UCLA
UCLA Electronic Theses and Dissertations

Title
Essays on Spatial Disparities in Labor and Housing Markets

Permalink
https://escholarship.org/uc/item/8ks4h38k

Author
Freyd, Benjamin

Publication Date
2022

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA

Los Angeles

Essays on Spatial Disparities in Labor and Housing Markets

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Economics

by

Benjamin Etienne Freyd

2022

© Copyright by
Benjamin Etienne Freyd
2022

ABSTRACT OF THE DISSERTATION

Essays on Spatial Disparities in Labor and Housing Markets

by

Benjamin Etienne Freyd
Doctor of Philosophy in Economics
University of California, Los Angeles, 2022
Professor Moshe Buchinsky, Co-Chair
Professor Maurizio Mazzocco, Co-Chair

My first chapter studies the origins of gentrification. I propose a mechanism through which a wave of gentrification can be triggered in a neighborhood: the opening of large offices of technology firms. Using information on nine such events and transaction-level housing data, I develop a difference-in-differences strategy that compares house prices in the vicinity of the new office to those in closely-matched neighborhoods slightly further away. I find that property prices rise 11% in treatment areas relative to control areas within two years after the opening. This difference subsides somewhat but remains at +8% five years later. These findings are substantially stronger than in the existing literature and thus suggest that such office openings can have a major impact on their neighborhood. I investigate two mechanisms, agglomeration forces and the development of consumption amenities, through which the impact of a single establishment opening can be amplified and sustained.

My second chapter analyses the causes of the shift towards low-income service jobs that American non-college workforce has experience in recent decades. Together with the rise of college-educated workers' incomes, this has strongly contributed to the overall increase in wage inequality. Two main explanations, routinization and consumption spillovers, have been proposed to explain this

occupational shift. Although the determinants of these two theories are highly spatially correlated, studies that have exploited regional variations to identify these mechanisms have so far only considered each explanation in isolation, raising confounding concerns. I highlight these concerns and provide reduced-form evidence that both theories operate simultaneously to drive growth in service employment. To strengthen my case, I extend the structural framework proposed in Autor and Dorn (2013) to include consumption spillovers through non-homothetic preferences. I estimate key parameters and assess the relative importance of each theory using simulations from the model. Relative to a model featuring homothetic preferences, my specification yields 57% more regional disparities in the growth of service occupations, which can be interpreted as the contribution of consumption spillovers. While the routinization hypothesis quantitatively dominates, my reduced-form and structural evidence point to sizeable consumption spillover effects that cannot be neglected.

My third chapter studies the effects of liberalizing the use of short-term contracts on the labor market outcomes of all workers. It specifically examines the impact of a 2003 change French jurisprudence. This decision from the French Civil Supreme Court — hereafter called the "reform" — extends the scope of a specific type of temporary contract in France, the CDDU (*contrat à durée déterminée d'usage*), to jobs that are not necessarily temporary by nature. This type of contract is allowed in 16 service sectors and is not restricted in terms of length or number of renewals, giving a lot more flexibility to employers than the standard temporary contract does (CDD, *contrat à durée déterminée*). I find that this change is associated with an increase in the share of temporary contracts of 2.9 percentage points in the CDDU sectors, while other service sectors only show a 0.7 point increase. This result echoes a frequent finding of temp for permanent substitution in the literature about the deregulation of temporary contract. A second and novel finding is that this reform weighs on young permanent workers' wages, who experience drop of nearly 4% in their wages over the two years following the reform, and little catch up afterwards. I show some evidence that something changed in the relationship between employers and employees in CDDU sectors: the employer's bargaining power seems to have risen as the opportunity to substitute temporary contracts for costly and protected permanent ones has increased.

The dissertation of Benjamin Etienne Freyd is approved.

Michela Giorcelli

Daniel Haanwinckel

Stuart Gabriel

Moshe Buchinsky, Committee Co-Chair

Maurizio Mazzocco, Committee Co-Chair

University of California, Los Angeles

2022

Contents

1 Large Tech Office Openings and the Onset of Gentrification **1**

1.1 Introduction 1

1.2 Data 3

1.2.1 Major Office Openings 3

1.2.2 House Prices 3

1.2.3 Demographics 4

1.2.4 Final Dataset 4

1.3 Empirical Analysis 5

1.3.1 Treatment and Control Design 5

1.3.2 Baseline Comparisons 6

1.3.3 Event Study Model 7

1.3.4 Endogeneity Discussion 8

1.4 Results 9

1.4.1 Event Study Results 9

1.4.2 Discussion 10

1.5 Robustness 12

1.5.1 Sensitivity of Results to Design Parameters 12

1.5.2 Time Placebo Tests . . . 12
1.5.3 Geographical Placebo Tests . . . 13
1.6 Conclusion . . . 13
1.7 Appendix . . . 13
Figures . . . 15
Tables . . . 23

2 Labor Market Polarization and the Growth of Service Employment: Routinization or Consumption Spillovers? 29

2.1 Introduction . . . 29
2.2 Data . . . 33
2.2.1 Employment and Wages . . . 34
2.2.2 Occupational Classifications . . . 34
2.2.3 Consumer Expenditures Data . . . 36
2.2.4 Capital Income Data . . . 37
2.2.5 Price of Capital . . . 38
2.3 Empirical Evidence . . . 38
2.3.1 Noncollege Service Employment Has Grown Faster than Other Occupations 39
2.3.2 The Adoption of Computer Technology is Correlated with Routine Jobs Destruction and College Wage Growth . . . 39
2.3.3 Evidence for the Routinization Theory . . . 40
2.3.4 Evidence for the Consumption Spillovers Theory . . . 42
2.3.5 Confronting the two theories in a regression framework . . . 44
2.4 Model . . . 46
2.4.1 Production . . . 47

2.4.2 Capital . . . 48
2.4.3 Workers . . . 48
2.4.4 Equilibrium . . . 50
2.5 Model Estimation . . . 51
2.5.1 Mapping between Model and Data Locations . . . 52
2.5.2 Calibrated Parameters . . . 53
2.5.3 Estimated Parameters and Identification Conditions . . . 53
2.5.4 Data Moments and GMM Conditions . . . 56
2.5.5 Estimation Results . . . 57
2.6 Discussion of Model Results . . . 58
2.6.1 Simulations . . . 58
2.6.2 Capital Adoption and Routine Labor Displacement . . . 58
2.6.3 The Growth of Service Employment and Wages . . . 58
2.6.4 Impact of Consumption Spillovers . . . 59
2.6.5 Robustness . . . 61
2.7 Conclusion . . . 61
2.8 Appendix . . . 62
2.8.1 Data Appendix . . . 62
2.8.2 Model Appendix . . . 63
Figures . . . 67
Tables . . . 85

3 Deregulated Temporary Contracts and Workers' Outcomes: Evidence from a 2003 French Reform 100

3.1 Introduction 100
3.2 Context 104
3.2.1 Overview of the French labor contract system 104
3.2.2 The specific case of CDDU 106
3.2.3 The change in the CDDU jurisprudence 107
3.3 Data 108
3.3.1 Data sources and sample selection 108
3.3.2 Weighting and seasonal corrections 111
3.3.3 Descriptive statistics on the Employment Survey 112
3.4 Results 113
3.4.1 Substitution between temporary and permanent contracts 113
3.4.2 Wages 115
3.5 Conclusion and future research 119
Figures 122
Tables 132

List of Figures

1.1 Sample composition: number of property transactions happening in the treatment and control groups 15

1.2 Baseline comparison of selected property characteristics between treatment and control groups 16

1.3 Event study results 17

1.4 Evolution of the workforce residing in treatment areas in three sectors (finance, management and professional/scientific services) 18

1.5 Evolution of the workforce working in treatment areas in two sectors (arts and entertainment, and accomodation and food) 19

1.6 Robustness tests on 27 empirical design configurations. 20

1.7 Time placebo tests 21

1.8 Treatment and control areas selected in the main specification 22

2.1 1980 share of noncollege routine employment and 1980-2005 change in noncollege employment share of service occupations 67

2.2 Relation between the share of a household's consumption expenditures allocated to services and the hourly wage of its head 68

2.3 Spatial distributions of college and noncollege wages in 1980 and 2005 69

2.4 1980-2005 growth in college wages contemporaneous change in noncollege employment share of service occupations 70

2.5 Relation between the determinants of the routinization and consumption spillover theories: initial share of noncollege routine employment and growth of college wages 71

2.6 Diagram for mechanisms of routinization and consumption spillovers 72

2.7 Diagram for the effect of non-homotheticity on service expenditure shares 73

2.8 Map of the correspondence between commuting zones in the data and locations (1,2) in the model 74

2.9 Price of capital, 1980-2005 75

2.10 Model simulations: Capital and Routine Employment 76

2.11 Model simulations: Service Employment and Wages 77

2.12 Model simulations: College Wages and College Expenditure Shares on Services . . 78

2.13 Model simulations: Routine and Service workers' Expenditure Shares on Services . 79

2.14 Model simulations: Contributions of worker types to increases in service consumption and employment 80

2.15 Model robustness: simulations and effect decomposition with a lower production elasticity $\mu = 0.15$ 81

2.16 Model robustness: simulations and effect decomposition with a higher production elasticity $\mu = 0.35$ 82

2.17 Model robustness: simulations and effect decomposition with a lower consumption elasticity $\sigma = 0.15$ 83

2.18 Model robustness: simulations and effect decomposition with a higher consumption elasticity $\sigma = 0.25$ 84

3.1 Shares of the three main contract types in total employment under contract. 122

3.2 Employment in CDDU and non-CDDU sectors, as a share of total service sector employment. 123

3.3 Shares of CDD and CDI in CDD+CDI total, in CDDU sectors and in other service sectors 124

3.4 Average share of employment in different CDDU sectors over 1998-2003 and over 2004-2008, as part of total CDDU employment. 125

3.5 Shares of CDD in CDD+CDI, in CDDU sectors, before and after the reform 126

3.6 Number of hires by contract type and by sector group 127

3.7 Evolution of log wages by age bracket and by sector group 128

3.8 Coefficient estimates β_t from the year-fixed-effects difference-in-differences regression 129

3.9 Evolution of log wages by sectoral subgroups that contain both CDDU and non-CDDU service sectors 130

3.10 Oaxaca decompositions for occupation type and education level for young workers in permanent contracts 131

List of Tables

1.1 Sample of tech office openings. 23

1.2 Baseline comparison of selected demographic characteristics between treatment and control groups 24

1.3 Regression estimates for the effect of office openings on local house prices 25

2.1 Top 10 service occupations by their 1980 employment share within services. 85

2.2 Top 10 routine occupations by their 1980 employment share, within routine occupations. 86

2.3 Expenditure categories in the CEX 2013-2019 classified as service expenditures. . 87

2.4 Employment and wage data for workers without a college degree, by broad occupation groups in 1980 and 2005. 88

2.5 Effect of personal computer adoption in businesses on college wages and routine employment 89

2.6 Share of routine employment by broad occupational groups. 90

2.7 Regression estimates: Initial Routine Employment Share, College Wage Growth and Growth of Service Employment 91

2.8 Regression estimates: Initial Routine Employment Share, Other Noncollege Wage Growth and Growth of Service Employment 92

2.9 Regression estimates: Routine Employment Growth, College Wage Growth and Growth of Service Employment 93

2.10 Calibrated parameters for all models 94

2.11 GMM estimates for models with homothetic and non-homothetic preferences, without labor mobility 95

3.1 Sectors allowed to use the CDDU contract 132

3.2 Effect of adding quarters 2-4 data on the wage gap between CDDU and other service sectors 133

3.3 Summary statistics from the French Employment Survey, by period (pre/post reform) and by sector group 134

3.4 Log wages difference-in-differences estimates for different age-brackets 135

3.5 Log wages difference-in-differences estimates by sectoral subgroups containing both CDDU and non-CDDU sectors 136

3.6 Log wages difference-in-differences by contract type for workers aged 15-34 y.o. . . 137

ACKNOWLEDGEMENTS

I am deeply grateful to everyone who has helped me through this journey – family, professors, fellow students and friends. First and foremost, I would like to thank my two main advisors, Moshe Buchinsky and Maurizio Mazzocco, for their invaluable guidance and support throughout the program. I would like to thank Stuart Gabriel for being always around, offering his help, financial support and opportunities to work on extremely interesting projects. I would also like to thank Daniel Haanwinckel for his special generosity in helping me with the conceptual and technical aspects of my job market paper, as well as reassuring me regarding the quality of my work on multiple occasions. Finally, I would like to thank Michela Giorcelli for her kindness, support and outstanding pieces of advice on research presentations. Together, this dissertation committee has allowed me to spend six amazing years at UCLA and to pull through every struggle inherent to doctoral endeavors. I would also like to thank Lee Ohanian, Pierre-Philippe Combes, Zsofia Barany, Clémence Tricaud, Gregor Schubert, Lorenzo Cattivelli, Tomas Guanziroli, Mark Edmonds and seminar participants at UCLA Econ and CCPR for insightful conversations and feedback.

I am grateful to the Howard and Irene Levine Program in Housing and Social Responsibility for financially supporting the research that makes the first chapter of my dissertation. I would also like to thank Stuart Gabriel and the Ziman Center for Real Estate at UCLA for additional support on this project.

I would like to thank my parents, Geneviève and Denis, for their unconditional support. I am indebted to Alexandre Tiard for being invariably understanding and supportive, for his outworldly cooking skills and his occasionally funny jokes. Special thanks also go to Alvaro Boitier for his help on the modeling part of my job market paper and his taste for lesser-known music venues. Many friends have also contributed to making this journey delightful – listing them here would be too long, but they will know I thought of them.

VITA

2007-2011	B.S. Mathematics and Physics, Ecole Centrale de Paris
2011-2015	M.S. Applied Mathematics, Ecole Centrale de Paris
2013-2015	M.A. Economics, Paris School of Economics
2015-2016	Economist, OFCE - Sciences Po
2016-2020	Economics Department Fellowship, UCLA
2017-2022	Teaching Fellow, Economics Department, UCLA
2018-2019	Research Assistant, Anderson Business School, UCLA
2019	TA Award for Outstanding Performance (Spring 2019), Economics Department, UCLA
2019	Howard and Irene Levine Program in Housing and Social Responsibility Research Grant, UCLA
2019	Ziman Center for Real Estate Data Grant, UCLA

Chapter 1

Large Tech Office Openings and the Onset of Gentrification

1.1 Introduction

Gentrification has recently re-emerged as a major question in several US cities. Former low-income neighborhoods, especially in California and New York City, have seen a fresh inflow of wealthy residents bringing along substantial neighborhood changes. This phenomenon has a priori ambiguous welfare effects. On the one hand, it tends to increase property values, benefiting incumbent homeowners; it also improves local economic dynamism, with the potential creation of more businesses, especially in the service sector (Glaeser, Kim, and Luca, 2018). Both of these are seen as positive impacts of gentrification. On the other hand however, gentrification can out-price lower-income renters, thereby reinforcing geographical income segregation (Couture, Gaubert, et al., 2019; Berkes and Gaetani, 2018). Households who have to move out may not only lose their place, but they may have to move somewhere with worse economic opportunities (Ganong and Shoag, 2017). Overall, the evidence about the effects of gentrification remains mixed, suggesting a lot of heterogeneity effects (Meltzer, 2016; Meltzer and Ghorbani, 2017).

In the economic literature, gentrification is mostly seen as an endogenous, snow-balling effect: wealthy residents attract more of their kind through the development of specific neighborhood amenities. This endogenous amenity development is at the core of modern urban models (see e.g.

Diamond, 2016). While many factors have been evoked (Hwang and Lin, 2016)little is known about what initially triggers a wave of gentrification. Policymakers could highly benefit from knowing whether their city is about to go through such a wave, in order to enact protective policies for renters or ensure sufficient new housing supply, both of which may take years to reach. So far, the presence of artistic and creative businesses has been associated with the future gentrification of a neighborhood (Behrens, Boualam, et al., 2018; Schuetz, 2014). Supermarket and coffee shop openings have also been associated with mild increases in house prices Glaeser, Kim, and Luca, 2018; D. G. Pope and J. C. Pope, 2015. But a clear causal link is hard to establish between the appearance of these businesses and a slowly ensuing gentrification process. In a lot of cases, either a fine enough spatial or temporal scale is missing and that is a threat to identification.

In this work, I study the impact of major technology office[1] openings on house prices in their neighborhood. These events have the potential of bringing a substantial amount of well-paid workers around the office location, thereby abruptly changing a neighborhood. Numerous anecdotal accounts of such effects have been reported (see e.g. Dave and Vincent, 2017; Rosenberg, 2018). Major industrial openings have also been shown to have sizeable effects on local labor and housing markets at a coarser scale (Greenstone, Hornbeck, and Moretti, 2010). I gather data on a number of major office openings, including their exact address and the year of opening. I look at the impact of these events on house prices by using transaction-level data on properties, controlling for housing quality through multiple property characteristics, and using fine-grained location fixed effects. To control for the endogeneity of firm location choices, I design treatment and control groups based on (i) their geographical proximity and (ii) their similarity in house prices prior to the opening event.

My results suggest a strong causal link between the office opening event and a rise in local house prices. In the main specification, I find that house prices rise 11% within 1 km of the new office, relative to matched areas located between 1 and 3 km away, within 2 years of the office opening. This difference persists at around 8% five years after. These effects are substiantially larger those found in the case of supermarket (D. G. Pope and J. C. Pope, 2015) and coffee shop (Glaeser, Kim, and Luca, 2018) openings. The results are qualitatively robust to a number of changes in the design of the treatment and control groups. I explore two important potential mechanisms behind

[1] e.g. a Google office dedicated to software engineering.

these findings, to understand their size and persistence. I assess the role of agglomeration forces, whereby new technology offices attract similar companies, creating a long-lasting snow-balling effect (Davis and Dingel, 2019; Gaubert, 2018; Behrens, Duranton, and Robert-Nicoud, 2014; Ellison, Glaeser, and Kerr, 2010; Rosenthal and Strange, 2001). I also assess the role of the development of consumption amenities around the office location, that can then attract a broader wealthy population (Couture and Handbury, 2017).

The paper is organized as follows. Section 3.3 presents the data. Section 1.3 describes the design of treatment and control groups and the main model specification for the empirical analysis. Section 1.4 presents the results and dicusses them. Section 1.5 explores the robustness of the results.

1.2 Data

1.2.1 Major Office Openings

I obtain data from various online and press sources about major openings of large technology offices. I obtain their exact addresses as well as their opening or opening-announcement year. In cases where I have announcement date information, I confirm that the office has opened and assume that it opened a year after the announcement. Table 1.1 describes the sample I use in this paper. The sample is made of 9 major openings spread out across the US and the last two decades. In the main specification, I regroup Amazon's and Apple's openings in Los Angeles due to strong their geographical and timing proximity.

1.2.2 House Prices

House prices are obtained from Zillow's ZTRAX database[2], a transaction level data set covering the entire US since the mid 1990's. ZTRAX provides sale amounts as well as an extensive set of property characteristics that I use to control for housing quality. I compute the sale price per built square foot and use it as the main outcome in this analysis. I further use information on the exact location (latitude/longitude) of the property, the number of bedrooms and bathrooms,

[2]Zillow Group, 2020

the type (single family or condominium), I restrict the same to sales of single-family homes and condominiums with a minimum amount of $100,000 and a maximum of $3 million (2019 dollars). I also focus on square footage ranging between 300 and 6000. I restrict the sample to properties built afte 1850. Since my analysis looks at highly urban areas, these restrictions only remove outliers. I only keep property transactions that happen within 5 kilometer of one of the office I have in my sample. At this stage, the sample contains 527,778 transactions. It will be further reduced through the design of the identification strategy, explained in section 1.3.1. I experiment further restrictions, removing pre-20th century building and very large homes: these do not affect the results significantly.

1.2.3 Demographics

I obtain demographic data from the Census LEHD. These data contain yearly counts of individuals down the the Census block level, split by broad demographic groups: age, race, industry of work. These counts are performed at residence and workplace levels, allowing a distinction between industrial and residential compositions. This is the only public data of the Census offering such high geographic granularity along with a yearly frequency, which are both needed for my analysis. These data are used to investigate the underlying mechanisms behing the changes in house prices observed following the opening of a major technology office. The industry of work information allows us to measure the composition of the workforce at a fine level and detect possible broader changes initiated by the opening of the office. This, in turn, can inform us about the development of the local consumption amenity sector, or the presence of agglomeration forces attracting firms that benefit from the presence of the new office.

1.2.4 Final Dataset

I generate a novel dataset by merging office opening information with the aforementioned property price and demographic data. To do so, I first obtain each office latitude/longitude coordinates from their exact address using Google Maps. I construct a 5 kilometer buffer around each office and include every property transaction that happens within that buffer (using their coordinates provided

in ZTRAX) and within 5 years of the opening year (before and after). Using a geographical shapefile for 2010 Census block groups, I associate to every property transaction the code of the block group in which it happens. I can then merge the demographic information contained in the LEHD data to obtain the final dataset.

1.3 Empirical Analysis

1.3.1 Treatment and Control Design

Treatment and control groups are primarily designed based on the geographical distance from the office opening. The desirability of a short commute[3] the development of consumption amenities around large firms[4] and small-scale agglomeration economies[5] are factors that suggest that the housing market response should be strongest around the opening location and wane over short distances. While commute time is a more direct measure of the how close two places are from a human perspective, my analysis is run at a scale small enough (a few kilometers) that commute times and geographical distances are interchangeable.

In the baseline specification, the treatment group is defined as transactions happening in Census block groups lying within a 1 kilometer radius of the opening event. Even at such small scales, there can be important differences in housing markets. Two close neighborhoods may have different dynamics that need to be taken into account. I use a pre-trend matching algorithm to select most control areas most comparable with the treatment area. The procedure first isolates potential control block groups witin a certain radius of the treatment area. It then calculates a 3-year pre-opening trend in log property price per square foot for the treatment area and all potential controls. It finally picks the bottom 10% (in the main specification) controls in terms of distances in trends with the treatment. More details on this procedure are given in appendix 1.7.

Figure 1.1 shows the size of treatment and control groups for each office opening event.

[3]See e.g. Frenkel, Bendit, and Kaplan, 2013

[4]See e.g. Behrens, Boualam, et al., 2018

[5]See e.g. Buzard et al., 2017

1.3.2 Baseline Comparisons

By comparing areas with similar property price pre-trends, our approach already accounts for most relevant differences in treatment and control groups. For the sake of completeness, and because assignment to treatment is never fully random in such analyses, I report a comparison of relevant characteristics of treatment and control groups as of one year prior to the opening event. I divide these characteristics into housing quality ones, measured at property level, and demographic ones, measured at block-group level. The offices in my sample open in highly different housing markets: some are in very dense urban locations, while some are in the surburbs. The treatment and control groups are also not well-balanced in terms of sample sizes. Doing a single, overall comparison of property characteristics would thus bear the risk of comparing urban areas with suburban ones. To avoid this, I report property-level comparisons within each office area. The results are shown in figure 1.2. For most characteristics and offices, treatment and control groups are not statistically distinguishable. While some differences appear, they should not directly be seen as a threat to the analysis. Having slightly different housing stocks does not preclude treatment and control groups from being counterfactually on the same price trend.

