatively called State Economic Areas (SEAs, in 1940 and 1950), County Groups (CGs, in 1970 and 1980), or Public Use Microdata Areas (PUMAs, in 1990 and

[32] The Census Bureau can, however, report average population characteristics for smaller geographic units. The third essay of this thesis uses such averaged data for Census tracts, a geographic unit of about 4000 residents.

2000). The Census Bureau did not report comparably detailed geographic information in 1960. All other geographic units reported in IPUMS data, including states and MSAs, are multiples of SEAs/CGs/PUMAs and do not provide additional geographic precision.[33] A direct identification of CZs in IPUMS data would violate confidentiality laws as some CZs do not reach the required population threshold of $100,000$ persons.[34]

We developed a procedure to match Census microdata to 1990 CZs. The key step of this procedure relates every PUMA (or SEA/CG) $j = 1, ..., J$ to every CZ $k = 1, ..., 741$ by computing the probability that a resident of j lives in k in Census year t, i.e.,

$$\alpha_{jkt} = \sum_{c=1}^{C} \frac{r_{jct}}{r_{jt}} \frac{r_{ckt}}{r_{ct}} \qquad \text{where } c = 1, ..., 3141 \text{ and } t = 1950, 1970, 1980, 1990, 2000 \tag{A.11}$$

In equation (A.13), r_{jt} is the number of residents in PUMA j in year t, r_{ct} is the number of residents in county c in year t, r_{jct} is the number of residents in the overlap between PUMA j and county c, and r_{ckt} is the number of residents in the overlap between county c and CZ k.

The share $\frac{r_{ckt}}{r_{ct}}$ of county c's population that falls into CZ k is either zero or one and can easily be determined since every county matches to exactly one CZ by definition. The fraction $\frac{r_{jct}}{r_{jt}}$ denotes the share of PUMA k's population that overlaps with a county c. While r_{jt} is observed in the microdata, r_{jct} is not. The Census Bureau provides files with population counts for each PUMA-county overlap r_{jct} for 1990 and 2000 but not for other years. In 1950 and 1970, the SEAs and CGs either map into a single county which implies $r_{jct} = r_{jt}$ when j is in c, or they are composed by multiple entire counties and hence $r_{jct} = r_{ct}$ when j contains c. In either case, r_{jct} can be determined based only on the available population counts for SEAs, CZs, and counties. The same procedure can be used for most of the 1980 CGs. A few of the 1980 CGs, particularly in the state of Connecticut, however combine parts of different counties. Since data for r_{jct} is not directly available for 1980, we aggregate these split-county CGs to aggregated CGs and then use county population counts to compute $r_{jct} = r_{ct}$ when an aggregated CG j contains a county c. Finally, the 2005 American Community Survey (ACS) uses the same PUMA definitions as the 2000 Census. We apply the α_{jkt} values of the year 2000 to the PUMAs of the 2005 ACS.

The computation of (A.13) makes use of the fact that the county structure of the United States is very stable over time. The counties that form a 1950 SEA or 1970 CG

[33]PUMAs and 1980 CGs do not always perfectly sum up to MSAs and MSA codes reported in IPUMS data hence sometimes only identify the more central parts of an MSA.

[34]With view to this constraint, Tolbert and Killian (1987) and Tolbert and Sizer (1996) aggregated smaller CZs to create Labor Market Areas (LMAs) whose population exceeds $100,000$ residents. Two special IPUMS data releases for 1980 and 1990 report these LMA codes instead of MSA codes. Unfortunately, however, the LMA samples only include microdata for one percent of the population which creates a considerably larger sampling error as opposed to the standard five percent samples with MSA codes.

are typically the same geographic units as the counties that make up the target geography, 1990 CZs. We adjust for the few counties that split or merged between 1950 and 2000 by mapping each year's county structure to the 3141 counties of the year 1990. Most of these changes occurred in Alaska which is divided into political units called boroughs and a set of Census Areas that serve statistical only purposes. We do not map data from Alaska and Hawaii to CZs in 1950 because the Census did not designate SEAs for these territories prior to statehood. The empirical analysis of CZs between 1950 and 2005 therefore focuses on the 722 CZs that cover the 48 mainland states.