Demographic characteristics being obtained from the Census LEHD at block group level, I have a fairly small sample of them. I thus report baseline comparisons in these characteristics for the overall sample, without distinguishing by office area. At this fine geographic scale, LEHD data reports only a few demographic characteristics. I report the proportion of high-skill workers, as defined by those who work at companies with a two-digit NAICS code falling in the top quartile in terms of college-educated workforce share (education data is not directly available for most years). I also report comparisons in terms of the percentage of residents who are white, under 30 years old and whose monthly wage exceed \$3,333[6]. As shown in table 1.2, treatment and control block groups do not significantly differ along any of these dimensions.

[6]Wages above this value are the highest wage bracket available in LEHD data.

1.3.3 Event Study Model

I harness a difference-in-differences strategy to evaluate the impact of a large office opening on the local housing market. The main specification is an event study model, where I estimate differences in treatment and control house prices for each year around the opening event. This specification uses the following structure:

$$y_{iot} = \sum_{s=-3}^{5} \beta_s \Big[T_i \times \mathbf{1}(\tau_{ot} = s) \Big] + \sum_{s=-3}^{5} \gamma_s \, \mathbf{1}(\tau_{ot} = s) + \alpha_{c(i)} + \lambda_t + X'_{iot}\delta + \epsilon_{iot}$$

where:

* y_{iot} is the log sale price per square foot of property i, located around office o and sold in year t
* $T_i = 1$ if i is treated, 0 otherwise
* τ_{ot} is the difference, in years, between t and the opening of office o
* γ_s are a set of leads and lags fixed effects
* $\alpha_{c(i)}$ are Census tract fixed effects, λ_t year fixed effects
* X_{iot} are property-level controls: number of bedrooms, number of bathrooms, categorical variable for the decade the property was built, property type (SF home or condo)

The coefficients of interest in this specification are the β_s. For each year s relative to the office opening year, β_s measures the average difference between house prices in the treatment and control groups. These estimates are obtained controlling for census tract fixed effects[7], absolute time fixed effects (controlling for general housing market trends) and fixed effects for time relative to opening (capturing any other time varying effect that would systematically happen around an office opening event and have a common impact on treatment and control groups). β_s also account for differences in the housing stock pertaining to the controls on property characteristics listed above. For $s \leq 0$, we expect β_s to be systematically close to and not significantly different from 0, which suggests that the treatment and control groups were on the same house price trend before the office opening

[7]Census tract fixed effects also indirectly control for which office opening event the property relates to, because tracts are more granular. We thus don't need additional office or city fixed effect.

event. For $s > 0$, β_s measures, in each year after the opening, the possible difference in house prices between treatment and control groups induced by the office opening event.

1.3.4 Endogeneity Discussion

Technology firms decide on office locations based on a number of factors. Among the most important of them are the corporate tax regime, the availability of office space and the availability of skilled workers[8]. We believe that the fine spatial granularity of this analysis controls for these potential confounders. Corporate taxes are usually a state-level issue, and do not change across neighborhoods of the same city; they thus do not affect my analysis. The local availability of skilled workers is most likely the same within a radius of a few kilometers because such a distance can easily be covered by walking or biking; there is no reason to see a sharp difference in commute times for workers if the office instead opened in a control area.

A more general endogeneity concern is the possibility that other events caused both the companies to move in, as well as house prices to go up relative to nearby neighborhoods. Any such event that would affect both treatment and control areas would be controlled for by the leads/lags terms in the event-study specification. Regarding events that would affect treatment and control groups differentially, it seems unlikely that they happened systematically, given the very diverse timings and locations of openings in the sample. Moreover, the fact that house price increases are very focal to the office opening location (see results in section 1.4) leaves little room for other events to have caused them. To test that in a more thorough way, we run a set of geographical placebo tests, by randomizing the location of new offices. The results are presented in section 1.5.3.

[8]For impact of taxes on firm location choice, see e.g. Giroud and Rauh, 2019. Regarding the availability of skilled labor, see e.g. Audretsch, Lehmann, and Warning, 2004. Regarding the availability of land and other common factors, see e.g. Buczkowska and Lapparent, 2014.

1.4 Results

1.4.1 Event Study Results

Figure 1.3 shows the results of the main specification (I report also detailed results in table 1.3). It shows the estimates for the main coefficients of interest β_s around the opening year. Pre-opening trends in house prices are almost idential in the treatment and control groups, suggesting good comparability. The differences clearly appear following the opening of an office. Within two years after opening, property prices in the treatment group gain about 10% relative to those in the control group. This short term effect slightly subsides over time, to reach around 7% five years after opening. This slight dip could reflect some sprawl, with the aura of the major office opening eventually reaching further outwards. But a large difference persists after several years. That suggests the ignition of endogenous mechanisms following the opening event and the fact these effects are hyper-local[9]. The potential role of different mechanisms is discussed in section 1.4.2.

We find much stronger property price effects that the existing literature working at a similar geographic scale. D. G. Pope and J. C. Pope (2015) find that Walmart and Target supermarket openings are associated with a 1–2% increase in prices within 0.5–1 mile (0.8–1.6 kilometers) over a 2 year period, while I find increases in the order of 10%. Glaeser, Kim, and Luca (2018) show that coffee shops openings are associated with a 0.5% increase in property prices within the same ZIP code over a 1-year period, while I find increases in the order of 7%. There are clearly major differences between the types of establishments that the literature has studied and those in my analysis. Coffee shop openings are most likely the consequence of an ongoing gentrification process and do not play a large part in that process individually. Supermarkets are more likely to play an important role, but not directly through the workforce they attract, being mostly composed of low-skill workers. Tech establishments tend to attract a large and highly skilled (hence highly paid) workforce, giving them a greater potential impact on the neighborhood they set foot in.

[9] A result that is in line with e.g. Behrens, Boualam, et al., 2018

1.4.2 Discussion

This section considers the role of two distinct mechanisms in generating a large and persistent house price difference between treatment and control groups. We succesively discuss the roles of agglomeration forces (the new firm attracting other firms because of production spillovers) and consumption amenities (the new firm and its workers pushing the development of certain local services).

At small geographic scales (e.g. a few kilometers here) the microeconomic reasons for agglomeration economies are dominated by labor market pooling and knowledge spillovers (Rosenthal and Strange, 2001). In other words, firms benefit from close proximity with other firms either because they use the same type of labor, which can reduce job search costs, or because they use similar knowledge and technology for their production, which can benefit from more frequent human interactions. Both these factors suggest that small-scale agglomeration forces happen between firms within the same industry. Major offices have an anecdotal potential to be the trigger of an agglomeration phenomenon. Since they belong to well-established companies, these offices are fairly autonomous: they may not require a local ecosystem of technology firms to function properly, which allows them to set foot in locations that do not yet experience tech agglomerations. On the other hand, they often serve as incubators or venture capital providers for smaller promising firms. That gives smaller firms a reason to come and locate near these big establishements to benefit from the opportunities they offer.

Agglomeration Forces. A way to assess the important of agglomeration forces in our case is thus to see whether the tech industry locally booms following the opening of a major office. To proxy for this boom, I use Census LEHD data and look at the relative evolution of the workforce between treatment and control areas in three key sectors: finance & insurance, management, and professional/scientific services[10]. Professional/scientific services notably include a number of activities associated to the technology sector, and is therefore an industry that we expect to boom

[10]Finance is defined by individuals working in the sector with NAICS code 52. Management by NAICS code 55. Professional/Scientific Services by NAICS code 54.

as a result of agglomeration forces around the major opening. For each block group b in my sample, I compute the total number of people working in each of these three sectors and residing in b. I aggregate these counts to entire treatment areas and normalize their share within the entire workforce to 0 at the year of office opening. I plot this index for each sector over time in figure 1.4. The pattern is clear: treatment areas experience a substantial gain in residents related to the tech industry after the opening, while other major high-skill sectors such as finance and management do not show a comparable boom. This points towards an agglomeration mechanism. While the opening office can account for part of this rise, it most likely does not account for the entirety of it, suggesting the entry of additional similar firms.

Consumption Amenities. I perform a similar analysis as for agglomeration forces above. This time, I focus on industries that provide services to local residents: arts and entertainment, and accomodation and food. Moreover, workers in these sectors are often low-income and likely not resident in the affluent or gentrifying areas they work in. Therefore, I used *workplace-level* data (WAC data from the LEHD) instead of the residence-level data used for the analysis of agglomeration mechanisms. An inflow of highly-paid tech workers following the opening of a new office could boost the demand for such services, whose development could further attract new wealthier residents — not necessarily working in the tech industry. The graph presented on figure 1.5 does not support that mechanism. The share of residents working in these service industries exhibits an overall flat, although erratic trend over the entire period, and the opening of a tech office does not alter that trend much. It is to be noted that the changes observed on the graph are also much smaller than those reflecting potential agglomeration forces on figure 1.4. The maximum observed change relative to the opening year is -0.15 percentage points, in the arts and entertainment workforce three years after the opening. Overall, this analysis suggests that the development of local consumption amenities is a much weaker potential explanation than agglomeration for the rise in house prices following the opening of a new large tech office.

1.5 Robustness

1.5.1 Sensitivity of Results to Design Parameters

I first assess the robustness of my results to the choice of parameters in the empirical design. Three parameters govern the definition of control and treatment groups: the treatment radius, the control eligibility radius, and the share of control blocks retained after pre-trend matching. To make the sensitivity analysis of results to these parameters more concise, I collapse the event-study model into a simple pre/post difference-in-differences. The specification used here is:

$$y_{iot} = \beta \left[T_i \, \mathbf{1}(\tau_{ot} \geq 0) \right] + \gamma \, T_i + \lambda_t \alpha_{c(i)} + X'_{iot} \delta + \epsilon_{iot}$$

where all elements are defined as in the main specification. The main coefficient of interest is now β. I report estimates from this specification in figure 1.6. Top panels show estimates for the coefficient γ of the model. γ reflects average pre-opening differences between treatment and control groups. Small, unsignificant estimates are obtained in nearly all parameters configurations, showing that pre-trends are parallel and thus validating the identifying assumption of the difference-in-differences approach. Bottom panels show estimate for the coefficient β of the model, reflecting post-opening differences between treatment and control. The estimates are large and significant across the board, showing that my results are not qualitatively sensitive to the empirical design parameters. The results are generally more significant than in the full event-study specification (see table 1.3) because of a gain in power.

1.5.2 Time Placebo Tests

In this section I assess the relevance of opening events for triggering house price spikes. To do so, I run placebo tests by artificially shifting office opening years forward and backward. For this experiment to be directly comparable to my main result, I re-run the entire procedure on the artificially modified data, including the pre-trend matching phase. Figure 1.7 shows the result of these placebo tests in bottom panels, along with the main results in the top panel for reference. There are no significant changes in house prices around forward and backward placebo years. This

reinforces the idea that office opening event are the true reason why we see house prices spike in the main result.

1.5.3 Geographical Placebo Tests

I further run geographical placebo test. Instead of using the actual office location, I redefine the center of the treatment area by a randomly drawn point within a 1 to 2-kilometer donut around the actual opening location. In other words, the office opening is now artificially moved to an outer ring around its actual location. I run 200 such random draws for each opening event and run the same regression analysis using these artificial locations. In the results, only 19% of the draws lead to significant estimates of the effect of office openings on house prices at the 5% confidence level. While a proportion around 5% would have been an ideal result to fully reject unobserved geographical factors, I take this result as a strong suggestion that these undesired effects are limited.

1.6 Conclusion

To be completed.

1.7 Appendix

Pre-trend Matching Algorithm

Consider an individual office opening event. The treatment area is defined as all block groups that intersect a buffer of radius d_T around the office location (defined by its latitude/longitude coordinates). $d_T = 1$ km in the main specification. I include all property transactions in a block group into the treatment, even when only part of that block group intersects with the treatment buffer. The reason to do so is twofold. First, block groups are very homogenous areas defined by the Census, and I do not want to create artificial boundary effects. Second, I use demographic data given at block-group level for balancing tests. To start the selection process of control block groups, I first consider as "eligible controls" the block groups that fall within a radius d_C of the treatment

area (excluding, of course, the treatment area itself). $d_C = 2$ km in the main specification. Again, an entire block group is considered as soon as part of it intersects the d_C-radius buffer.

For each eligible control block group b, I compute the pre-opening trend in its average log property price per square foot $P_{C,b,t}$. I do so too for the treatment area *as a whole*. I get:

$$\forall t \leq \text{opening year} : \begin{cases} \Delta P_{T,t} = P_{T,t} - P_{T,t-1} & \text{for the treatment} \\ \Delta P_{C,b,t} = P_{C,b,t} - P_{C,b,t-1} & \text{for each control block } b \end{cases}$$

Each eligible control block b's trend is then compared to the treatment trend through a mean squared distance (MSD):

$$\text{MSD}(b) = \sum_{t \leq \text{opening}} \left(\Delta P_{T,t} - \Delta P_{C,b,t} \right)^2$$

I select as a final control group the bottom λ_C share of eligible control blocks in terms of their MSD with the treatment. In the baseline specification, $\lambda_C = 10\%$. In the Google Venice example, this leads to the selection of the block groups colored in gray. This procedure is applied to every opening event separately.Figure 1.8 shows the final treatment and control areas for each opening event in my sample.

Figures

Figure 1.1: Sample composition: number of property transactions happening in the treatment and control groups, 3 years prior to 5 years after the opening, for each office in the sample.

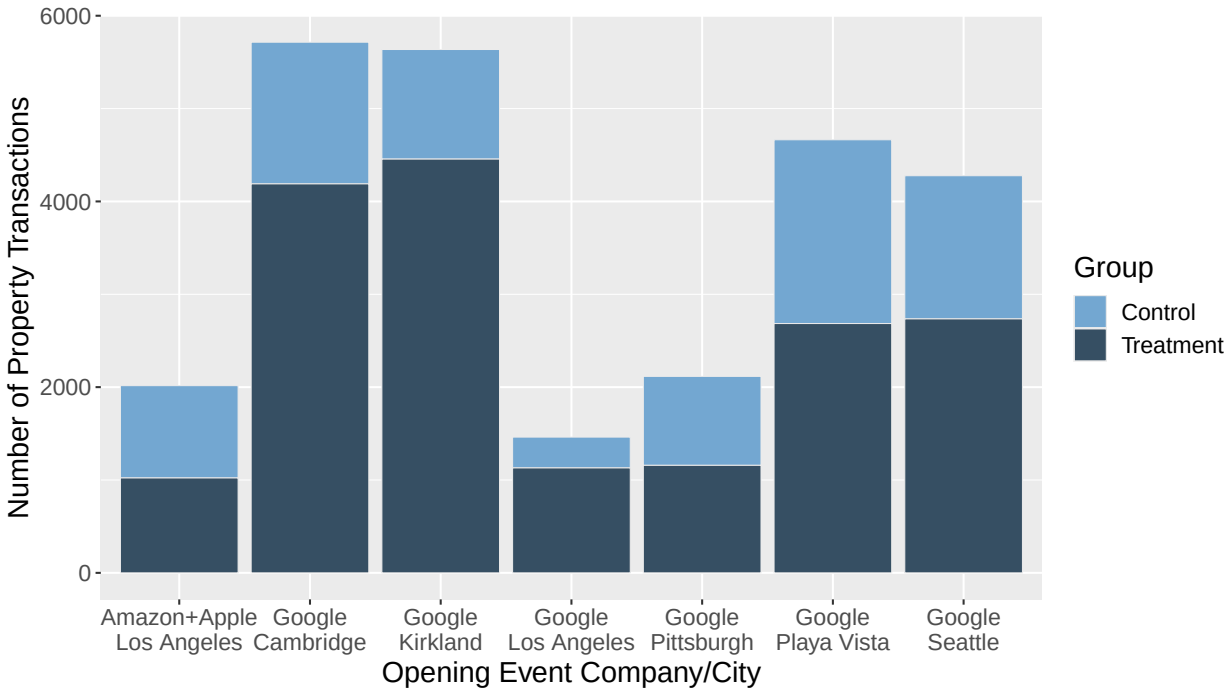


Figure 1.2: Baseline comparison of selected property characteristics between treatment and control groups, within each office area. Based on data one year prior to the office opening event. Dots represent mean estimates. Bars represent 95% confidence intervals.

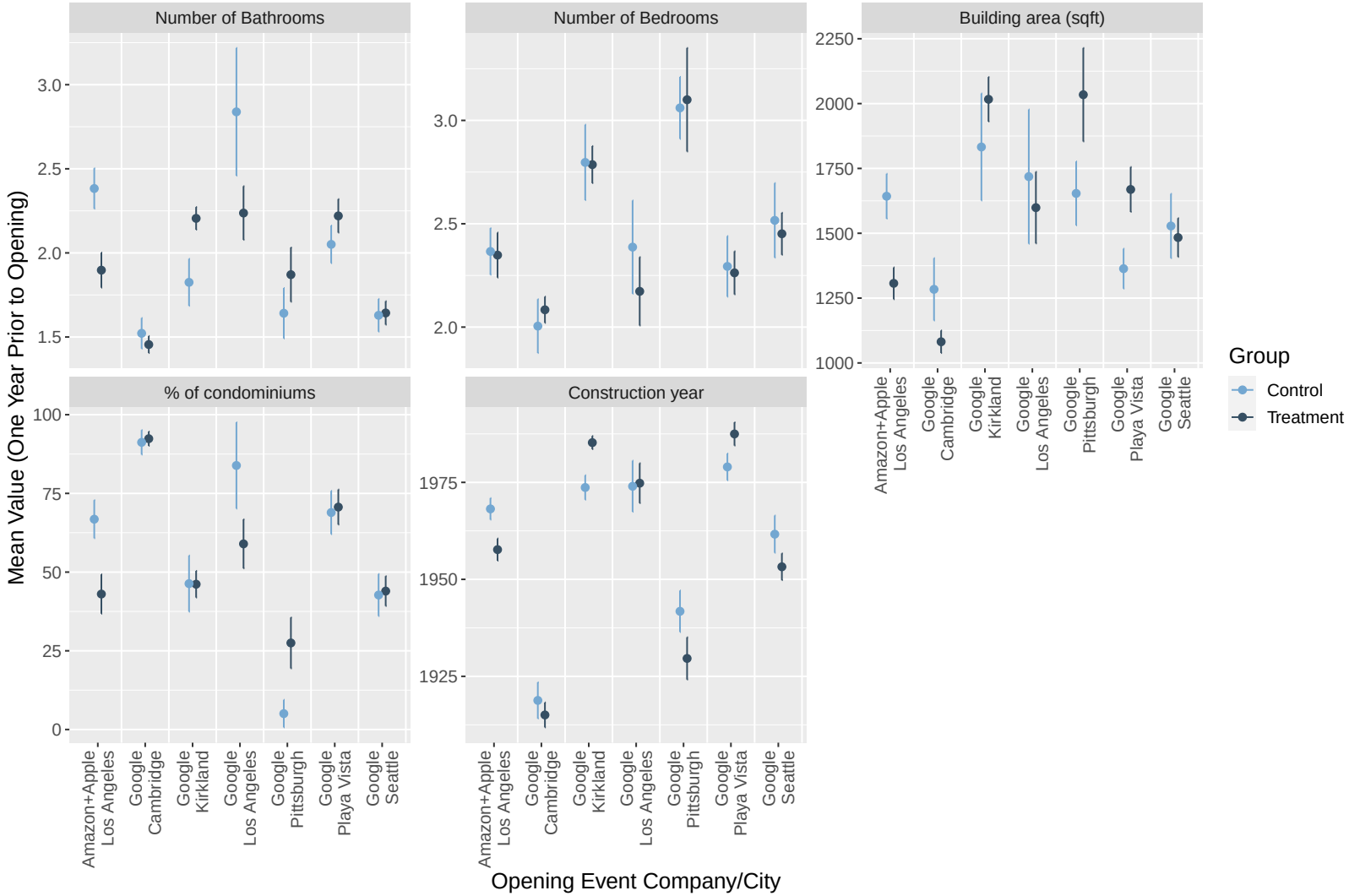


Figure 1.3: Event study results: estimates for β_s, $s \in [-3, 5]$. $N = 25{,}844$. $R^2 = .67$. Treatment group is defined as properties sold within 1 km of the office locations. Control group is defined as properties within a 2 km buffer of treatment area, belonging to block groups in the top 20% of pretrend matching. Property controls include number of beds and baths, an indicator for the decade of construction, a condo/single-family dummy. Error bars denote 95% confidence intervals. Robust standard errors clustered at block-group level.

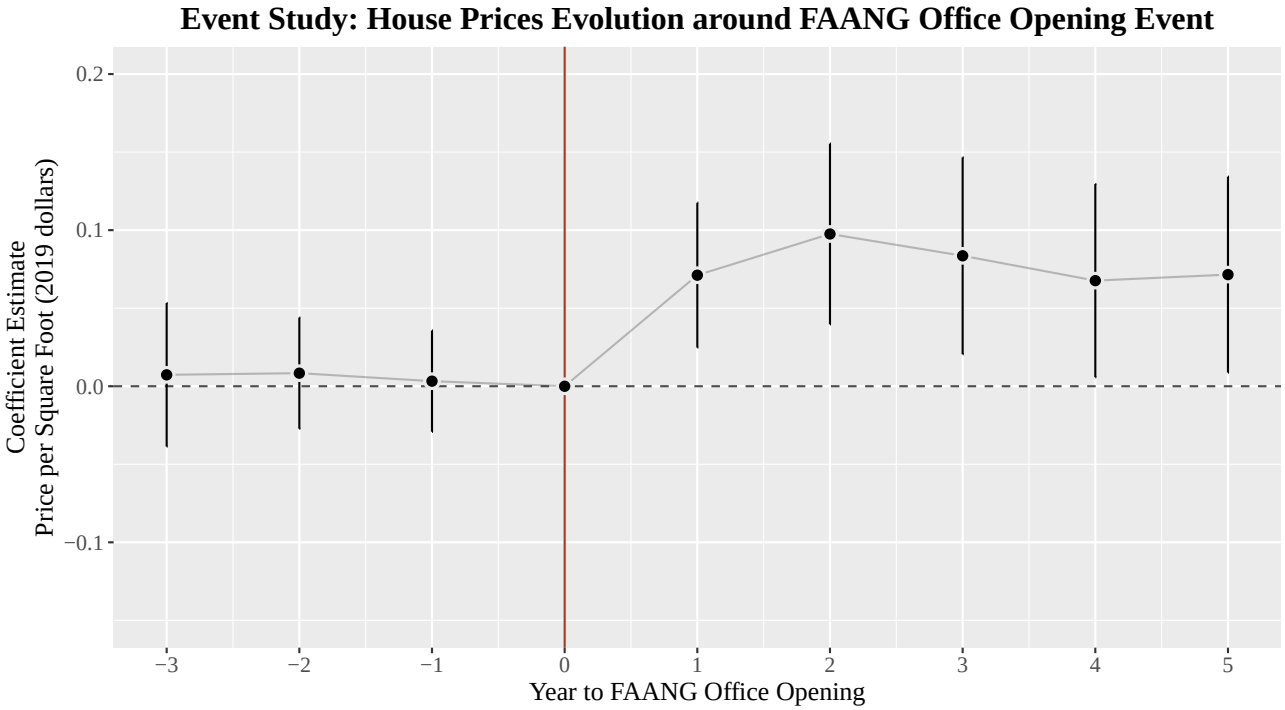


Figure 1.4: Evolution of the workforce residing in treatment areas in three sectors (finance, management and professional/scientific services). This plots the percentage shares of specific sectors' workforce within the entire workforce. Shares are normalized to their opening-year level, so the graph shows percentage points changes relative to that year.

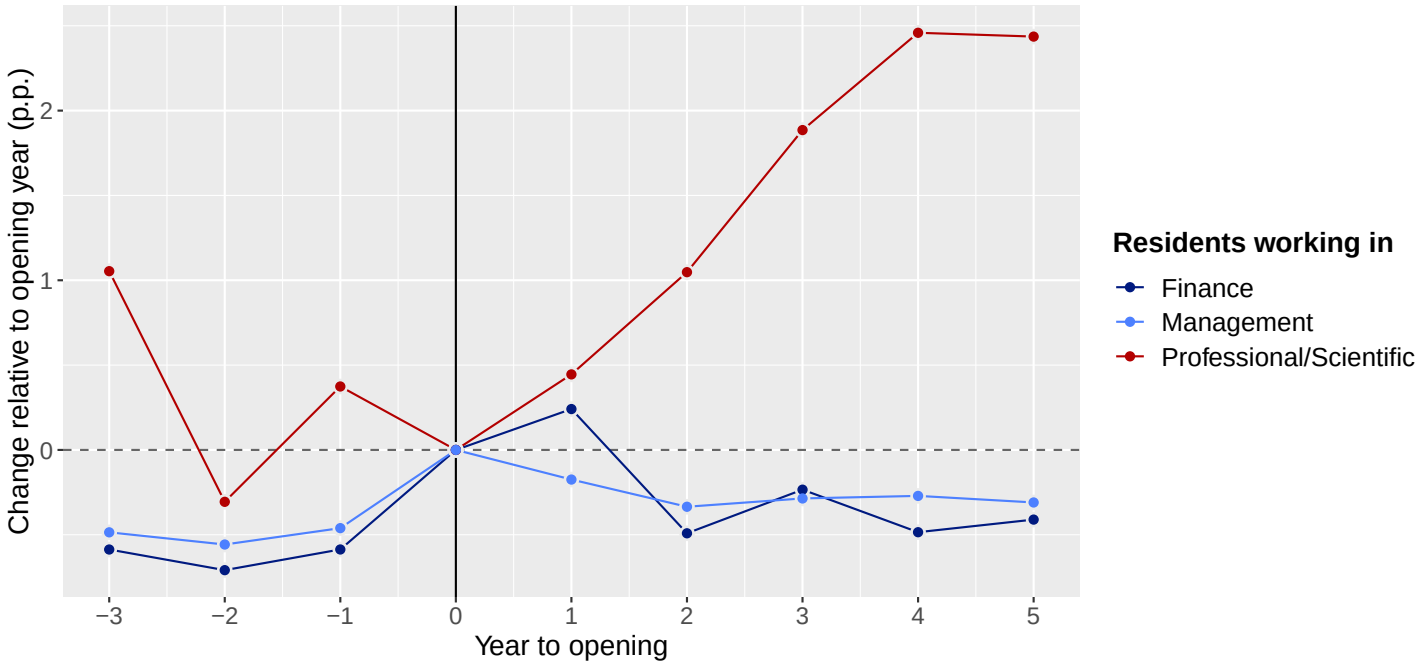


Figure 1.5: Evolution of the workforce working in treatment areas in two sectors (arts and entertainment, and accomodation and food). This plots the percentage shares of specific sectors' workforce within the entire workforce. Shares are normalized to their opening-year level, so the graph shows percentage points changes relative to that year.

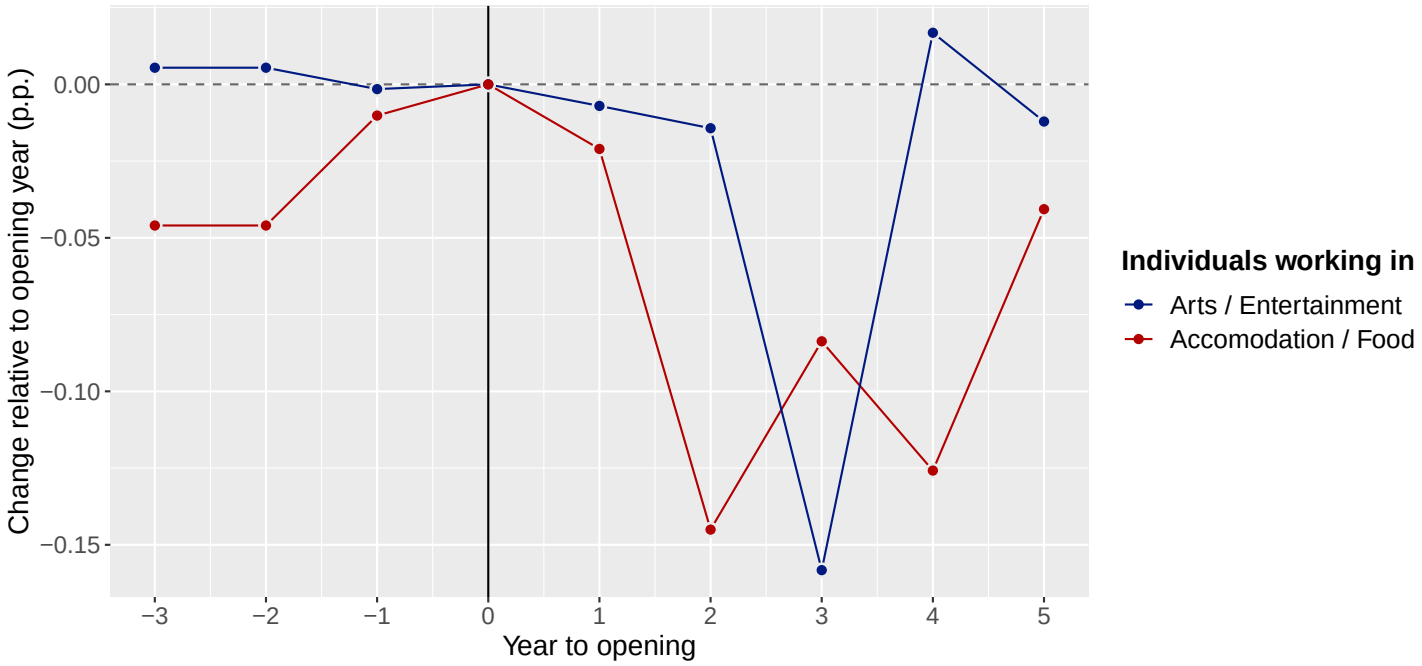


Figure 1.6: Robustness tests on 27 empirical design configurations. Varying treatment area radius (0.5, 1 and 1.5 km), control buffer size (1, 2 and 3km) and percentage of control block-groups kept in pretrend matching (10, 20 and 30%). The top panels show the coefficient on the treatment dummy, which measures the average log price difference between treatment and control, prior to office opening. The bottom panels show the coefficient on the interaction between treatment dummy and post dummy, measuring the additional difference in prices following the opening.

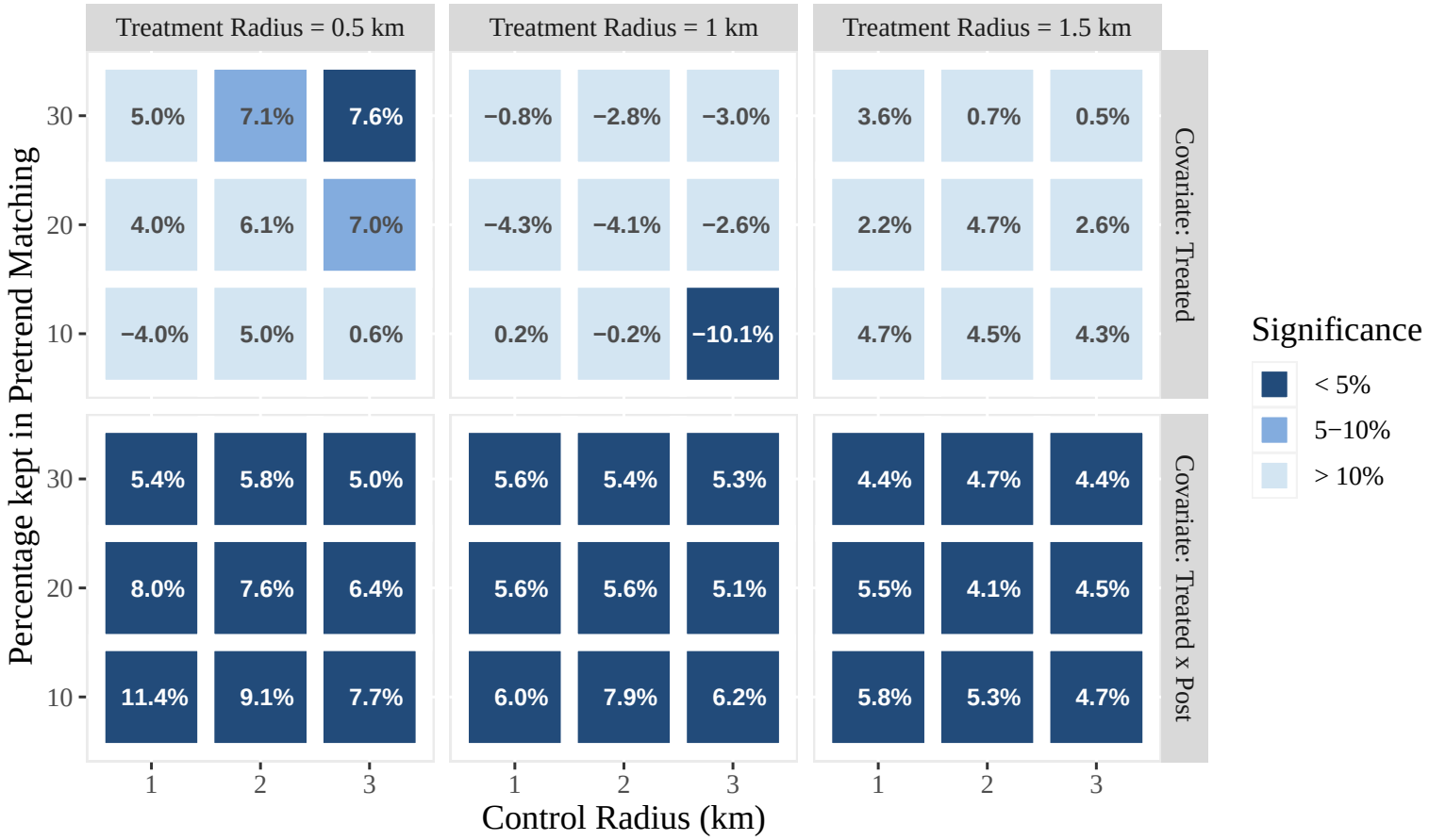


Figure 1.7: Placebo tests. Entire experiment, including pre-trend matching, is re-run using placebo years for office openings. The top panel shows the main results for reference. The bottom left panel shifts opening years 3 years forward. The bottom right panel shifts them 3 years backward.

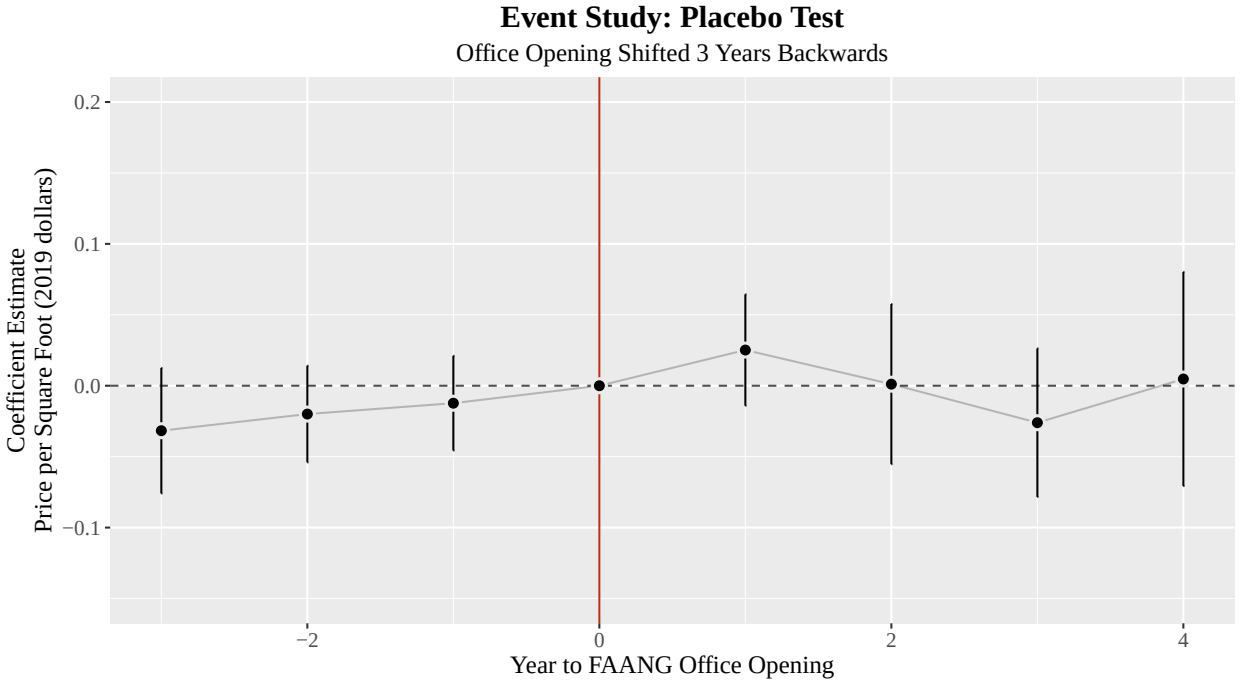
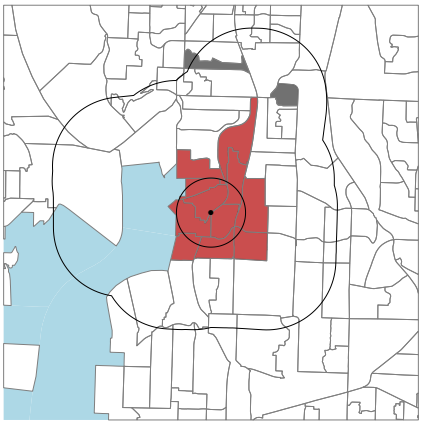


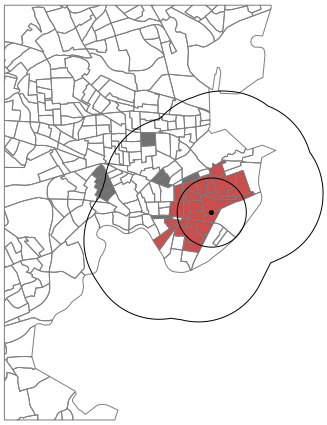
Figure 1.8: Treatment and control areas selected in the main specification. The dots represent new offices. The smaller circles represent treatment radius (1 km here). The red block groups are those that are considered treated. The larger black boundaries represent control eligibility areas (2 km around the treatment area here). The gray block groups are controls selected through pre-trend macthing procedure (top 10% here).

(a) Google – Kirkland



(b) Amazon and Apple – Los Angeles

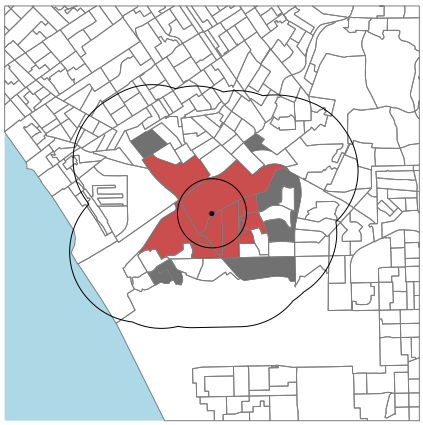
(c) Google – Cambridge



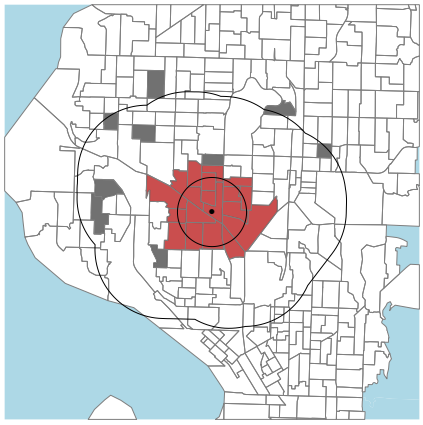
(d) Google – Los Angeles

(e) Google – Pittsburgh

(f) Google – Playa Vista



(g) Google – Seattle



Tables

Table 1.1: Sample of tech office openings.

Company	City	State	Opening Year
Google	Kirkland	Washington	2004
Google	Seattle	Washington	2006
Google	Cambridge	Massachussetts	2008
Google	Pittsburgh	Pennsylvania	2010
Google	Playa Vista	California	2011
Google	Los Angeles	California	2011
Amazon	Los Angeles	California	2017
Apple	Los Angeles	California	(ann.) 2018

Table 1.2: Baseline comparison of selected demographic characteristics between treatment and control groups. The treatment and control columns represent proportion estimates in each group. Difference = treatment - control. The p-value is from a Student t-test on the difference, with H_0: difference = 0. Characteristics measured at block group level. Based on data one year prior to the office opening event. High-skill industries are defined as 2-digit NAICS codes in the top 25% in terms of share of college-educated workers.

Demographic Characteristic	Treatment	Control	Difference	P-value	N. Obs.
High Skill (%)	37.37	35.1	2.27	0.204	181
White (%)	76.49	76.06	0.43	0.887	109
Under 30 y.o. (%)	27.76	26.51	1.25	0.503	181
Monthly Wage > $3,333 (%)	47.54	45.81	1.72	0.367	181

Table 1.3: Estimates for the coefficients of interest β_s in the event-study specification. The baseline specification is Model 2, with a treatment radius of 1 km. I also report results under two alternative treatment radii in Model 1 and 3. Robust standard errors clustered at the block group level.

	Model 1	Model 2	Model 3
Treatment Buffer:	0.5 km	1 km	1.5 km
Treat × (YtO = −3)	−0.06	0.01	0.01
	(0.06)	(0.02)	(0.02)
Treat × (YtO = −2)	−0.02	0.01	0.00
	(0.03)	(0.02)	(0.02)
Treat × (YtO = −1)	−0.03	0.00	−0.01
	(0.03)	(0.02)	(0.01)
Treat × (YtO = 1)	0.06	0.07***	0.06***
	(0.05)	(0.02)	(0.02)
Treat × (YtO = 2)	0.08*	0.10***	0.07***
	(0.04)	(0.03)	(0.03)
Treat × (YtO = 3)	0.04	0.08***	0.04
	(0.04)	(0.03)	(0.03)
Treat × (YtO = 4)	0.08**	0.07**	0.04
	(0.03)	(0.03)	(0.03)
Treat × (YtO = 5)	0.03	0.07**	0.05*
	(0.05)	(0.03)	(0.03)
Property Controls	✓	✓	✓
Census Tract F.E.	✓	✓	✓
Adj. R^2	0.63	0.67	0.67
Num. obs.	14,934	25,844	39,266

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Bibliography

Audretsch, David B, Erik E Lehmann, and Susanne Warning (2004). *University Spillovers and New Firm Location*. en. Tech. rep. 0204. Papers on Entrepreneurship, Growth and Public Policy.

Behrens, Kristian, Brahim Boualam, et al. (2018). *Gentrification and pioneer businesses*. Tech. rep.

Behrens, Kristian, Gilles Duranton, and Frédéric Robert-Nicoud (2014). "Productive Cities: Sorting, Selection, and Agglomeration". In: *J. Polit. Econ.* 122.3, pp. 507–553.

Berkes, Enrico and Ruben Gaetani (2018). *Income Segregation and Rise of the Knowledge Economy*. Tech. rep. 213.

Buczkowska, Sabina and Matthieu de Lapparent (2014). "Location choices of newly created establishments: Spatial patterns at the aggregate level". In: *Reg. Sci. Urban Econ.* 48, pp. 68–81.

Buzard, Kristy et al. (2017). "Localized Knowledge Spillovers: Evidence from the Spatial Clustering of R&D Labs and Patent Citations".

Couture, Victor, Cecile Gaubert, et al. (2019). "Income Growth and the Distributional Effects of Urban Spatial Sorting".

Couture, Victor and Jessie Handbury (2017). "Urban Revival in America, 2000 to 2010".

Dave, Paresh and Roger Vincent (2017). "Snapchat has changed Venice, and the neighborhood isn't changing back". In: *Los Angeles Times*.

Davis, Donald R and Jonathan I Dingel (2019). "A Spatial Knowledge Economy". In: *Am. Econ. Rev.* 109.1, pp. 153–170.

Diamond, Rebecca (2016). "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000". In: *Am. Econ. Rev.* 106.3, pp. 479–524.

Ellison, Glenn, Edward L Glaeser, and William R Kerr (2010). "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns". In: *Am. Econ. Rev.* 100.3, pp. 1195–1213.

Frenkel, Amnon, Edward Bendit, and Sigal Kaplan (2013). "Residential location choice of knowledge-workers: The role of amenities, workplace and lifestyle". In: *Cities* 35, pp. 33–41.

Ganong, Peter and Daniel Shoag (2017). "Why has regional income convergence in the U.S. declined?" In: *J. Urban Econ.* 102, pp. 76–90.

Gaubert, Cecile (2018). "Firm Sorting and Agglomeration". In: *Am. Econ. Rev.* 108.11, pp. 3117–3153.

Giroud, Xavier and Joshua Rauh (2019). "State Taxation and the Reallocation of Business Activity: Evidence from Establishment-Level Data". In: *J. Polit. Econ.* 127.3, pp. 1262–1316.

Glaeser, Edward L, Hyunjin Kim, and Michael Luca (2018). "Measuring Gentrification: Using Yelp Data to Quantify Neighborhood Change".

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti (2010). "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings". In: *J. Polit. Econ.* 118.3, pp. 536–598.

Hwang, Jackelyn and Jeffrey Lin (2016). "What Have We Learned About the Causes of Recent Gentrification?" In: *Cityscape* 18.3, pp. 9–26.

Meltzer, Rachel (2016). "Gentrification and Small Business: Threat or Opportunity?" In: *Cityscape*.