In order to map Census microdata to CZs, we replace every individual microdata observation $n = 1, ..., N$ of Census year t with 722 observations that are identical to the initial observation except for the person weight which is adjusted from the initial Census person weight w_{nt} to

$$\omega_{nkt} = w_{nt}\alpha_{jkt} \qquad \text{where } k = 1, ..., 722 \text{ and } t = 1950, 1970, 1980, 1990, 2000, 2005 \tag{A.12}$$

Observations are dropped whenever $\alpha_{jkt} = 0$ and the number of observations per person is thus equal to the number of CZs that overlap with the PUMA in which the person resides. The adjusted person weights ω_{nkt} in the resulting dataset sum to the original weights w_{nt} for each person n, i.e.,

$$\sum_{k=1}^{722} \omega_{nkt} = w_{nt} \qquad \forall n = 1, ..., N \tag{A.13}$$

This procedure essentially 'splits' a person into multiple parts whenever an individuals' PUMA cannot be uniquely assigned to a CZ. In practice, however, many PUMAs match fully into a single CZ. Of the 2071 PUMAs in the 2000 Census, only 19 percent overlap with several CZs.

The individual-level CZ data is observed with noise due to the need to match PUMAs to CZs. However, it allows studying wage and employment dynamics in local labor markets that are defined according to a tractable economic definition and with fixed boundaries. While an analysis of MSAs over time is usually biased by the inclusion of an ever larger geographic area, the CZ geography consistently covers the entire area of the United States.

References

Acemoglu, Daron (1999), "Changes in Unemployment and Wage Inequality: An Alternative Theory and Some Evidence." *American Economic Review*, 89 (December), 1259–1278.

Ananat, Elisabeth Oltmans (2007), "The Wrong Side(s) of the Tracks: Estimating the Causal Effects of Segregation on City Outcomes."National Bureau of Economic Research Working Paper No. 13343.

Angrist, Joshua D. and Kevin Lang (2004), "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program." *American Economic Review*, 94 (December), 1613–1634.

Autor, David H. and David Dorn (2008), "Inequality and Specialization: The Growth of Low-Skilled Service Employment in the United States."Mimeograph Massachusetts Institute of Technology, November 2008.

Autor, David H., Frank Levy, and Richard J. Murnane (2003), "The Skill Content of Recent Technological Change: An Empirical Investigation." *Quarterly Journal of Economics*, 118 (November), 1279–1333.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney (2006), "The Polarization of the U.S. Labor Market." *American Economic Review Papers and Proceedings*, 96 (May), 189–194.

——— (2008), "Trends in U.S. Wage Inequality: Revising the Revisionists." *Review of Economics and Statistics*, 90 (May), 300–323.

Bajari, Patrick and Mathew E. Kahn (2005), "Estimating Housing Demand with an Application to Explaining Racial Segregation in Cities." *Journal of Business and Economics Statistics*, 23 (January), 20–33.

Baum-Snow, Nathaniel (2007), "Did Highways Cause Suburbanization?" *Quarterly Journal of Economics*, 115 (May), 775–805.

Baum-Snow, Nathaniel and Byron Lutz (2008), "School Segregation, School Choice and Changes in Residential Location Patterns by Race."Mimeograph, August 2008.

Baumol, William J. (1967), "Macroeconomics of Unbalanced Growth: Anatomy of an Urban Crisis." *American Economic Review*, 57 (June), 415–426.

Bayer, Patrick, Hanming Fang, and Robert McMillan (2005), "Separate when Equal? Racial Inequality and Residential Segregation."National Bureau of Economic Research Working Paper No. 11507, October 2005.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan (2007), "A Unified Framework for Measuring Preferences for Schools and Neighborhoods." *Journal of Political Economy*, 115 (August), 588–638.

Beaudry, Paul and David A. Green (2003), "Wages and Employment in the United States and Germany: What Explains the Differences?" *American Economic Review*, 93, 573–602.

Beaudry, Paul, Mark Doms, and Ethan G. Lewis (2006), "Endogenous Skill Bias in Technology Adoption. City-Level Evidence from the IT Revolution."NBER Working Paper 12521, September 2006.

Becker, Gary S. and Kevin M. Murphy (2000), *Social Economics: Market Behavior in a Social Environment*, Cambridge MA: Harvard University Press.

Benabou, Roland (1993), "Workings of a City: Location, Education, and Production." *Quarterly Journal of Economics*, 108 (August), 619–652.

Blanchard, Olivier J. and Lawrence F. Katz (1992), "Regional Evolutions." *Brookings Papers on Economic Activity*, 1992, 1–75.

Blinder, Alan (2007), "How Many U.S. Jobs Might be Offshorable?"Princeton University Center for Economic Policy Studies Working Paper No. 142, March 2007.