Meltzer, Rachel and Pooya Ghorbani (2017). "Does gentrification increase employment opportunities in low-income neighborhoods?" In: *Reg. Sci. Urban Econ.* 66, pp. 52–73.

Pope, Devin G and Jaren C Pope (2015). "When Walmart comes to town: Always low housing prices? Always?" In: *J. Urban Econ.* 87, pp. 1–13.

Rosenberg, Mike (2018). "Will Amazon's HQ2 sink Seattle's housing market?" In: *Seattle Times*.

Rosenthal, Stuart S and William C Strange (2001). “The Determinants of Agglomeration”. In: *J. Urban Econ.* 50.2, pp. 191–229.

Schuetz, Jenny (2014). “Do art galleries stimulate redevelopment?” In: *J. Urban Econ.* 83, pp. 59–72.

Zillow Group (2020). *ZTRAX*. URL: https://www.zillow.com/research/ztrax/.

Chapter 2

Labor Market Polarization and the Growth of Service Employment: Routinization or Consumption Spillovers?

2.1 Introduction

Recent decades have seen an important reallocation of noncollege employment towards service occupations in the US (Autor and Dorn, 2013). Because these occupations pay among the lowest wages, this reallocation has been an important contributing factor to wage inequality (Autor, 2019). Two main theories have been proposed to explain this phenomenon: routinization and consumption spillovers. Routinization, initially developed by Autor, Levy, and Murnane (2003), refers to the replacement of automatable ("routine") workers by machines and computers, leading to a reallocation of labor towards other occupations, notably services. The consumption spillovers theory, first proposed by Manning (2004), refers to the labor demand increase for service occupations stemming from an income effect: real wages have increased (notably those of college-educated workers) and now richer households disproportionately demand services (in other words, the demand for services is not homothetic in income).

Do both theories really play a role in the growth of service occupations, and what are their relative importance in explaining this phenomenon? Previous studies have used regional variations in the determinants of these mechanisms to study each theory separately (see e.g. Autor and Dorn (2013) and Mazzolari and Ragusa (2013) for the two main contributions on each theory). We show that

these determinants are highly spatially correlated, raising confounding concerns when studying them in isolation. In turn, we propose reduced-form and structural approaches that alleviate these confounding concerns and allow us to separately assess the quantitative relevance of routinization and consumption spillovers. Decomposing the contributions of each theory has significant policy implications. If routinization is the main culprit, efforts should be made to improve the prospects of affected workers. For instance, investment in training of displaced routine workers could give them a better chance of moving up the skill distribution, instead of down towards less-paid service occupations[1]. If instead, consumption spillovers matter more, the focus should be on improving the conditions of service workers. For instance, increases in minimum wages would have substantial effects, since many of these workers are paid close to the floor[2].

This paper makes three distinct contributions. First, we provide reduced-form evidence that both the routinization and consumption spillovers channels operate at the same time. Using micro data from the Census and American Community Survey over the period 1980-2005, and following the strategy of Autor and Dorn (2013) and Mazzolari and Ragusa (2013), we exploit regional variations in the initial share of routine employment (for routinization) and in the growth of college wages (for consumption spillovers) and associate them with changes in the share of service occupations. We find strong evidence supporting both theories. In our main specification, the two mechanisms have effects of similar magnitudes: a one-standard-deviation increase in the initial share of routine employment is associated with a 0.56% increase in the noncollege employment share of service occupations, while an equivalent increase in the growth rate of college wages is associated with 0.42% change. These findings are qualitatively robust to the addition of a number of labor market controls, although the magnitude and significance of the consumption spillover channel decreases substantially. More importantly, accounting for both theories in the same regression model reveals a confounding concern. The coefficients associated with each theory significantly decrease (by 13% for routinization and 38% for consumption spillovers) relative to a model that includes only

[1]Investment in re-training workers is one of the main ideas to face the expected labor displacement effect from future technologies (Casey, 2020). Nedelkoska and Quintini (2018) finds that workers who are the most vulnerable to automation are three times less likely to benefit from on-the-job training than other workers.

[2]For instance, restaurant workers have always been considered the typical minimum-wage earner, from the seminal works of Card and Krueger (1994) to recent studies on the topic by Jardim et al. (forthcoming)

either one of the two theories. By studying these mechanisms in isolation, previous literature has likely overlooked these confounding issues and produced biased estimates.

While our regression evidence suggest that both routinization and consumption spillovers matter for the growth of noncollege service employment, a unified structural model is needed to correctly assess their relative importance. The second contribution of this paper is therefore theoretical. Building on Autor and Dorn (2013), who develop a parsimonious spatial framework to study routinization, we introduce a consumption spillover channel by allowing preferences to be non-homothetic. We use this model to simulate the effect of an exogenous decrease in the price of capital. This leads to two effect. On the one hand, firms adopt more capital that substitutes for routine labor. On the other hand, capital-skill complementarity leads to increases in the wages of college workers, whose demand for services disproportionately increases because of non-homotheticity. Our model is able to reproduce spatial disparities in service employment growth across US commuting zones through both the routinization and consumption spillover channels. This two channels are governed by separated parameters, making the model well-suited to study their relative importance in service employment growth. We provide a characterization for the solution of the model[3], as well as identification conditions for its key parameters.

The third contribution is to estimate the model and run simulations to quantitatively assess the importance of routinization and consumption spillovers. More specifically, we estimate two models: one where preference homotheticity is imposed, and one where we estimate a parameter governing non-homotheticity. The former model only reflects the routinization channel, while the latter combines it with potential consumption spillovers. We then compare the implications of these two models in terms of service employment. Our estimation procedure relies on 1980 commuting-zone-level data on employment and wages from the US Census Bureau, as well as data on consumption of goods and services from the Consumer Expenditure Survey (administered by the Bureau of Labor Statistics.). The former data source allows us to estimate parameters that are key to the routinization mechanism, while the latter is relevant for parameters related to the consumption spillovers channel. Our main estimate for the non-homotheticity parameter is large; it implies that the least-paid workers

[3]The model does not admit a closed-form solution, but we provide a set of non-linear equations characterizing its solution and implement an algorithm to solve the model numerically.

spend 95% of their income on goods, where the most-paid spend only about 20%. Using data from the Penn World Tables on the real price of capital over 1980-2005, we simulate a fall in that price in each model. We find that with non-homotheticity, regional disparities in service employment growth are nearly 57% larger than under the homotheticity assumption. Finally, decomposing the effect of non-homotheticity on service employment into the contributions of different worker types, we find that college-educated workers' demand for services is its main driver. Overall, the evidence from our reduced-form and structural analyses points to the importance of consumption spillovers, where routinization is too often considered as the sole explanation.

Related Literature. This paper contributes to the literature that studies the determinants of labor market polarization. Initially documented by Acemoglu (1999), the phenomenon has since received widespread attention. The routinization explanation remains most prevalent explanation in the literature. Autor, Levy, and Murnane (2003), Goos and Manning (2007), and Autor and Dorn (2013) provide support for this explanation, by exploiting regional variation in the presence of routine jobs in the context of the US, the UK, or internationally. They focus on automation shocks that target specific occupations. More recently, Barany and Siegel (2020) note that occupation-specific and sector-specific shocks can be difficult to separate and can easily be mistaken for one another. They propose a framework to test whether it is sector-specific, occupation-specific or sector-occupation-specific technological shocks that drove the occupational reallocation of employment observed in the recent decades. They find that occupation-specific shocks play a central role, consistent with the routinization hypothesis of the previous papers. The main alternative explanation for polarization relies on consumption spillovers and was introduced by Manning (2004). Mazzolari and Ragusa (2013) provide evidence that growth of wages in the top decile are associated with growth in specific service occupations that substitute for home production. Comin, Mestieri, and Danieli (2020) note that expenditure elasticities and the employment share of low-skill occupations are positively correlated across sectors: as income grows, demand for low-skill sectors disproportionately rises and this causes polarization. Boppart (2014) and Comin, Lashkari, and Mestieri (2021) similarly note the importance of income effects on labor reallocation patterns across sectors and occupations. More generally, the literature has considered a range of spillover effects across skill groups. In a

seminal contribution, Moretti (2004) shows that in cities with a higher share of college-educated workers, less-educated ones earn higher wages. More recently, Liu and Yang (2021) show that the presence of college-educated workers increases the local share of noncollege workers in services, but not in manufacturing. They interpret this finding as evidence for consumption spillovers, as opposed to production spillovers (see Lee (2019) for a constrasting view). Other notable explanations that have been proposed rely on shocks that affect certain sectors and occupations through trade (Autor, Dorn, and Hanson, 2013; Autor, Dorn, and Hanson, 2015) and offshoring (Goos, Manning, and Salomons, 2014). These studies largely focus on a single explanation at a time and give little attention to the possibility that multiple explanations may coexist and be intertwined, raising omitted variable bias concerns. This paper takes the spatial correlation of both theories more seriously and provides a structural framework allowing to evaluate their respective contributions.

The paper is organized as follows. Section II describes the data. Section III provides descriptive and reduced-form evidence that both the routinization and consumption spillover channels operate in the data. Section IV presents the model. Section V provides identification conditions for key parameters of the model, shows estimation results and simulations under estimated parameters. Section VI discusses the results. Section VII concludes.

2.2 Data

We use four primary sources of data. Data from the Census and the American Community Survey (ACS) is used for employment and wages, as well as the demographic characteristics of workers. Data from the Consumer Expenditures Survey (CEX) is used to analyze consumption patterns of goods and services. Data from the Bureau of Economic Analysis (BEA) is used to build a measure of local capital income. Finally, data on the price of capital is recovered from the Penn World Tables for model simulations.

2.2.1 Employment and Wages

Census/ACS employment and wage data from the IPUMS USA database (Ruggles et al., 2020) is collected for the years 1980, 1990, 2000 and 2005. The Census samples for 1980, 1990, and 2000 include 5 percent of the US population and the ACS sample includes 1 percent of the population. Theses samples are restricted to individuals aged 18-64 at the time of survey, who were not institutionalized, in the nonfarm sector, and who worked positive hours and earned positive labor income in the year prior to survey. Labor is defined as total hours worked, and computed as the product of three components: the individual's sample weight, the number of weeks worked and the usual weekly work hours. For individual with missing weeks worked or work hours, we infer that information from the mean of individuals in the same year and educational-occupational cell. Hourly wages are calculated by dividing annual wages by the product of weeks worked and weekly work hours, excluding self-employed workers. Wages are adjusted for inflation using the 2005 Consumer Price Index.

Our analysis requires a time-consistent definition of labor markets. In line with the Economic Geography literature, we rely on the commuting zone (CZ) geography developed by Tolbert and Sizer (1996). Commuting zones are a complete partition of the mainland US, where each unit is a county or a group of counties. These units are defined such that residence-workplace commuting patterns are strong within units and weak across units. They thus coincide well with the concept of local labor market. For each year of Census/ACS data, it is possible to crosswalk the finest available geography to 1990 commuting zones with small approximations. For 1980 data, we use a crosswalk from county groups to commuting zones. For other years, we use a crosswalk from Public Use Microdata Areas (PUMAs) to commuting zones. All crosswalks are obtained from the data page of the website of *David Dorn* (2021).

2.2.2 Occupational Classifications

Studying the importance of the routinization and consumption spillover hypotheses to explain growth in service employment requires defining a specific occupational classification and applying it to Census/ACS data. In particular, it is crucial to define service workers, as well as routine

workers that are most impacted by technological change. I thus categorize occupations into four categories, following Autor and Dorn (2013). Workers with a high-school degree or less are categorized into three occupational groups: service, routine, and other. Workers with at least some college education make up the fourth category, which we will refer to as college workers. This section presents more detail on these classifications.

Service occupations are defined based on the broadest "service" classification available directly in Census data. Table 2.1 shows the ten detailed occupations that make the biggest employment shares within the service category in 1980. Together, they make more than 80% of noncollege service employment in both 1980 and 2005, which are the beginning and end of our analysis period. They typicall require low skill levels and they all consist in caring for and assisting others. For the most part, they provide households with services that they could otherwise produce at home (e.g. eating out versus cooking at home).

Routine occupations are defined following Autor, Levy, and Murnane (2003)[4]. Their distinctive characteristic is to rely heavily on repetitive or codifiable tasks, that have the potential of being executed by a computer. As such, workers in routine occupations are more vulnerable to automation than other workers. Table 2.2 shows the ten detailed occupations making the largest employment shares within the routine category. Routine occupations largely fall into two subgroups: clerical and production/operator jobs. Clerical occupations, like secretaries and bookkeepers, have been impacted by advances in computer technology that allowed processing more information with fewer workers. Production and operator jobs typically involve numerous repetitive tasks that can be performed by machines.

Noncollege workers that are not employed in either service or routine occupations are classified in the "other noncollege" category. Workers with at least some college education are regrouped in the fourth and last occupational category.

[4]They use measures of task contents from the Dictionary of Occupational Titles (United States Employment Service, 1977) for each detailed occupations present in the Census classification and build a simple score aimed at assessing the "routine" aspect of an occupation. In their main specification, occupations are classified as routine if they fall in the top third of the routine score. More information is provided in the appendix.

2.2.3 Consumer Expenditures Data

The consumption spillover hypothesis builds on the premise that preferences for goods and services exhibit non-homotheticity: richer households disproportionately demand services. To test this premise and quantitatively measure the strength of the spillover channel, we need information on the consumption patterns of households with different income levels. The Diary part of the Consumer Expenditures Survey (henceforth CEX) provides household-level data on consumption of goods and services, as well as information on wages of household members, demographic data and residence location at the MSA level. We use the vintages of the CEX from 2013 to 2019 because it includes important geographical information used to control for local prices, as well as occupational information to map these data to the Census data. Since this period happens after the period of our analysis (1980-2005) the implicit assumption is that consumption pattern differences across income levels haven't changed significantly over time (see e.g. Boppart (2014) for evidence on that regularity). For consistency with the sample restriction on Census/ACS data, we focus on households headed by an individual aged 18-64, who worked for positive income in the past year. We first describe our classification of expenditures into goods and services, explain how we measures differences in consumption patterns across income, and finally explain how we map these data to our occupational classification in the Census.

We identify as "services" expenditure items that fall in the "home substitute services" defined by Mazzolari and Ragusa (2013). This classification was proposed to highlight service expenditures that households purchase as a replacement of potential home production. It is consistent with the consumption spillover premise that richer workers purchase more services to substitute for home production, because they have a higher opportunity cost of time. Table 2.3 describes the detailed categories of expenditures included in services. This classification of expenditures is remarkably consistent with the service occupations classification proposed in section 2.2.2. For instance, housekeeping and gardening maps naturally into janitor and gardener occupations, while expenditures related for food away from home maps into waiters and cooks. One can therefore expect that a change in demand for items belonging to service expenditures will translate into changes in labor for service occupations defined above.

Using our expenditure classification, CEX data allow us to compute household-level shares of expenditures in services by simply taking the ratio of service expenditures to total expenditures. We can then relate it to household economic, demographic and geographical information. The model that we present in this paper does not feature cross-sectional variations in income at the individual level, but only at the *worker type* level. Worker types in the model map into the service, routine and college workers defined in section 2.2.2. In order to use expenditure shares that are consistent between the model and the data, we need to classify workers the same way in the CEX. Two of these worker types are directly identifiable in the CEX data. Service occupations are categorized in the CEX exactly the same way as they are in the Census. College workers can also be identified based on education information, as in Census data. The level of detail in occupational classes in the CEX prevents us from defining routine workers in these data. However, we show in section 2.5.3 that we only need two worker types with different wage levels to identify the parameter governing preference non-homotheticity in our model.

2.2.4 Capital Income Data

To estimate the model presented in this paper, we need a measure of capital income by commuting zone. This measure is important to quantify how reliant regions are on capital in production in 1980. This is directly tied to the main goal of this paper, because initial capital intensity determines both (i) how a fall in the price of capital will impact substitution between capital and routine labor and (ii) how it will impact college workers through complementarity with capital. Capital income is not available in Census data, so we rely on data from the BEA.

The BEA provides data on investment in fixed assets by all private companies in 1980. While these data are not available at regional level, they are available at industry level (2-digit NAICS code). Combined with information on industrial composition at the commuting zone level, one can thus build a local measure of investment. Using BEA data, we compute aggregate investment amounts for each of the 22 2-digit NAICS codes covering the private sector, and divide them by the aggregate number of workers in each industry (obtained from Census data):

$$\text{CPW}_{i,1980} = \frac{\text{total investment }_{i,1980}}{\text{number of workers }_{i,1980}} \qquad (i \text{ is an industry})$$

We thus get a measure $\text{CPW}_{i,1980}$ of "capital per worker" by industry i in 1980. We then compute the number of workers in each industry and commuting zone from Census data. Multiplying these local number of workers by $\text{CPW}_{i,1980}$ and summing across industries gives us a measure of investment at the commuting zone level c:

$$\text{total investment}_{c,1980} = \sum_i \text{number of workers}_{c,i,1980} \times \text{CPW}_{i,1980}$$

The accuracy of this measures relies on the assumption that the variation of capital intensity across regions is mainly due to industrial composition differences.

2.2.5 Price of Capital

We use data on the price of capital to perform simulations with our structural model. These data come from the Penn World Tables (henceforth PWT; see Heston, Summers, and Aten (2012)). The data provide us with a measure of the price of capital p_k, as well as a measure of the price of consumption goods p_g. To match with the real price of capital in the model, we use the ratio p_k/p_g.

2.3 Empirical Evidence

This section presents reduced-form evidence in support of both the routinization and consumption spillovers theories. It closely follows the empirical strategy of Autor and Dorn (2013) and Mazzolari and Ragusa (2013) but gives more attention to the interactions between the theories they develop. The identification of these mechanisms hinges on geographical variations in their determinants across commuting zones in the US. For routinization, we exploit regional variations in the noncollege employment share of routine occupations. For consumption spillovers, we exploit variations in the growth rate of wages of college-educated workers. While the data provide support for both theories, we also find that their determinants are spatially correlated, raising confounding concerns. We start addressing these concerns by confronting both theories in a common regression framework.

2.3.1 Noncollege Service Employment Has Grown Faster than Other Occupations

Among college without a college degree, the share of employment in service occupations has dramatically increased over the 1980-2005 period. Table 2.4 shows that it grew from 13.1% to 19.1%, making it the occupation group with the highest employment growth. Service occupations are also paid the lowest wages in 1980, with an national average of $9.6 per hour (in 2005 dollars). They maintain this lowest ranking in 2005, despite being among the three occupation groups with positive wage growth over the period. Employment growth in service occupations thus contributes to a substantial reallocation of labor towards the least-paying jobs over 1980-2005.

2.3.2 The Adoption of Computer Technology is Correlated with Routine Jobs Destruction and College Wage Growth

Our theory builds on two premises that link the fall in the price of computer technology in the 1980s, and its rapid adoption by firms, to changes in local labor markets[5]. First, the adoption of such technology is associated with the substitution of routine labor by capital, which is the basis of the routinization channel. Second, the penetration of computers in firms is associated with productivity increases — hence wage increases — for college-educated workers, which works as the basis of the consumption spillover theory. We present evidence in support of these two premises, exploiting data on computer penetration initially put together by Beaudry, Doms, and Lewis (2006). These data are based on private sector surveys and measure the number of personal computers per employee at the firm level, in 1990 and 2002. They are purged from industry and establishment size fixed effects and aggregated to the commuting zone level. The data are directly obtained from Autor and Dorn (2013), who provide changes in the computer penetration measure over 1980-1990 and 1990-2002, assuming that penetration was zero in 1980.

We estimate two linear models to evaluate the relationship between the adoption of computers,

[5]For more specific and detailed evidence on the impact of computer technology adoption on the labor market, see e.g. Doms, Dunne, and Troske (1997), Beaudry, Doms, and Lewis (2010), and Michaels, Natraj, and Van Reenen (2014). On the other hand, Houseman (2018) debates the relevance of automation to explain the decline in US manufacturing employment and gives more emphasis to trade.

changes in routine employment and changes in college wages:

$$\Delta y_{c,t} = \alpha_{d(c),t} + \beta \Delta \text{PC}_{c,t} + \epsilon_{c,t} \tag{2.3.1}$$

where $\Delta y_{c,t}$ measures either (i) the change in routine occupations' share within noncollege employment or (ii) the growth in college-educated workers' wages in commuting zone c over period t = 1980-1990 or 1990-2000[6]; $\alpha_{d(c),t}$ is a division-by-period fixed effect; $\Delta \text{PC}_{c,t}$ is the change in the computer peneration measure; $\epsilon_{c,t}$ is an error term.

Table 2.5 shows the results of these regressions. A one standard deviation in the measure of computer penetration is associated with a -0.7% change in the share of routine employment, to be compared to a mean change of 0.08% over the sample; it is also associated with a 2.7% increase in college wages, compared to a mean increase of 6.0%. Computer penetration is strongly associated with both the growth of routine employment (negatively) and the growth of college wages (positively), consistent with our view of a common shock driving routinization and consumption spillovers.

2.3.3 Evidence for the Routinization Theory

The routinization hypothesis was initially proposed by Autor, Levy, and Murnane (2003). According to this theory, recent technological change does not simply favor high-skill relative to low-skill workers (unlike standard skill-biased technological change). It rather "substitutes for workers in performing cognitive and manual tasks that can be accomplished by following explicit rules [routine tasks]" (Autor, Levy, and Murnane, 2003). Routinization thus refers to labor-displacing automation of codifiable tasks. Workers in occupations that involve a lot of routine tasks (henceforth, routine occupations), like machine operators or administrative clerks, were particularly affected. On the other hand, workers in low-skill service occupations have been less affected by technology. These occupations typically consist of assisting and caring for other. They include e.g. food service workers, security guards, janitors and gardeners, cleaners, home health aides, child care workers, hairdressers and beauticians, recreation occupations, etc. Because these jobs require human contact, they are less automatable. Among noncollege workers, routine occupations are typically found in the middle of the wage distribution, while service occupations are found towards the bottom. The

[6]We match data on computer peneration in 2002 to American Community Survey data in 2000.

differential effect of technology on these two sets of occupations has, according to the routinization theory, generated a reallocation of noncollege employment from routine to service tasks (Autor and Dorn, 2013). Noncollege employment has thus shifted from the middle to the bottom of the wage distribution.

In 1980, there is substantial variation in the noncollege employment share of routine occupations across commuting zones, ranging from 21% at the 5th percentile to 37% at the 95th. According to the routinization theory, regions initially employing more routine labor are expected to benefit more from the adoption of new technologies (capital) that allow the automation of workers' tasks and a great reduction in production costs. Relative to other regions, the employment share of routine occupations should thus drop by more, and that of service occupations rise by more. To show support for this hypothesis, we use data on noncollege employment shares in broad occupation groups by commuting zone, in 1980 and 2005. We split commuting zones into two groups: the first group contains commuting zones with a share of noncollege routine employment below the national mean in 1980, while the second contains those with a share above the mean. We compute the changes in employment between 1980 and 2005 for each of these occupation groups and the two regional groups.

We show the results on figure 2.1, which provides two pieces of evidence in support of routinization. First, while Admin/Retail, Operators and Production occupations experience relative losses in employment in both regional groups, these losses are more pronounced in routine-intensive regions. These occupational categories are the ones that contain the largest concentration of routine employment, according to table 2.6. Second, commuting zones that are more routine-intensive experience greater growth in service employment, relative to other occupation groups. Routine-intensive regions experience on average a 6.7 p.p. increase in the share of noncollege service employment, whereas regions that are less reliant on routine labor experience a 5.0 p.p. increase. The change in share is thus 1.34 times higher in routine intensive regions. Taken together, these two pieces of evidence provide support for the routinization hypothesis. Noncollege labor reallocation towards service occupations is stronger in routine-intensive commuting zones, and this is associated with excess losses in categories that contain a large share of routine employment. The overall pattern is consistent with the displacement of routine workers towards service occupations.

2.3.4 Evidence for the Consumption Spillovers Theory

The consumption spillover hypothesis proposes an alternative explanation for the growth of non-college service occupations' employment share (Manning, 2004; Mazzolari and Ragusa, 2013; Liu and Yang, 2021). It builds on three observations: the real wages of college-educated workers have dramatically increased since the 1980s; and the share of a household's budget spent on services increases with income; services purchased by households are largely fulfilled by noncollege service workers. As described in Autor, Levy, and Murnane (2003), recent technological progress "complements workers in performing nonroutine problem-solving and complex communications tasks", which has led to strong wage growth among college-educated workers. Moreover, consumption data show that household spending on services (food away from home, maintenance services, child care, etc.), as a share of total expenditures, increases with income. In other words, there is non-homotheticity in consumption. Since services that households purchase generally substitute for home production, this observation is consistent with the fact that better-paid workers have a higher opportunity cost of time (Becker, 1965). The combination of these features paves the way for another explanation in the differential growth of noncollege service occupations: as college wages rise thanks to technological progress, non-homotheticity implies a disproportionate increase in households' demand for services; noncollege service workers experience a positive labor demand shock, lifting the employment share of these occupations typically found at the bottom of the wage distribution.

The consumption spillover theory builds on two premises. First, that consumption of goods and services exhibits non-homotheticity: the richer individuals are, the more their consumption expenditures will be concentraded on services. This view is consistent with the existence of a subsistance level of goods consumption (e.g. food, water, personal care products, etc.) and the relative luxury status of services (e.g. eating out, entertainment, home and car maintenance, etc.). The second premise is that real wages have increased, allowing the consumption spending shift from goods to services to happen. We show evidence that these two conditions have been met in recent decades.

To provide evidence for the existence of non-homotheticity in goods and services consumption, we

harness consumption data from the CEX. Figure 2.2 shows the relation between the hourly wage of a household head and the share of the household's consumption expenditures allocated to services. There is a clear positive association between the two. Expenditure shares on services range from about 11.5% on average for households with a head making $5 per hour (the federal minimum wage in 2005 is $5.15), to about 14% for households with a head making $50 per hour or more.

Turning to the second premise of the consumption spillover theory, we provide evidence that real wages have dramatically increased, in particular for workers with a college degree. In figure 2.3, we show the evolution of the spatial distribution of hourly wages[7], between 1980 and 2005, for college and noncollege workers. Over that period, the distribution of noncollege workers' wages shows little changes, going from a mean of $14.4 per hour to 14.6 and standard deviation of $1.67 per hour to 1.61. On the other hand, the distribution of college-educated workers experience a striking evolution. In 1980, college workers made on average $20.1 per hour, against 25.2 in 2005, or a 25% increase. The spatial standard deviation of their wages has doubled, from $2.36 per hour to 4.61. The increase in the mean wages of college workers creates space for the consumption spillover theory to operate. The increase in standard deviation is also interesting on its own: the important spatial disparities in the growth of college wages creates an opportunity to identify the strength of the consumption spillover channel by comparing regions with more or less growth. Henceforth, our proxy for the consumption spillover theory will thus be the local growth in college wages.

As for the routinization hypothesis in section 2.3.3, we provide a preliminary test for the consumption spillovers theory by comparing growth in noncollege service employment across regions with different growth rates of college workers' wages. We again split commuting zones into two groups: a first group with commuting zones experiencing college wage growth below the 1980-2005 national mean, and a second group with wage growth above the mean. Figure 2.4 shows the changes in employment shares of major occupation categories in these two geographical groups. Service employment grows more, relative to other occupations, in regions that experience stronger college wage growth: its employment increases by 6.8 p.p. on average in high-wage-growth regions, against 4.9 p.p. in other regions, which corresponds to a factor 1.39 between the two.

[7]For 1980 and 2005 and for each commuting zone, we compute the mean wages of college and noncollege workers. We then plot the distribution of these average wages across commuting zones.

2.3.5 Confronting the two theories in a regression framework

Preliminary evidence is consistent with both the routinization and consumption spillovers theories: changes in the noncollege employment share of service occupations is positively associated with the initial share of routine labor and the contemporaneous growth in college wages. However, the tests presented so far suffer from an important drawback: as shown on figure 2.5, the regional proxies for both theories are positively correlated across commuting zones. This raises confounding concerns: providing evidence for either theory using the above proxies puts us at risk of capturing the other theory as well. To alleviate this concern, we turn to an augmented version of the data used so far and employ a regression framework, as in Autor and Dorn (2013). Instead of considering the entire 1980-2005 period, we stack the data over three periods — 1980-1990, 1990-2000 and 2000-2005. We also take into consideration other labor market features that may impact the growth of service employment in our analysis.

The regression specification we adopt is the following:

$$\Delta \text{SVC}_{c,t} = \beta \text{SR}_{c,t} + \gamma \Delta \text{CW}_{c,t} + X'_{c,t}\delta + \alpha_{d(c),t} + \epsilon_{c,t} \tag{2.3.2}$$

$\Delta \text{SVC}_{c,t}$ denotes the change in noncollege employment share of service occupations in commuting zone c, over the period t = 1980-1990, 1990-2000 or 2000-2005. $\text{SR}_{c,t}$ measures the noncollege employment share of routine occupations at the beginning of period t; it is the proxy we retain for the routinization theory. $\Delta \text{CW}_{c,t}$ measures the growth of college-educated workers' wages over period t; it is the proxy we retain for the consumption spillover hypothesis. $X_{c,t}$ represents a number of additional labor market controls, measured either in levels at the beginning of period t or in contemporaneous changes over t. Finally $\alpha_{d(c),t}$ are division-by-year fixed effects ($d(c)$ denotes the Census division in which commuting zone c lies). All variables that describe changes over a period are transformed into decadal change equivalents.

The estimates from this regression are presented in table 2.7. Columns 1 and 2 focus on each theory separately and confirm the preliminary evidence presented so far. The coefficients on the proxies for both routinization and consumption spillovers theories are positive and significant at $p = 0.01$. The mean change in the share of noncollege service employment across all three periods is 2.63%. For comparison, a one standard deviation change in the initial share of routine employment is associated

with a $0.053 \times 0.106 = 0.56\%$ increase in service employment share; a standard deviation change in the growth rate of college wages is associated with an increase of $0.10 \times 0.042 = 0.42\%$. Studied in isolation, both theories are thus consistent with the data and are associated with increases in service employment of important and similar magnitudes.

Since the proxies for both theories are positively correlated across space, it is important to include them both in a regression model. Column 3 of table 2.7 does so and shows evidence consistent with the correlation: the coefficients are still both significant but lower in magnitude than in columns 1 and 2. The coefficient related to routinization decreases by 13% (relative to column 1) while the coefficient related to consumption spillovers decreases by 38% (relative to column 2). A simple calculation shows that a standard deviation change in the initial share of routine employment is associated with a change in the outcome that is 1.9 times bigger than the effect of an equivalent change in the growth rate of college wages. First, these findings suggest that the estimates in the first two columns suffered from upward bias, particularly so for the coefficient associated with consumption spillovers. Studying these two theories in isolation by exploiting geographical variations, which is what the literature has focused on so far, is thus prone to biases. Second, relative to the findings from columns 1 and 2 that both channels had effects of similar magnitudes, taking both channels into account suggests that consumption spillovers may matter less than routinization for the growth of service occupations. Columns 4 and 5 of table 2.7 test the robustness of the previous findings to a host of labor market controls. Despite some loss in significance, the main coefficient estimates remain mostly significant.

Alternative specifications. A concern with the main specification presented in 2.7 is whether the consumption spillover effect is specific to college-educated workers. To test this hypothesis, we run the same regression using the growth in wages of noncollege workers who do not work in service or routine occupations. The rationale for excluding these occupations is that their wages are mechanically related to the growth in service occupations: the routinization theory ties a drop in labor demand for routine workers to a rise in employment in services, while the consumption spillover theory ties a disproportionate increase in demand for services to employment growth in that sector. Noncollege workers in other occupations should however not be directly impacted,

which makes them better candidates for testing the hypothesis. Table 2.8 presents the results, using the same five configurations as in the main specification. The takeaway is clear: in all configurations, growth in the wages of other noncollege workers is not significantly associated with growth in service employment, and the coefficients are an order of magnitude smaller than those corresponding to the growth in college wages in table 2.7. This confirms our hypothesis that the growth in college workers' wages is a particular driver of growth in the service sector.

We further provide evidence that our main specification does not give a mechanical advantage to the consumption spillover theory. The outcome and the proxy for this theory are measured in changes, while the proxy for routinization is measured in levels at the beginning of each decade. This could give an unfair advantage to the former theory, since changes are more likely to be correlated with each other than with an initial level. To alleviate that concern, we run the same regression using the growth of routine emp0loyment as a proxy for the routinization theory. Table 2.9 presents the results. As expected, growth in routine occupations is negatively related to the change in service occupations' employment share in all configurations. Moreover, using this alternative proxy for routinization yields highly significant coefficients on the growth of college workers' wages, with magnitudes comparable to those found in table 2.7. This suggests that the main specification does not give a mechanical advantage to the consumption spillover theory.

The inclusion of determinants of both routinization and consumption spillovers in a common regression framework alleviates direct confounding concerns, and suggests that both mechanisms operate at the same time in the data. However, this analysis makes strong linearity assumptions and relies on a choice of proxies for both theories, which may introduce additional approximations and confounding concerns. To further strengthen our case, we design and estimate a model that embeds both mechanisms and allows us to evaluate their respective importance.

2.4 Model

We develop a spatial structural model that allows both routinization and consumption spillovers to play a role in the growth of service occupations. Our view is that these theories are distinct

consequences of a *common technological shock*[8] that (i) displaces routine workers and (ii) increases college workers' wages. The diagram in figure 2.6 summarizes this view. Building on Autor and Dorn (2013), who propose a model reflecting only the routinization mechanism, we add non-homotheticity in consumption, consistent with the observation that richer households disproportionately spend on services. This enrichment allows regions that see higher growth in college wages is stronger to experience stronger demand for locally-provided services, hence for service employment. In other words, it adds the possibility of consumption spillovers. Although the model features multiple locations, the description that follows is made for an arbitrary location and the corresponding indices are dropped for readability. The presentation of equilibrium in section 2.4.4 introduces location indices and makes the spatial aspect of the model explicit.

2.4.1 Production

There are two sectors indexed by j: sector g produces goods while sector s produces services. Each location has a representative, perfectly competitive firm in each sector. Factors of production consist of capital K and three types of labor: abstract labor L_a, routine labor L_r and manual labor L_m. There are two types of workers: high-skill and low-skill. High-skill workers can only provide abstract labor, while low-skill ones can provide either routine or manual labor.

The goods sector uses three of these factors, routine labor L_r, abstract labor L_a, and capital K, through the following production technology:

$$Y_g = L_a^{1-\beta}\left[\alpha_r L_r^{\mu} + (1-\alpha_r)K^{\mu}\right]^{\beta/\mu} \tag{2.4.1}$$

where $\beta, \mu \in (0,1)$. In this production function, the elasticity of substitution between abstract labor and the K-L_r aggregate is 1, while the elasticity of substitution between routine labor and computer capital is $1/(1-\mu)$ and, by assumption, is greater than 1. By implication, K is a relative complement to abstract labor and a relative substitute for routine labor. These assumptions are at the center of the routinization hypothesis: everything else equal, when the price of capital falls, the

[8]More precisely, we think of this technological shock as an exgenous real price decline of computers, that appears strikingly in the early 1980s.

optimal behavior of the firm implies that routine labor gets displaced by capital, while demand for abstract labor rises thanks to its complementarity with capital.

Taking goods as the numeraire, given wages for routine and abstract labor, w_r and w_a, and the rental price of capital p_k, the goods' firm solves the following problem:

$$\max_{L_a, L_r, K} \quad \pi_g = Y_g(L_a, L_r, K) - w_a L_a - w_r L_r - p_k K \tag{2.4.2}$$

The service sector uses only manual labor L_m through the following production technology:

$$Y_s = \alpha_s L_m \tag{2.4.3}$$

Given the price of services p_s (relative to the good numeraire) and wages w_m paid to a unit of manual labor, the service-producing firm solves the following problem:

$$\max_{L_m} \quad \pi_s = (\alpha_s p_s - w_m) L_m \tag{2.4.4}$$

2.4.2 Capital

The fall in the price of capital p_k, consistent with the observed decline in the cost of computers from the 1980s, is the single shock driving the model simulations that will be run. This price is assumed to be exogenous, and capital purchased from outside of the economy, akin to a small open economy model.

2.4.3 Workers

In each location, there is a unit-mass continuum of low-skill workers and a mass χ of high-skill workers. All workers supply a unit of labor inelastically, and low-skill workers must choose whether to provide routine or manual labor.

Occupational Choice of Low-Skill Workers

Without heterogeneity in low-skill workers' efficiency in performing manual and routine tasks, all their labor supply would go to the occupation that offers the highest wage. To have positive labor supply in each of the two low-skill occupations, we assume that these workers (i) all have the same skill level for performing manual tasks, but (ii) have individual skill-level η for performing routine tasks. η is assumed to follow an exponential distribution with parameter δ. The associated probability distribution function is thus $f(\eta) = \delta \exp(-\delta\eta)$. Let w_r be the per-skill-unit wage in routine ocupations, and w_m be the wage in manual occupations. A worker with routine skill η chooses to work in routine occupations if and only if

$$\eta w_r \geq w_m \tag{2.4.5}$$

This defines a cutoff efficiency $\eta^* = w_m/w_r$ such that all low-skill workers with $\eta \leq \eta^*$ provide manual labor, while those with effiencies above this cutoff provide routine labor. In efficiency units, labor provided in each low skill occupations are given by:

$$L_m^S = \int_0^{\eta^*} \delta \exp(-\delta\eta)\, d\eta = 1 - \exp(-\delta\eta^*) \tag{2.4.6}$$

$$L_r^S = \int_{\eta^*}^{\infty} \eta\delta \exp(-\delta\eta)\, d\eta = (1/\delta + \eta^*) \exp(-\delta\eta^*) \tag{2.4.7}$$

They can be linked through:

$$L_r^S = \frac{1}{\delta}\left(1 - \log(1 - L_m^S)\right)\left(1 - L_m^S\right) \tag{2.4.8}$$

Consumption

Workers consume both goods and services to maximize the following CES utility function

$$u(c_g, c_s) = \left(\theta_g (c_g - c_0)^{\frac{\sigma-1}{\sigma}} + c_s^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \tag{2.4.9}$$

A worker earning income I (accounting for efficiency units) thus solves the problem

$$\begin{aligned} \max_{c_g, c_s} \quad & u(c_g, c_s) \\ \text{s.t.} \quad & c_g + p_s c_s = I \end{aligned} \tag{2.4.10}$$

The c_0 term in the utility function is key to the consumption spillover channel. It generates non-homotheticity in demands for goods and services. When $c_0 > 0$, households with higher income spend a greater share of their income on services, consistent with the evidence presented in section 2.3.4. Solving the utility maximization problem for a worker earning a wage w and facing a relative price of services p_s allows to express the share of income that they spend on services:

$$x(p_s, w) \equiv \frac{p_s c_s}{p_s c_s + c_g} = \frac{\theta^{-\sigma} p_s^{1-\sigma}}{1 + \theta^{-\sigma} p_s^{1-\sigma}} \left(1 - \frac{c_0}{w}\right) \tag{2.4.11}$$

Plotting this function on figure 2.7 for different c_0 and θ offers important partial equilibrium insights. Moving from a configuration with $c = 0$ to $c > 0$ holding θ constant leads to a reduction in expenditure shares $x(p_s, w)$ at all wage levels. In this model, the share of services in total consumption and the share of service occupations in noncollege employment are tightly linked. θ is thus an essential parameter that allows adjusting the average expenditure share on services across the economy when changing the value of c_0. As shown on figure 2.7, reducing θ creates the desired situation where high-wage earners spends more on services that in the homothetic case, while low-wage earners spend less.

2.4.4 Equilibrium

Mobility assumptions. The model describes N locations, indexed by n. We make two important assumption regarding the mobility of goods, services across locations:

Assumption A1. Goods produced in each location are assumed to be perfect substitutes in consumption and costlessly tradable across regions.

Assumption A2. Services can instead only be consumed locally.

The perfect substitutability and tradability of goods is a simplifying assumption that we endorse since the focus of this paper is not to study trade. It implies that goods prices are the same

everywhere. Since we also take goods as the numeraire, goods have a unit price in all locations. All prices in the model are therefore understood as "per unit of the tradable good". The nontradability assumption for services guarantees that shocks to manual labor supply and demand have local effects. It is consistent with the nature of service occupations, whose vast majority requires proximity to consumers, if not face-to-face contact.

We assume that labor is immobile. This allows us to focus on the effects of routinization and consumption spillovers, which are theories that do not take a precise stance on the impacts of migration. Moreover, reduced-form evidence in table 2.7 showed that changes in the local share of college workers was not significantly associated with changes in noncollege service employment.

Assumption A3. Labor is immobile across locations.

Equilibrium definition. Equilibrium is defined by employment levels L_{zgn}, wage levels w_{zgn}, consumption amounts $\{c_{zgn}, c_{zsn}\}$ for each worker type $z \in \{a, r, m\}$ and each location $n \in [1, N]$, relative service prices p_{sn} and capital amounts K_n for each location n such that

- consumption amounts solve the utility maximization problems of workers given wages and prices
- employment levels and capital amounts solve goods and services firms' profit maximization problems, given wages, p_{sn} and p_k
- the markets for goods and services output clear
- all labor markets clear

2.5 Model Estimation

The goal of the model is to assess the relative importance of the routinization and consumption spillover channels in explaining spatial disparities in the growth of service employment. In the context of the model, that is equivalent to asking whether including a realistic level non-homotheticity — by estimating c_0 — quantitatively matters for that outcome. To answer that question, we estimate two versions of our model: one where homothetic preferences are imposed ($c_0 = 0$ is fixed, as

in Autor and Dorn (2013)) and one where c_0 is instead estimated based on the non-homothetic consumption patterns observed in the data. In each of these models, we simulate a fall in the price of capital, and compare the implications in terms of service employment shares.

2.5.1 Mapping between Model and Data Locations

To match with the spirit of the regressions in section 2.3.5, we estimate each model with multiple locations and compare the changes in service employment across locations as the price of capital falls. The only parameter we allow to vary across locations is β, and we can thus use this parameter to characterize a location in our model. β drives the strength with which a drop in the price of capital affects a given location. A high value for β implies more substitution of routine labor by capital, as well as a stronger increase in abstract labor demand through complementarity with capital. In a model with homothetic preferences, differences in β across locations will only drive a strength differential in the routinization channel. In a model with non-homotheticity, consumption spillover differentials will add to the spatial disparities.

For simplicity, we consider models with two locations: a high-β location and a low-β one. Each of these model locations is mapped to multiple commuting zones in the data. Due to the Cobb-Douglas aspect of the goods production function, β can be directly linked to the income share of

As described at the beginning of this section, we need to map each of the two locations in our model to a group made of multiple CZs in the data. To highlight the roles of routinization and consumption spillovers in the growth of service employment, we need to split commuting zones along dimensions that proxy for these theories. Due to the structure of the model, the share parameter β governs both proxies. On the one hand, the location with the higher β is more reliant on routine labor and capital in goods production; it will thus substitute more capital for routine labor as the price of capital falls, and noncollege workers will be displaced towards service occupations; this is the routinization theory. On the other hand, the location with the higher β also has stronger complementarity between capital and abstract labor; college workers will thus concentrate and experience stronger wage growth in that location, which is the vector of consumption spillovers. As the identification conditions for β showed, it is closely tied to the income share of college labor.

We thus decide to split the data along that dimension: the first location in the model is mapped to commuting zones in the data with a 1980 college labor income share below the national mean (weighting by 1980 population shares) and the second location is mapped to commuting zones above the mean. The mapping is shown on figure 2.8.

2.5.2 Calibrated Parameters

We calibrate a subset of the parameters in both models and provide sensitivity checks on these choices in the appendix. In the main specification, the production elasticity of substitution between routine labor and capital, μ, is set to 0.25, following estimates provided in Krusell et al. (2000). The consumption elasticity of substitution between goods and services, σ, is set to 0.2 following Barany and Siegel (2020). α_s and θ both drive the overall output share of services, from the supply and demand side respectively. Restrictions imposed by the structure of the model imply that we cannot estimate them separately. We choose to calibrate $\alpha_s = 1$ and estimate θ[9]. Since θ is a preference parameter, we estimate a single value for all locations in the model. We also calibrate the skill distribution parameter $\delta = 1$ following (Autor and Dorn, 2013). The mass of high-skill workers χ is assumed constant across location and time and is calibrated to match the national average share of college-educated workers over 1980-2005, which equals 51%. That implies $\chi/(1-\chi) = 0.51$, or $\chi = 0.34$. Finally, we also calibrate the initial and final value for the price of capital p_k to match data from the Penn World Tables. Table 2.10 summarizes our preferred calibration.

2.5.3 Estimated Parameters and Identification Conditions

Locations in the model are characterized by their β, which is the only location-dependent parameter. We choose this restriction to focus on the relevant mechanisms: a location with high-β will adopt capital faster as its price decrease, leading to more routine labor displacement and college wage growth, which are the sources of routinization and consumption spillovers respectively. We estimate α_r, θ and c_0 in a location-independent way. The rest of this section provides identification conditions

[9]Our rationale for choosing not to estimate α_s in a location-specific way, instead of a spatially constant θ, is to avoid adding other location differentials beyond the scope of this paper.

for these key parameters.