Bound, John and Harry J. Holzer (2000), "Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s." *Journal of Labor Economics*, 18 (January), 20–54.

Boustan, Leah Platt (2007a), "Escape from the City? The Role of Race, Income, and Local Public Goods in Post-War Suburbanization."National Bureau of Economic Research Working Paper No. 13311, August 2007.

——— (2007b), "Was Postwar Suburbanization 'White Flight'? Evidence from the Black Migration."National Bureau of Economic Research Working Paper No. 13543, October 2007.

Card, David (2001), "Immigrant Inflows, Native Outflows, and the Local Market Impacts of Higher Immigration." *Journal of Labor Economics*, 19 (January), 22–44.

Card, David and Jesse Rothstein (2007), "Racial Segregation and the Black-White Test Score Gap." *Journal of Public Economics*, 91 (December), 2158–2184.

Card, David, Alexandre Mas, and Jesse Rothstein (2008a), "Tipping and the Dynamics of Segregation." *Quarterly Journal of Economics*, 123 (February), 177–218.

——— (2008b), "Are Mixed Neighborhoods Always Unstable? Two-Sided and One-Sided Tipping."Mimeograph, March 2008.

Charles, Kerwin Kofi and Jonathan Guryan (2008), "Prejudice and Wages: An Empirical Assessment of Becker's *The Economics of Discrimination*." *Journal of Political Economy*, 116 (October), 773–809.

Charles, Kerwin Kofi and Erik Hurst (2002), "The Transition to Home Ownership and the Black-White Wealth Gap." *Review of Economics and Statistics*, 84 (May), 281–297.

Clark, Colin (1957), *The Conditions of Economic Progress*, London: Macmillan.

Collins, William J. and Robert A. Margo (2001), "Race and Home Ownership: A Century-Long View." *Explorations in Economic History*, 38 (May), 68–92.

Cortes, Patricia (2008), "The Effect of Low-Skilled Immigration on US Prices: Evidence from CPI Data." *Journal of Political Economy*, 116, 381–422.

Coulson, N. Edward (1999), "Why are Hispanic- and Asian-American Homeownership Rates so Low? Immigration and Other Factors." *Journal of Urban Economics*, 45 (March), 209–227.

Coulson, N. Edward and Eric W. Bond (1990), "A Hedonic Approach to Residential Succession." *Review of Economics and Statistics*, 72 (August), 433–444.

Cutler, David M. and Edward L. Glaeser (1997), "Are Ghettos Good or Bad?" *Quarterly Journal of Economics*, 112 (August), 827–872.

Cutler, David M., Edward L. Glaeser, and Jacob L. Vigdor (1999), "The Rise and Decline of the American Ghetto." *Journal of Political Economy*, 107 (June), 455–506.

——— (2008), "Is the Melting Pot Still Hot? Explaining the Resurgence of Immigrant Segregation." *Review of Economics and Statistics*, 90 (August), 478–497.

DiNardo, John E., Nicole M. Fortin, and Thomas Lemieux (1996), "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica*, 64, 1001–1044.

DiPasquale, Denise and Edward L. Glaeser (1999), "Incentives and Social Capital: Are Homeowners Better Citizens." *Journal of Urban Economics*, 45 (March), 354–384.

Doms, Mark E. and Ethan G. Lewis (2006), "Labor Supply and Personal Computer Adoption."Federal Reserve Bank of Philadelphia Working Paper No. 06-10, June 2006.

Doms, Mark E., Timothy Dunne, and Kenneth R. Troske (1997), "Workers, Wages, and Technology." *Quarterly Jornal of Economics*, 112, 253–290.

Dustmann, Christian, Johannes Ludsteck, and Uta Schoenberg (2007), "Revisiting the German Wage Structure."Mimeograph, University of Rochester, February 2007.

Easterly, William (2005), "Empirics of Strategic Interdependence: The Case of the Racial Tipping Point."New York University DRI Working Paper No. 5.

Farley, Reynolds, Elaine L. Fielding, and Maria Krysan (1997), "The Residential Preferences of Blacks and Whites: A Four-Metropolis Analysis." *Housing Policy Debate*, 8 (4), 763–800.

Frankel, David M. and Ady Pauzner (2002), "Expectations and the Timing of Neighborhood Change." *Journal of Urban Economics*, 51 (June), 295–314.

Fu, Yuming (1991), "A Model of Housing Tenure Choice: Comment." *American Economic Review*, 81 (June), 381–383.