Identification of β. β is a location-specific parameter. For a given location n in the model, the first order condition of the goods-producing firm with respect to abstract labor gives the following equation that identifies β_n:

$$\beta_n = 1 - \frac{w_{an} L_{an}}{Y_{gn}} \tag{2.5.1}$$

Identification of c_0. The size of c_0 influences the strength of the dependence between a worker's income and the share of their income they spend on services. Identification of c_0 will thus intuitively come from comparing the service expenditure shares of workers of different incomes (i.e. of different occupations in the model) and matching the differences between the model and the data. Using the optimal demand functions for goods and for services derived from the utility maximization problem, we can show that a worker earning income w and facing a price of services p_s will spend the following share of their income on services:

$$x(w, p_s) \equiv \frac{p_s c_s(w, p_s)}{c_g(w, p_s) + p_s c_s(w, p_s)} = \frac{\theta^{-\sigma} p_s^{1-\sigma}}{1 + \theta^{-\sigma} p_s^{1-\sigma}} \left(1 - \frac{c_0}{w}\right) \tag{2.5.2}$$

This equation shows that expenditure share has four determinants: the parameters θ and c_0, the local price of services and the income of the worker. We want to compare the service expenditure shares of workers of different incomes w, w' in the same location, so that they face the same local price of services. If we did so by taking the difference in expenditure shares $x(w, p_s) - x(w', p_s)$, the equation above shows that we would keep a dependence on the price p_s and on the parameter θ. Since we don't have direct data on local prices and we want to isolate the c_0 parameter, this approach is not appropriate. Instead, we consider the ratio of service expenditure shares across different incomes:

$$\frac{x(w, p_s)}{x(w', p_s)} = \frac{1 - \frac{c_0}{w}}{1 - \frac{c_0}{w'}} \tag{2.5.3}$$

Here, the dependence on local prices has canceled out. Inverting this equation leads to

$$c_0 = \frac{x(w, p_s) - x(w', p_s)}{\frac{x(w, p_s)}{w'} - \frac{x(w', p_s)}{w}} \tag{2.5.4}$$

Identification of θ. Intuitively, θ governs the relative preference for goods vs services in consumption. We can use the service expenditure shares of college workers $x(w_a, p_s)$ to identify this parameter. Since the model imposes that $p_s = w_m$, we can write this expenditure share as

$$x(w_a, p_s) = \frac{\theta^{-\sigma} p_s^{1-\sigma}}{1 + \theta^{-\sigma} p_s^{1-\sigma}} \left(1 - \frac{c_0}{w_a}\right) = \frac{\theta^{-\sigma} w_m^{1-\sigma}}{1 + \theta^{-\sigma} w_m^{1-\sigma}} \left(1 - \frac{c_0}{w_a}\right)$$

We can invert this equation to obtain

$$\theta = \left(\frac{1 - c_0/w_a - x(w_a, p_s)}{x(w_a, p_s)}\right)^{\frac{1}{\sigma}} w_m^{\frac{1-\sigma}{\sigma}}$$

All terms appearing on the right-hand-side of this equation are known. c_0 can be pinned down independently from other parameters following the identification procedure described above. We obtain the wages of service and college workers, as well as the expenditure shares of college ones from CEX data.

Identification of α_r. Taking the ratio of first order conditions for L_r and K from the profit maximization problem of goods producing firms (see appendix 2.8.2 for details) yields

$$\frac{\alpha_r}{1 - \alpha_r} \left(\frac{L_r}{K}\right)^{\mu - 1} = \frac{w_r}{p_k}$$

L_r does not have a direct empirical counterpart because it mixes number of workers and efficiency units. K is also an abstract measure of the amount of capital and does not have a clear data counterpart. However, we have data on the prices of these inputs (w_r and p_k) as well as on the income they generate ($w_r L_r$ and $p_k K$). Rearraging the above equation, we can write

$$\frac{\alpha_r}{1 - \alpha_r} = \left(\frac{w_r L_r}{p_k K}\right)^{1-\mu} \left(\frac{w_r}{p_k}\right)^{\mu}$$

Inverting this equation gives us identification of α_r

$$\alpha_r = \frac{1}{1 + \left(\frac{w_r L_r}{p_k K}\right)^{1-\mu} \left(\frac{w_r}{p_k}\right)^{\mu}}$$

2.5.4 Data Moments and GMM Conditions

We estimate the model solely based on initial, 1980 data. This data consists of the 722 commuting zones covering the entirety of the mainland US. For each of these locations, we compute four moments that are the basis of our estimation procedure: the income shares of abstract, routine and manual labor, and the mean difference in service expenditure shares between manual and abstract workers. The following paragraphs describe how these moments are computed.

Calculation of income shares. Census data allow us to compute total labor income for abstract, routine and manual workers by CZ, by simply multiplying workers' annual wages by employment. For capital income, we use data from the BEA and our imputation method based on local industry composition, described in section 2.2.4.

Calculation of expenditure shares. The identification of the non-homotheticity parameter c_0 relies on differences in service expenditure shares between workers with different incomes. CEX data include information on education and occupation that perfectly matches with the Census broad occupational and educational categories. We can thus identify noncollege service workers and college workers in the CEX (but we do not have detailed enough information to identify routine workers) and compute differences in their service expenditure shares. To the extent that the model predicts different wage levels for manual and abstract labor, we can identify c_0 by matching the differences in service expenditure shares between these workers in the model to the equivalent moment in the data. Since the data on income shares is at the commuting zone level, we need to build expenditure share moments at the same level. To do so, we use a subset of the CEX for which information on the MSA of residence is available (this subset is made of surveyed individuals that live in large US cities). We compute the difference between the average service expenditure share of college workers with that of noncollege service workers. We then map MSAs into CZs and obtain moments to identify c_0 on a subset of CZs.

In all, we have two models with two locations each. In the first model, where preference homotheticity is imposed, there is one location-dependent parameter (β) and two spatially-constant

parameters (α_r, θ) to estimate, for a total of four parameters. In the model with non-homothetic preferences, we moreover estimate c_0, for a total of five parameters. Following the identification conditions outlined above, we thus build five moment conditions that allow us to exactly identify the models.

2.5.5 Estimation Results

Table 2.11 shows the parameter GMM estimates and standard errors in the case of the models without labor mobility. The first panel shows results in the case of imposed homothetic preferences, while the second shows the non-homothetic case, where we estimate c_0. The estimates are all significant at the 1% level.

Two important features of the results are readily noticeable from this table. First, the estimates for β in each location change very little across the two model specification. In the homothetic model, the gap in β across locations reflects the importance of the routinization channel in explaining spatial differences in service employment growth. The fact that these estimates change little when adding non-homotheticity suggests that consumption spillovers play a role that is separate from routinization.

Second, in the non-homothetic model, the estimate for c_0 is positive and highly significant. To get a better sense of the size of this coefficient, it is useful to compare it to goods consumption levels by the three types of workers in the initial period. Let $\bar{c}_{zg,1980}$ be the goods consumption amount of worker type z in year 1980, averaged across both locations in the model. We compute

$$\frac{c_0}{\bar{c}_{zg,1980}} \qquad \text{for } z = a, r, m$$

and obtain 18.6% for college workers, 22.7% for routine workers and 94.8% for service workers. These numbers highlight strong non-homotheticity. Service workers are close to the subsistance level c_0, while abstract workers consumes more than fives times more that amount of goods. These findings are qualitatively consistent with estimates provided in the recent macroeconomic literature (Comin, Lashkari, and Mestieri, 2021; Barany and Siegel, 2020).

The following section presents simulations of a fall in the price of capital in both models, with

and without non-homotheticity in consumption. It discusses the effect differences on service employment and wages, and explores the mechanisms behind these differences in light of the routinization and consumption spillover theories.

2.6 Discussion of Model Results

2.6.1 Simulations

The simulations we present here consist of solving the two models, with homothetic and non-homothetic preferences, using calibrated and estimated parameters from section 2.5, along a path of falling price of capital p_k. This path is taken from the Penn World Tables data described in section 2.2.5 and plotted in figure 2.9.

Our model is stylized. Its goal is not to provide a precise fit of many aspects found in the data, but rather to give a simple and insightful quantification of the relative importance of routinization and consumption spillovers in explaining the growth in service employment.

2.6.2 Capital Adoption and Routine Labor Displacement

We first show that simulations are consistent with the premises of the routinization hypothesis. The location with a higher β is more reliant on the routine-capital aggregate in goods production. It thus experiences faster capital adoption and greater displacement of routine labor as the price of capital falls. Figure 2.10 plots the corresponding simulations. Capital is plotted in percentage change relative to 1980, while routine employment is plotted in absolute differences relative to 1980 (because routine employment is already expressed as a share in the model).

2.6.3 The Growth of Service Employment and Wages

Our main objects of interest are the evolution of noncollege service workers' employment share and wages. Figure 2.11 plots the simulations of these two variables along the falling path of capital price. Consistently with regression specifications in section 2.3.5, service employment is shown

in absolute differences relative to 1980 (because it is already expressed as a share in the model), while wages are shown in percentage change relative to 1980.

The homothetic model (solid lines) yields results consistent with the routinization hypothesis. The greater displacement of routine labor in the high-β location induces an stronger inflow of noncollege workers into service occupations. Over time, the gap in service employment between the two locations thus widens. The rise in wages relies on the imperfect substitutability between goods and services in consumption: as the relative price of goods falls (because with cheaper capital, goods cost less to produce), a dominant income effect raises the demand for services. The non-homothetic preferences model (dashed lines) yields similar patterns but with employment and wage gaps that widen faster over time. This pattern is consistent with coexistence and complementarity of the routinization and consumption spillover theories. In the model with non-homotheticity, the employment gap widens 57% faster and the wage gap 25% faster.

2.6.4 Impact of Consumption Spillovers

How does the introduction of non-homotheticity, hence of consumption spillovers, contribute to the increased spatial dispersion in service employment and wage growth? Figure 2.12 shows the evolution of college workers' wages w_a and expenditure shares on services

$$x_{as} = \frac{p_s c_{as}}{p_s c_{as} + c_{ag}}$$

across both models. College wages are shown as percentage change relative to 1980, while the expenditure share is shown are absolute changes relative to 1980. Both models predict that college wages rise faster in the location with the higher β, due to a higher rate of capital adoption in that location, combined with L_a-K complementarity. Interestingly, both models yield extremely similar evolutions of college wages, consistent with similar estimates for β in both locations.

Turning to expenditure shares on services, both models predict that these shares grow faster in the high-β location. In the model with homothetic preferences, wages do not influence expenditure shares: the difference in their evolutions across locations is thus solely due to different paths of local prices of services p_s. In the model with non-homotheticity, the spatial gap between service expenditure shares of college workers widens faster, by a factor 1.47. This reflects two phenomena:

(i) the gap in the local price of services widens faster, and (ii) workers with higher wages tend to spend a larger share of their income on services at equal prices.

Both models predict a pattern of wages and expenditure shares for service and routine workers that is similar to that of college workers (see figure 2.13). In particular, the non-homothetic utility function induces an increasing and concave relationship between wages and expenditure shares on services (keeping prices constant): workers experiencing wage growth from a low starting point thus increase their expenditure share of services relatively more than richer workers. In theory, it could thus be that routine and service workers' changes in service spending are driving growth in the service sector.

We thus need to decompose the spatial gap in growth of service employment into contributions from each worker type's changes in consumption. We rely on the following model identity that states that service output, hence service employment, equals the sum of service consumptions by different worker types:

$$L_m = Y_s = c_{as}L_a + c_{rs}L_r + c_{ms}L_m \equiv C_{as} + C_{rs} + C_{ms} \qquad \text{for each location}$$

and decompose the spatial gap in the following way:

$$\Delta L_m = \Delta C_{as} + \Delta C_{rs} + \Delta C_{ms}$$

where Δ is an operator performing a double difference between time and location:

$$\begin{aligned}\Delta X = &\Big[X(t = 2005, \text{location} = \text{high-}\beta) - X(t = 2005, \text{location} = \text{low-}\beta)\Big] \\ &- \Big[X(t = 1980, \text{location} = \text{high-}\beta) - X(t = 1980, \text{location} = \text{low-}\beta)\Big]\end{aligned}$$

To compare the relative contributions of each worker type's changes in consumption of services, we take the ratio of these changes over the change in service employment[10]:

$$\forall z \in \{a, r, m\} : \quad \text{contrib}\,(z) = \frac{\Delta C_{zs}}{\Delta L_m}$$

Figure 2.14 shows the results of this decomposition by plotting $\text{contrib}\,(z)$ for each worker type in each model. In both models, college workers' consumption demand contributes the most to changes

[10]This way of comparing relative contributions is only valid when all components of change are of the same sign, which is the case here: consumption of services increases for all workers between 1980 and 2005.

in service employment, followed by routine and service workers. Looking at the differences between the two models, we see that the relative contribution of college workers increases when adding non-homotheticity. The fact that the cross-location gap in service employment widens faster under non-homotheticity is thus mostly driven by the evolution of college workers' consumption, who shift towards service expenditures as their income rises. This is qualitatively consistent with the reduced-form evidence presented earlier in this paper, which showed that the growth in college workers' wages was associated with service employment growth (table 2.7) but that growth in other workers' wages was not (table 2.8).

2.6.5 Robustness

This section presents robustness test on the calibration choices for μ and σ, which are the parameters with the biggest potential of affecting the results. We re-run model estimation for each alternative calibration. We show simulations of the main variables of interest, as well as the decomposition of service employment changes, much like what has been shown in previous sections. Figures 2.15 through 2.18 present these additional results. The main takeaway is that our results are qualitatively robust to changes in these parameters.

2.7 Conclusion

The growth in employment in low-skill service occupations is an important driver of labor market polarization and of wage inequality. The existing literature provides different views on the causes of this phenomenon and often uses regional variations to provide evidence for their proposed mechanisms. So far, it has not considered multiple explanations in the same framework. However, the two most prevalent explanations, routinization and consumption spillovers, exhibit strong spatial correlations. US regions that experience large losses in routine employment over 1980-2005 also experience stronger growth of college wages. A probable reason for this link is capital adoption intensity: regions where computers had a stronger penetration could replace more routine labor with new technologies, and also experienced shaper productivity gains among high-skill workers. This spatial correlation raises confounding concerns for studies that exploit regional variations to

identify these effects but focus on a single explanation at a time. In this paper, we fill this gap by considering these two theories together and assessing their respective importance in unified reduced-form and structural models. Our regression results show that both channels operate at the same time. It also suggests that accounting for a single channel at a time produces upward biased estimates. We develop a parismonious structural model embedding routinization and consumption spillovers to obtain a joint quantitative assessment of both theories and, as well as information on the mechanisms at play. Under our preferred specification, we find that while the routinization hypothesis dominates, consumption spillovers explain about a third of the regional disparities in service employment growth. Their contribution is thus not negligible and should be considered in future work on labor market polarization.

Our results call for developing frameworks embedding multiple mechanisms when studying labor market polarization. Our model is parsimonious. Future research would primarily benefit from enriching it with more detailed heterogeneity. This would allow a closer fit between the model and data. Second, future research could develop frameworks that encompass other channels of polarization, such as trade shocks and industry-specific changes. This would allow a more complete account of the forces at play. Other developments can inform us further about the policy side of the polarization phenomenon. For example, introducing the possibility for routine workers to get training and move up the skill distribution would allow us to understand the benefits of this policy in the face of automation. The interaction between a minimum wage and occupational shifts of labor towards the bottom of the wage distribution would provide important insight into the relevance of this policy.

2.8 Appendix

2.8.1 Data Appendix

Measuring the routine content of occupations. We use the classification found in Autor and Dorn (2013), which is based on data from the Dictionary of Occupational Titles (DoT). For each occupation in the Census classification, this dictionary provides information on the task content

of these jobs. In an earlier paper, Autor, Levy, and Murnane (2003) categorize the different tasks documented in the DoT into three categories: routine, manual and abstract. The information in the DoT allows them to compute measures of each of these tasks' intensities for each occupation in the Census. Formally, for occupation k, they obtain T_k^r, T_k^m, T_k^a as the intensity of routine, manual and abstract tasks in that occupation. They form a simple index of routine task intensity (RTI) at the occupation level through:

$$\text{RTI}_k = \log(T_k^r) - \log(T_k^m) - \log(T_k^a)$$

They finally consider that an occupation k is a "routine" occupation if its routine task intensity score RTI_k is greater than the 66th percentile of the distribution of RTI_k's:

$$\text{Occupation } k \text{ is considered routine if} \quad \text{RTI}_k > Q_{66}(\text{RTI}_j)$$

We take this classification as is, and crosswalk occupation codes provided in Autor and Dorn (2013) to codes provided in the IPUMS data.

2.8.2 Model Appendix

Derivations

Competitive equilibrium in the model is characterized by the following equations:

- Optimality conditions (location indices omitted)
  - Goods-producing firm (first order conditions of profit maximization)

$$w_a = (1-\beta) L_a^{-\beta} X^{\beta} \tag{2.8.1}$$

$$w_r = \beta\, \alpha_r L_a^{1-\beta} X^{\beta-\mu} L_r^{\mu-1} \tag{2.8.2}$$

$$p_k = \beta\, (1-\alpha_r) L_a^{1-\beta} X^{\beta-\mu} K^{\mu-1} \tag{2.8.3}$$

  - Services-producing firm

$$w_m = \alpha_s p_s \tag{2.8.4}$$

  - Noncollege workers' occupational choice

$$L_r^S = \frac{1}{\delta}\left(1 - \log(1 - L_m^S)\right)\left(1 - L_m^S\right) \tag{2.8.5}$$

- Consumer demand functions (from first order conditions of utility maximization)

$$c_{zg} = \left(1 - p_s \frac{(\theta p_s)^{-\sigma}}{1 + p_s(\theta p_s)^{-\sigma}}\right) w_z + p_s \frac{(\theta p_s)^{-\sigma}}{1 + p_s(\theta p_s)^{-\sigma}} c_0 \qquad \text{for } z = a, m \tag{2.8.6}$$

$$c_{zs} = \frac{(\theta p_s)^{-\sigma}}{1 + p_s(\theta p_s)^{-\sigma}}(w_z - c_0) \qquad \text{for } z = a, m \tag{2.8.7}$$

$$c_{rg} = \frac{1}{L_r} \int_{\eta^*}^{\infty} \left[\left(1 - p_s \frac{(\theta p_s)^{-\sigma}}{1 + p_s(\theta p_s)^{-\sigma}}\right) \eta w_r + p_s \frac{(\theta p_s)^{-\sigma}}{1 + p_s(\theta p_s)^{-\sigma}} c_0\right] \delta e^{-\delta\eta}\, d\eta \tag{2.8.8}$$

$$c_{rs} = \frac{1}{L_r} \int_{\eta^*}^{\infty} \frac{(\theta p_s)^{-\sigma}}{1 + p_s(\theta p_s)^{-\sigma}} (\eta w_r - c_0)\, \delta e^{-\delta\eta}\, d\eta \tag{2.8.9}$$

- Equilibrium conditions
  - Service output market clearing

$$\forall n: \quad L_{an} c_{asn} + L_{mn} c_{msn} + L_{rn} c_{rsn} = Y_{sn} \tag{2.8.10}$$

  - College workers' spatial equilibrium

$$\forall n > 1: \quad u(c_{agn}, c_{sgn}) = u(c_{ag1}, c_{sg1}) \tag{2.8.11}$$

  - College workers' labor market clearing

$$\sum_n L_{an} = \frac{\chi}{1 - \chi} N \tag{2.8.12}$$

Solution Algorithm

The solution algorithm finds the solution of the model given a set of parameters, through a numerical solver. To improve its performance, the model equations above are combined so that the smallest amount of work is left to the solver. Specifically, we only ask the solver to find values for w_a, w_r, w_m and L_a for each location. We consider the following ratioed FOCs from the goods firm's problem to help us pin down routine labor and capital:

$$[L_a]/[L_r]^{\lambda}: \quad \frac{1-\beta}{\beta^{\lambda}} \alpha_r^{-\lambda} L_a^{-\beta-\lambda(1-\beta)} L_r^{(1-\mu)\lambda} = \frac{w_a}{w_r^{\lambda}}$$

$$[L_a]/[K]^{\lambda}: \quad \frac{1-\beta}{\beta^{\lambda}} (1-\alpha_r)^{-\lambda} L_a^{-\beta-\lambda(1-\beta)} K^{(1-\mu)\lambda} = \frac{w_a}{p_k^{\lambda}}$$

where $\lambda = \dfrac{\beta}{\beta - \mu}$.

Conditional on guesses for w_a, w_r, w_m and L_a, we directly determine all the other model variables in the following way:

- $p_s = w_m/\alpha_s$
- $L_m = 1 - \exp(-w_m/w_r)$
- L_r through the first ratioed FOC above (depends on L_a and prices)
- K through the second ratioed FOC above (depends on L_a and prices)
- $Y_g = F(L_a, L_r, K)$ (production function)
- all consumptions are determined through their expression in the model equations presented in the derivations section 2.8.2

Conditions given to the solver to pin down the $4N$ unknowns w_{an}, w_{rn}, w_{mn} and L_{an} (one for each location) are the following $4N$ equations:

- Service output market clearing (N equations)

$$\forall n: \quad L_{an}c_{asn} + L_{mn}c_{msn} + L_{rn}c_{rsn} = Y_{sn} \tag{2.8.13}$$

- Abstract labor FOC (N equations)

$$\forall n: \quad w_{an} = (1-\beta)L_{an}^{-\beta}\left[\alpha_r L_{rn}^{\mu} + (1-\alpha_r)K_n^{\mu}\right]_n^{\beta} \tag{2.8.14}$$

- Routine labor market clearing (N equations)

$$\forall n: \quad L_{rn} = (1 + w_{mn}/w_{rn})\exp(-w_{mn}/w_{rn}) \tag{2.8.15}$$

- College workers' spatial equilibrium ($N-1$ equations)

$$\forall n > 1: \quad u(c_{agn}, c_{sgn}) = u(c_{ag1}, c_{sg1}) \tag{2.8.16}$$

- College workers' labor market clearing (1 equation)

$$\sum_n L_{an} = \frac{\chi}{1-\chi}N \tag{2.8.17}$$

The solver is constrained by bounds: L_a has to be positive and all wages have to be above c_0 to ensure utility above $-\infty$.

Estimation Algorithm

The estimation algorithm finds the set of parameters so that the model output matches selected moments in the data. The algorithm used is from the `gmm` package in R, implemented by Chaussé (2010). The procedure uses a two-step GMM (Hansen, 1982) with a heteroskedastic and auto-correlation consistent way of computing the weighting matrix (Newey and West, 1987).