Glaeser, Edward L. and Joseph Gyourko (2005), "Urban Decline and Durable Housing." *Journal of Political Economy*, 113, 345–375.

——— (2007), "Arbitrage in Housing Markets."National Bureau of Economic Research Working Paper No. 13704.

Goos, Maarten and Alan Manning (2003), "Lousy and Lovely Jobs: The Rising Polarisation of Work in Britain."CEPR Working Paper No. 0604, December 2003.

——— (2007), "Lousy and Lovely Jobs: The Rising Polarisation of Work in Britain." *Review of Economics and Statistics*, 89, 118–133.

Goos, Maarten, Alan Manning, and Anna Salomons (2008), "Job Polarization in Europe."Mimeograph London School of Economics, December 2008.

Gould, Eric D., Bruce A. Weinberg, and David B. Mustard (2002), "Crime Rates and Local Labor Market Opportunities in the United States: 1979-1997." *Review of Economics and Statistics*, 64 (February), 45–61.

Guryan, Jonathan (2004), "Desegregation and Black Dropout Rates." *American Economic Review*, 94 (September), 919–943.

Harrison, Bennett and Barry Bluestone (1988), *The Great U-turn: Corporate Restructuring and the Polarizing of America*, New York: Basic Books.

Henderson, J. Vernon and Yannis M. Ioannides (1983), "A Model of Housing Tenure Choice." *American Economic Review*, 73 (March), 98–113.

Hoff, Karla and Arijit Sen (2005), "Homeownership, Community Interactions, and Segregation." *American Economic Review*, 95 (September), 1167–1189.

Jacob, Brian A. (2004), "Public Housing, Housing Vouchers, and Student Achievement: Evidence from Public Housing." *American Economic Review*, 94 (March), 233–258.

Jaeger, David A., Susanna Loeb, Sarah E. Turner, and John Bound (1998), "Coding Geographic Areas across Census Years: Creating Consistent Definitions of Metropolitan Areas."National Bureau of Economic Research Working Paper No. 6772.

Jones, Lawrence D. (1990), "Current Wealth Constraints on the Housing Demand of Young Owners." *Review of Economics and Statistics*, 72 (August), 424–432.

Kain, John F. (1968), "Housing Segregation, Negro Employment, and Metropolitan Decentralization." *Quarterly Journal of Economics*, 82 (May), 175–197.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz (2007), "Experimental Analysis of Neighborhood Effects." *Econometrica*, 75 (January), 83–119.

Leonardi, Marco (2008), "Product Demand Shifts and Wage Inequality."Mimeograph, University of Milan, October 2008.

Manning, Alan (2004), "We Can Work it Out: the Impact of Technological Change on the Demand for Low-Skill Workers." *Scottish Journal of Political Economy*, 51 (November), 581–608.

Margo, Robert A. (1992), "Explaining the Postwar Suburbanization of Population in the United States: The Role of Income." *Journal of Urban Economics*, 31 (May), 301–310.

Mazzolari, Francesca and Giuseppe Ragusa (2008), "Spillovers from High-Skill Consumption to Low-Skill Labor Markets."Mimeograph, University of California at Irvine, May 2008.

Meyer, Peter B. and Anastasiya M. Osborne (2005), "Proposed Category System for 1960-2000 Census Occupations."Bureau of Labor Statistics Working Paper 383, September 2005.

Michaels, Guy (2007), "The Division of Labor, Coordination, and the Demand for Information Processing."CEPR Working Paper 6358, June 2007.

Miyao, Takahiro (1978), “Dynamic Stability of a Mixed City in the Presence of Neighborhood Externalities.” *American Economic Review*, 68 (June), 454–463.

——— (1979), “Dynamic Stability of an Open City with Many Household Classes.” *Journal of Urban Economics*, 6 (July), 292–298.

Moebius, Markus and Tanya Rosenblat (2001), “The Process of Ghetto Formation: Evidence from Chicago.”Mimeograph, Harvard University 2001.

Neal, Derek (1999), “The Complexity of Job Mobility among Young Men.” *Journal of Labor Economics*, 17, 237–261.

Ngai, Rachel L. and Christopher Pissarides (2007), “Structural Change in a Multisector Model of Growth.” *American Economic Review*, 97 (March), 429–443.

Office of Management and Budget (2000), “Standards for Defining Metropolitan and Micropolitan Areas.” *Federal Register*, 65 (December 27), 82228–82238.