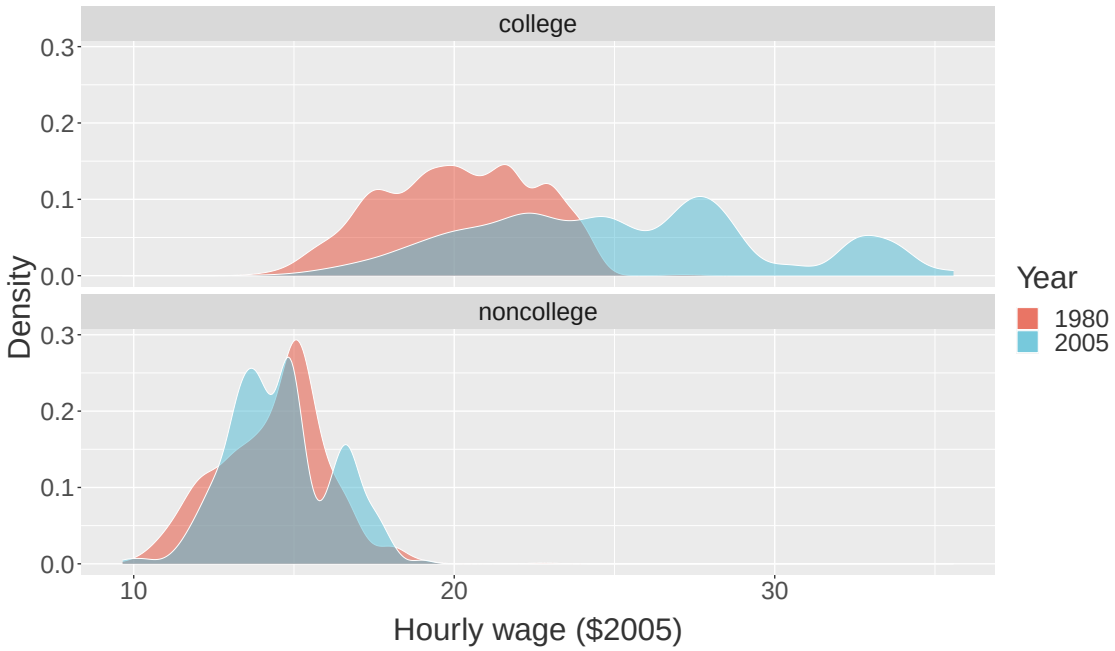
Figures

Figure 2.1: 1980 share of noncollege routine employment and 1980-2005 change in noncollege employment share of service occupations

Note: The horizontal axis represents broad occupation groups as defined by the Census Bureau. The vertical axis is expressed in percentages. All 722 commuting zones are split into two groups, depending on how their share of noncollege routine employment compares to the national mean in 1980 (weigthed by 1980 shares of total population).

Figure 2.2: Relation between the share of a household's consumption expenditures allocated to services and the hourly wage of its head (in 2005 dollars). N = 24,456. Data excludes households where the head did not work or made zero wages in the previous year. Data is binned into 46 buckets of $1 per hour, and hourly wages below $5 or above $50 are included in the first and last bins respectively. The lower bound of $5 is chosen after the $5.15 US federal minimum hourly wage in place in 2005. Bins are weighted by the sum of CEX weights of the households they represent. Share of expenditures shown are bin-specific means and are weighted by CEX household weights. The blue line represents the best linear fit across all binned points. Source: CEX 2013-2019.

Figure 2.3: Spatial distributions of college and noncollege wages in 1980 and 2005

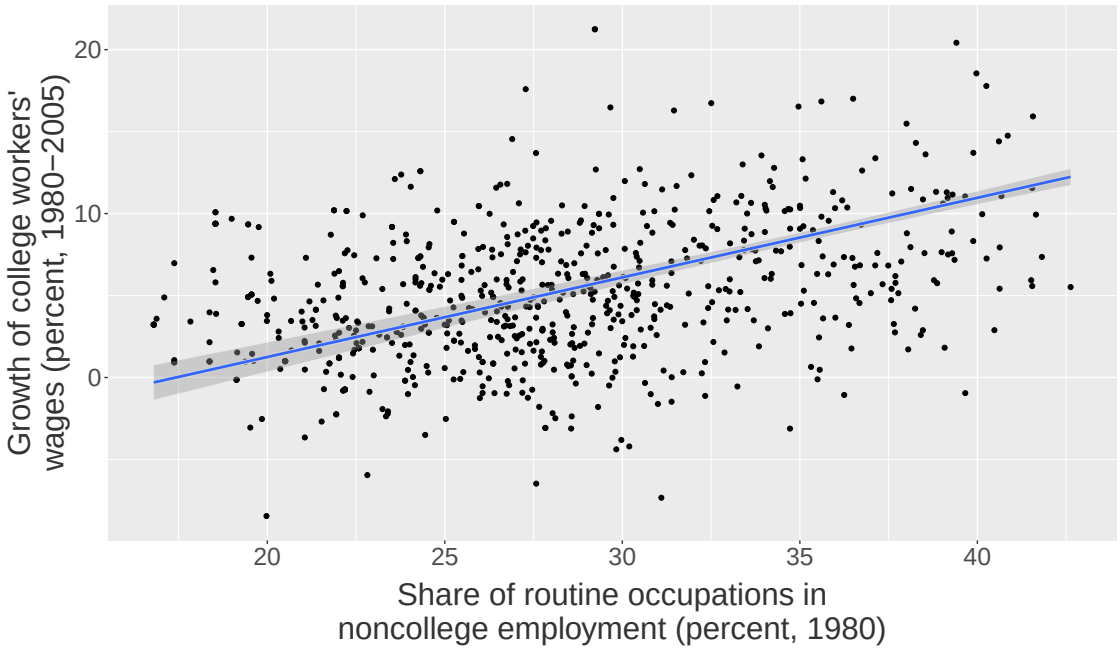


Note: Individual college and noncollege wages are averaged at the commuting zone level, using labor supply weights (yearly hours worked × individual sample weight) and deflated using the 2005 CPI. The graphs represent the spatial distributions of these average wages over commuting zones.

Figure 2.4: 1980-2005 growth in college wages contemporaneous change in noncollege employment share of service occupations

Note: The horizontal axis represents broad occupation groups as defined by the Census Bureau. The vertical axis is expressed in percentages. All 722 commuting zones are split into two groups, depending on how their 1980-2005 growth rate of average college hourly wages compares to the national mean (weigthed by 1980 shares of total population).

Figure 2.5: Relation between the determinants of the routinization and consumption spillover theories: initial share of noncollege routine employment and growth of college wages



Note: N = 722 commuting zones × 3 stacked time periods = 2166. Commuting zones are weighted by their beginning-of-period share of total population (represented by different circle sizes on the graph). The blue line is the best linear fit between the two variables.

Figure 2.6: Diagram illustrating the mechanisms of the two theories of interest: routinization and consumption spillovers

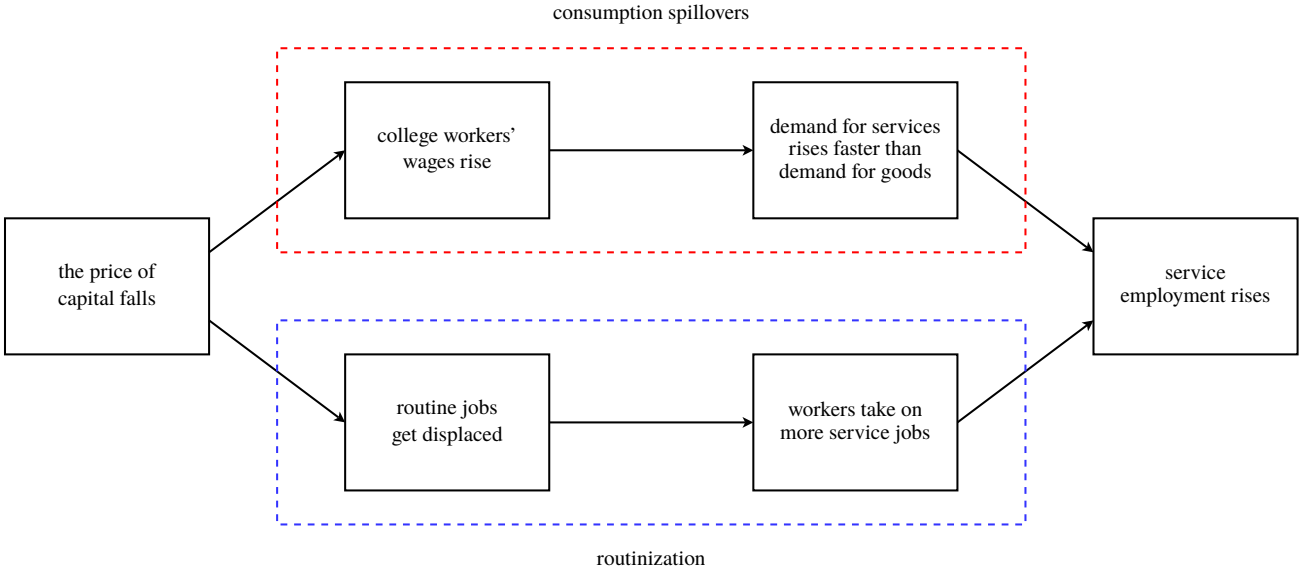


Figure 2.7: Diagram for the effect of non-homotheticity on service expenditure shares

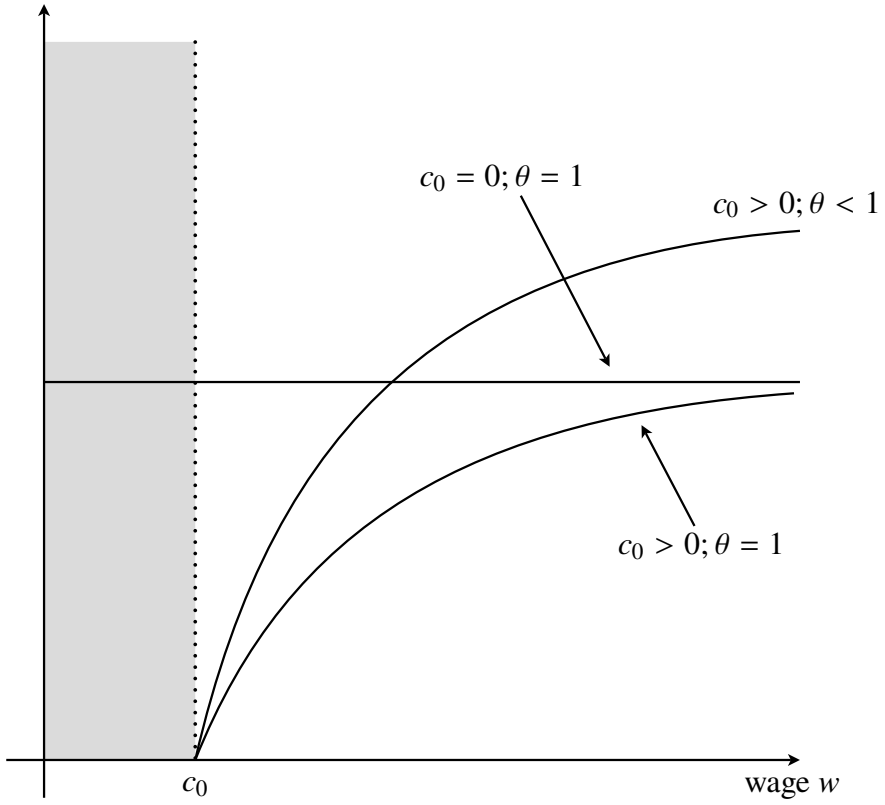
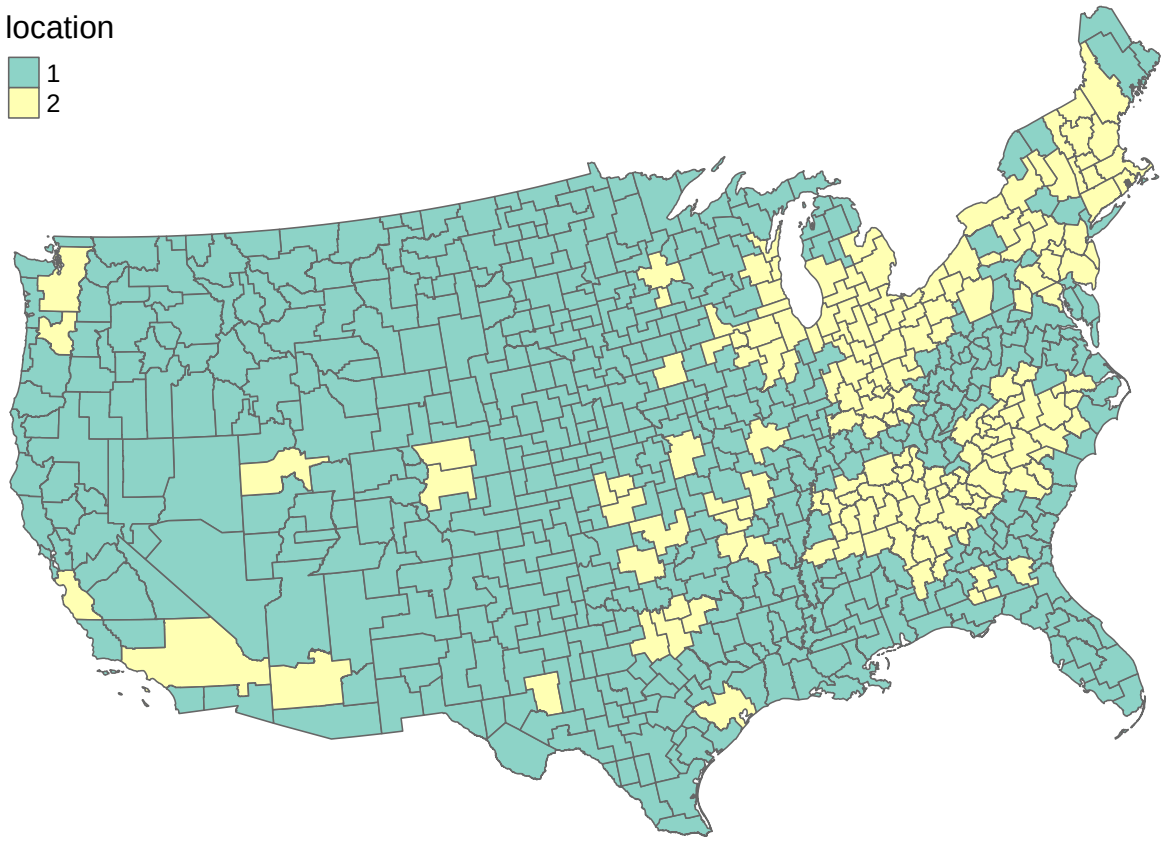
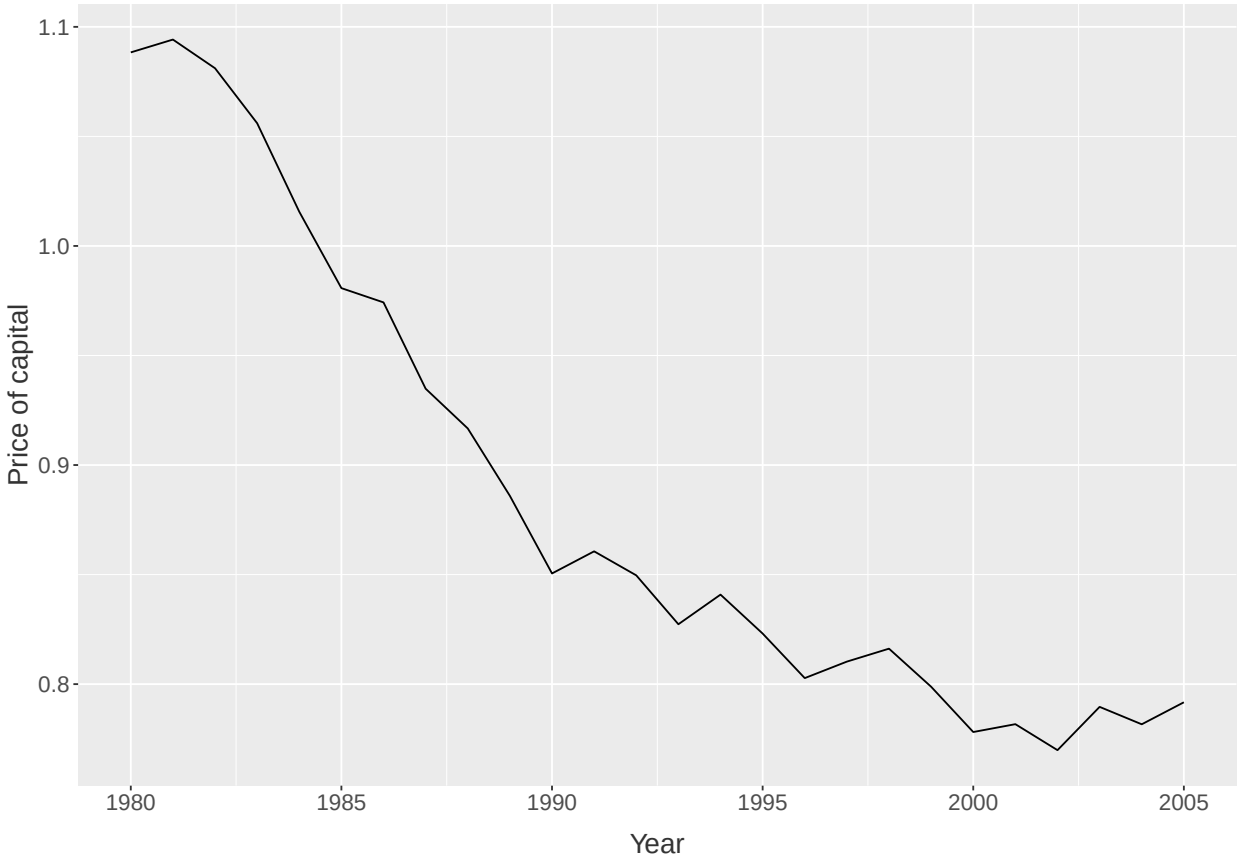


Figure 2.8: Map of the correspondence between commuting zones in the data and locations (1,2) in the model



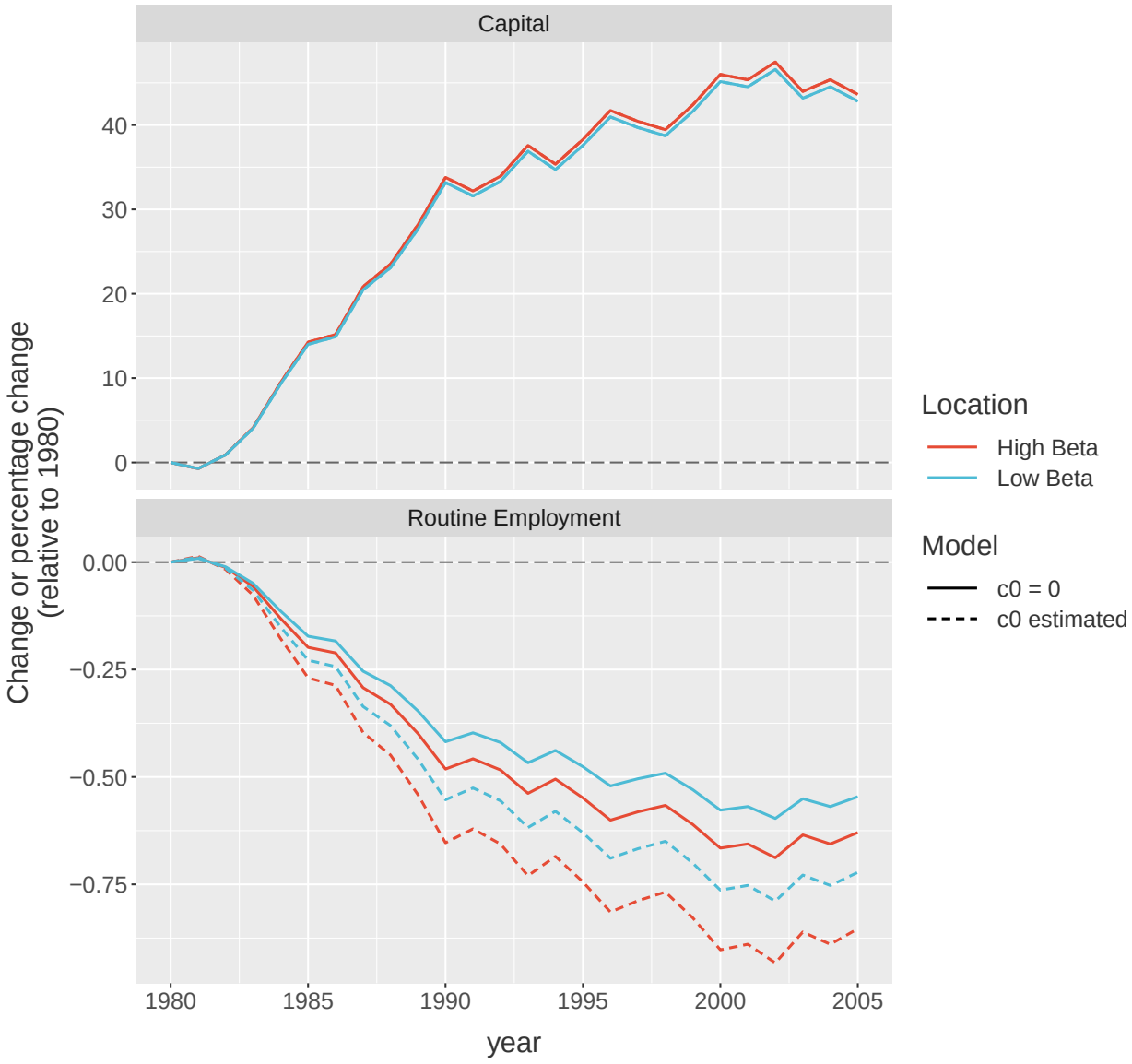
Note: Commuting zones mapped into model location 1 are those with an initial income share of college workers below the national mean (weighted by 1980 share of total population). They correspond to the location with a higher β value in the model. Location 2 consists of commuting zones with an initial income share of college workers share above the national mean (lower β).

Figure 2.9: Price of capital, 1980-2005



Note: Data obtained from the Penn World Tables, for the US over 1980-2005. The price of capital is deflated by the price of consumption goods (also obtained from PWT) to coincide with the real price of capital in the model.

Figure 2.10: Model simulations: Capital and Routine Employment



Note: The models exclude labor mobility across locations. The top panel shows the path of capital (K in the model), expressed in percentage change relative to 1980. The bottom panel shows the path of routine employment ($1 - L_m$ in the model), expressed in absolute changes relative to 1980, since it is already expressed as a share in the model.