Oliner, Stephen D. and Daniel E. Sichel (2000), “The Resurgence of Growth in the Late 1990s: Is Information Technology the Story?” *Journal of Economic Perspectives*, 14, 3–22.

Oreopoulos, Philip (2003), “The Long-Run Consequences of Living in a Poor Neighborhood.” *Quarterly Journal of Economics*, 118 (November), 1533–1575.

Peri, Giovanni and Chad Sparber (2007), “Task Specialization, Comparative Advantages and the Effects of Immigration on Wages.”Mimeograph, University of California at Davis, July 2007.

Pinker, Steven (2007), *The Stuff of Thought: Language as a Windowinto Human Nature*, New York: Viking Adult (Penguin).

Ruggles, Steven, Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander (2004), *Integrated Public Use Microdata Series: Version 3.0 [Machine-readable database]*, Minneapolis: Minnesota Population Center [producer and distributor].

Saiz, Albert and Susan Wachter (2006), “Immigration and the Neighborhood.”Institute for the Study of Labor (IZA) Discussion Paper No. 2503.

Sampson, Robert J., Stephen W. Raudenbush, and Felton Earls (1997), “Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy.” *Science*, 277 (August), 918–924.

Schelling, Thomas C. (1969), "Models of Segregation." *American Economic Review*, 2 (May), 488–493.

——— (1971), "Dynamic Models of Segregation." *Journal of Mathematical Sociology*, 1 (July), 143–186.

——— (1978), *Micromotives and Macrobehavior*, New York: Norton.

Sinai, Todd and Nicholas S. Souleles (2005), "Owner-Occupied Housing as a Hedge Against Rent Risk." *Quarterly Journal of Economics*, 120 (May), 763–789.

Smith, Christopher L. (2008), "Implications of Adult Labor Market Polarization for Youth Employment Opportunities."Mimeograph, Massachusetts Institute of Technology, July 2008.

Spitz-Oener, Alexandra (2006), "Technical Change, Job Tasks and Rising Educational Demands: Looking Outside the Wage Structure." *Journal of Labor Economics*, 24 (April), 235–270.

Tiebout, Charles (1956), "A Pure Theory of Local Expenditures." *Journal of Political Economy*, 64 (October), 416–424.

Tolbert, Charles M. and Molly Sizer Killian (1987), "Labor Market Areas for the United States."Staff Report No. AGES870721. Washington, DC: Economic Research Service, US Department of Agriculture.

Tolbert, Charles M. and Molly Sizer (1996), "U.S. Commuting Zones and Labor Market Areas. A 1990 Update."Economic Research Service Staff Paper No. 9614.

Topel, Robert H. (1986), "Local Labor Markets." *Journal of Political Economy*, 94, 111–143.

U. S. Department of Labor, Employment and Training Administration (1977), *Dictionary of Occupational Titles: Fourth Edition*, Washington DC: US Government Printing Office.

Vigdor, Jacob and Jens Ludwig (2007), "Segregation and the Black-White Test Score Gap."National Bureau of Economic Research Working Paper No. 12988.

Weiss, Matthias (2008), "Skill-biased Technical Change: Is There Hope for the Unskilled?" *Economics Letters*, 100 (September), 439–441.

Zax, Jeffrey S. and John F. Kain (1996), "Moving to the Suburbs: Do Relocating Companies Leave Their Black Employees behind." *Journal of Labor Economics*, 14 (July), 472–504.

Curriculum Vitae

Personal

Born April 4, 1979.

Swiss Citizen.

Education

PhD in Economics (Dr. oec.), University of St. Gallen, 2009.

Thesis Committee: David H. Autor, Monika Bütler, Alfonso Sousa-Poza.

Program for Doctoral Students, Swiss National Bank Study Center Gerzensee, 2006.

M.Sc. in Internat. Management, Community of European Management Schools, 2004.

M.Sc. in Economics (lic. oec.), University of St. Gallen, 2004.

Exchange Student, University of North Carolina at Chapel Hill, 2004.

Exchange Student, ESADE Barcelona, 2002.

Research Positions and Affiliations

Research Affiliate, Institute for the Study of Labor (IZA Bonn), since 2008.

Visiting Researcher, Boston University Department of Economics, 2008-2009.

Visiting Researcher, MIT Department of Economics, 2007-2008.

Visiting Researcher, University of Chicago Harris School of Public Policy Studies, 2007.

Research Assistant, University of St. Gallen FAA Institute, 2000-2006.