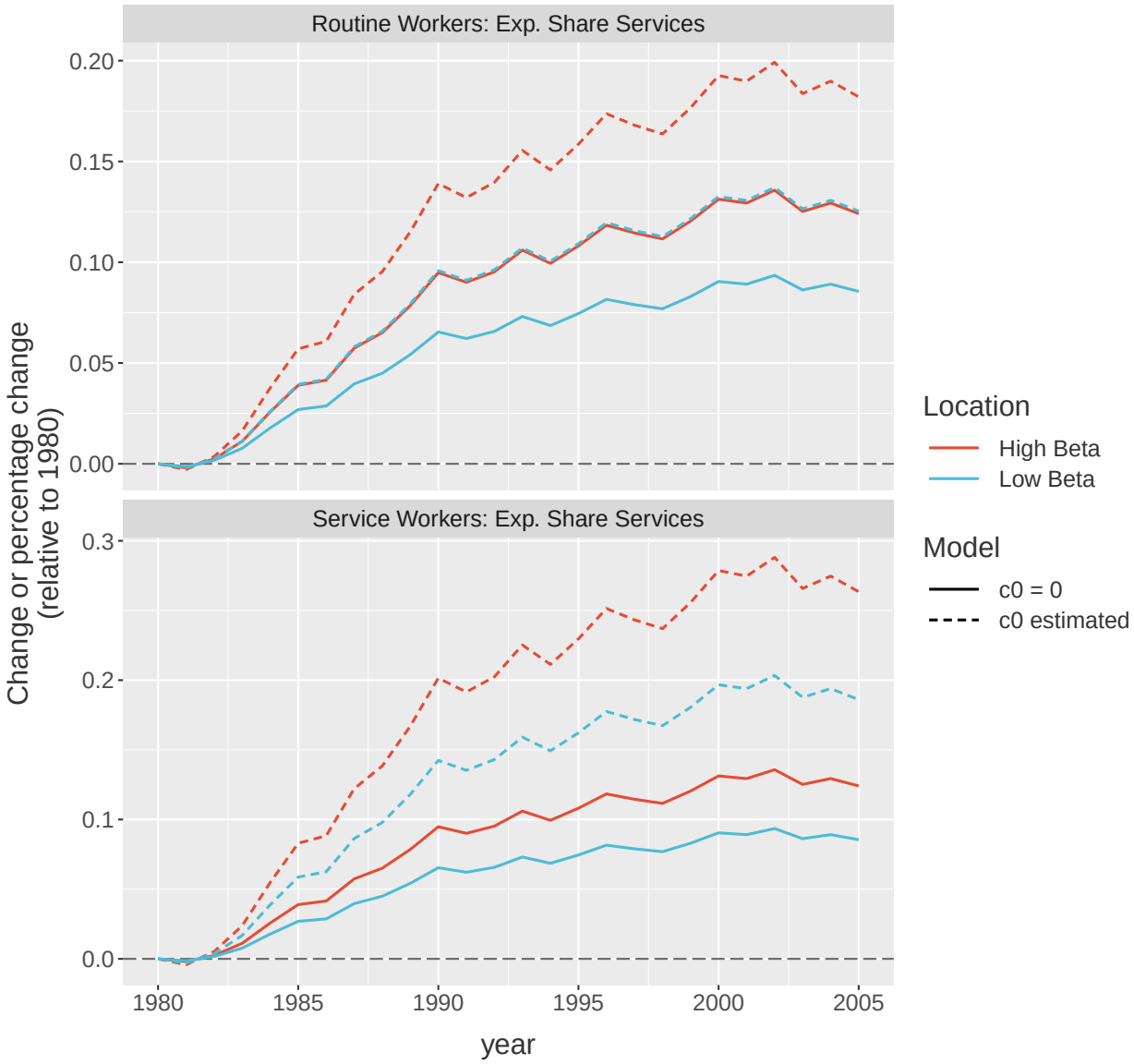
Figure 2.11: Model simulations: Service Employment and Wages

Note: The models exclude labor mobility across locations. The top panel shows the path of service employment (L_m in the model), expressed in absolute changes relative to 1980, since it is already expressed as a share in the model. The bottom panel shows the path of service workers' wages (w_m in the model), expressed in percentage change relative to 1980.

Figure 2.12: Model simulations: College Wages and College Expenditure Shares on Services

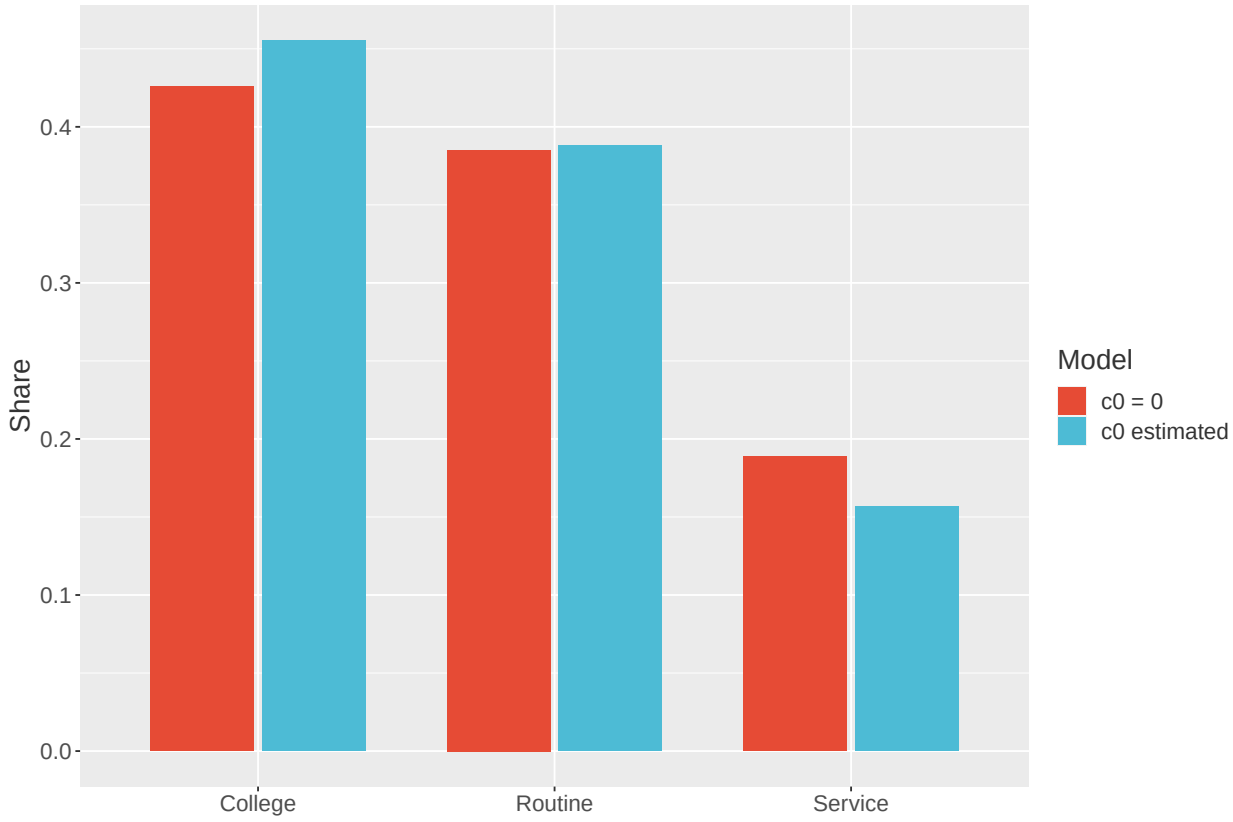
Note: The models exclude labor mobility across locations. The top panel shows the path of college wages (w_a in the model), expressed in percentage change relative to 1980. The bottom panel shows the path of the share of expenditures of college workers use to purchase services ($x_{as} \equiv p_s c_{as}/(p_s c_{as} + c_{ag})$ in the model), expressed in absolute changes relative to 1980.

Figure 2.13: Model simulations: Routine and Service workers' Expenditure Shares on Services



Note: The models exclude labor mobility across locations. The top panel shows the path of the share of expenditures of routine workers use to purchase services (x_{rs} in the model), expressed in absolute changes relative to 1980. Since routine workers are heterogeneous, their expenditure shares are constructed using aggregate consumption: $x_{rs} \equiv p_s C_{rs}/(p_s C_{rs} + C_{rg})$. The bottom panel shows the path of the share of expenditures of service workers use to purchase services ($x_{ms} \equiv p_s c_{ms}/(p_s c_{ms} + c_{mg})$ in the model), expressed in absolute changes relative to 1980.

Figure 2.14: Model simulations: Contributions of worker types to increases in service consumption and employment



Note: Models exclude labor mobility across locations.

Figure 2.15: Model robustness: simulations and effect decomposition with a lower production elasticity $\mu = 0.15$

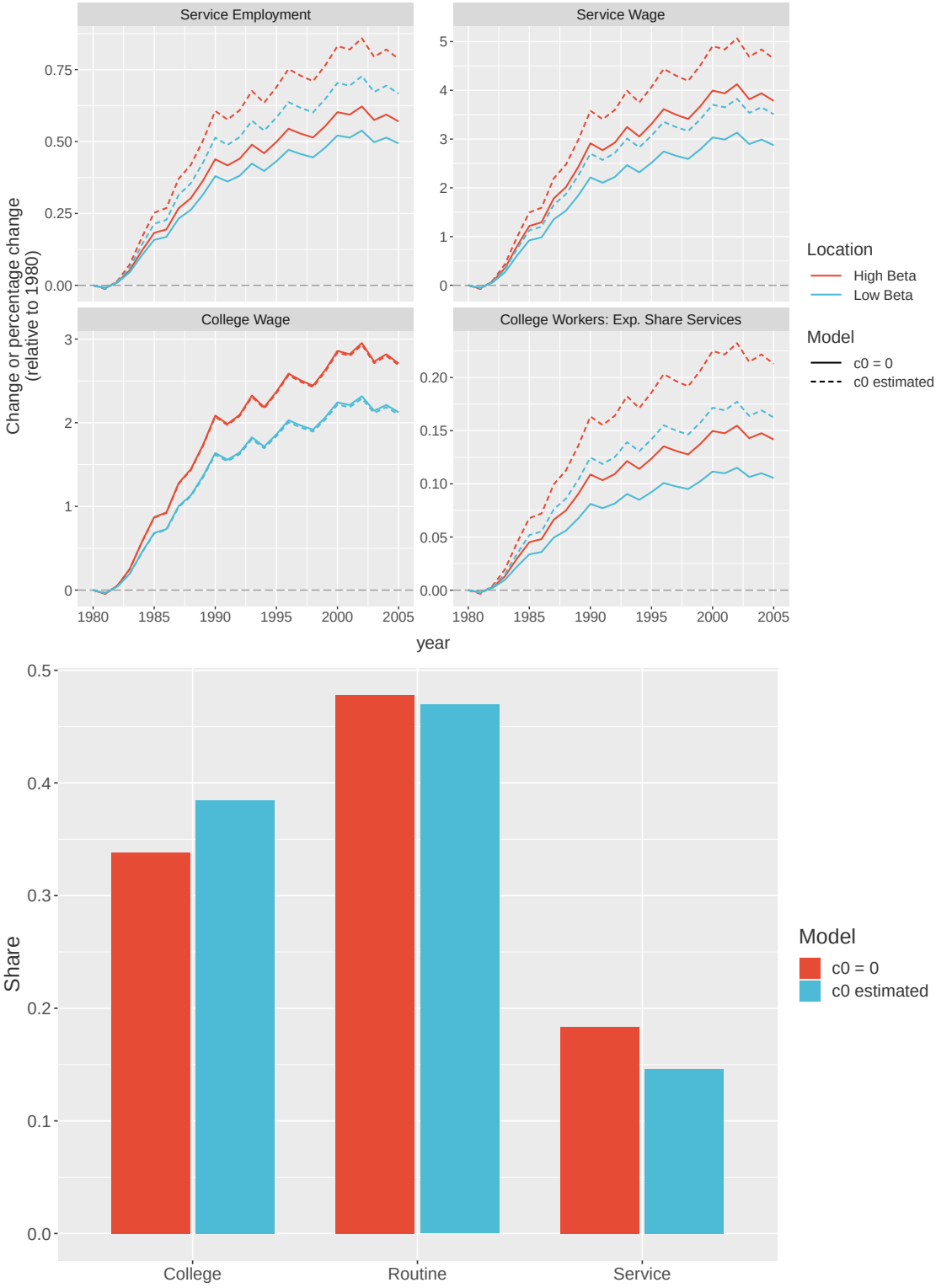


Figure 2.16: Model robustness: simulations and effect decomposition with a higher production elasticity $\mu = 0.35$

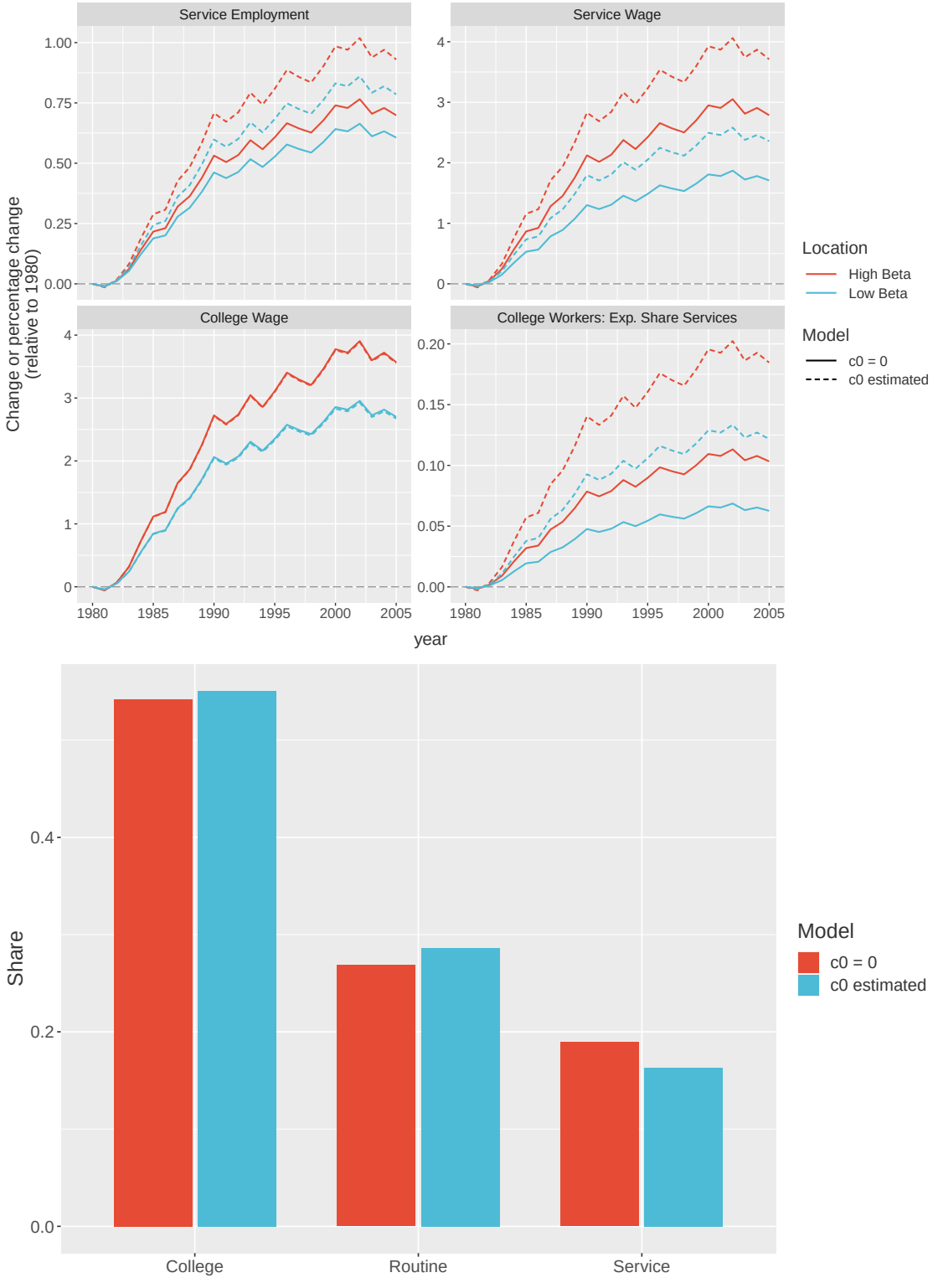


Figure 2.17: Model robustness: simulations and effect decomposition with a lower consumption elasticity $\sigma = 0.15$

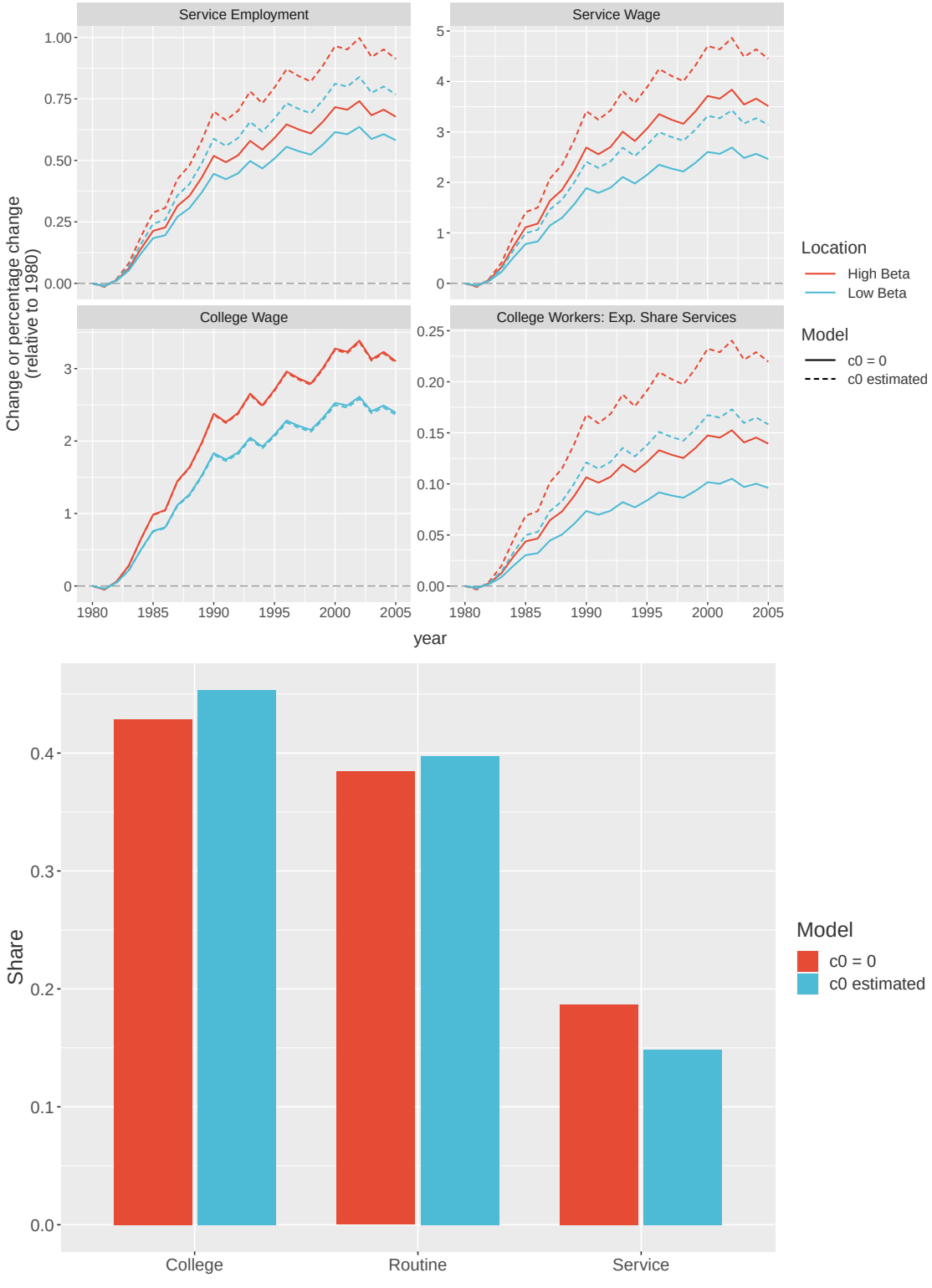
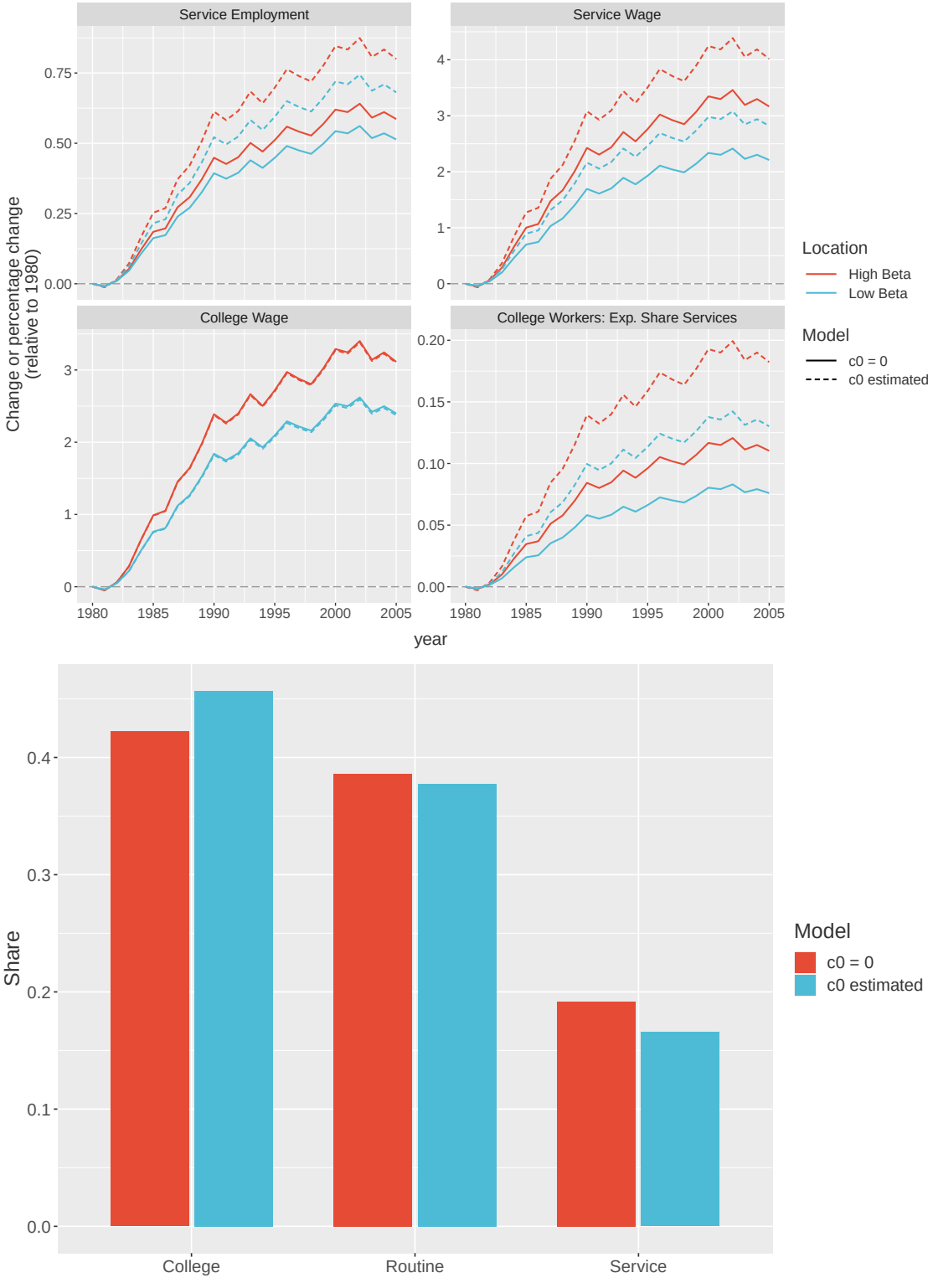


Figure 2.18: Model robustness: simulations and effect decomposition with a higher consumption elasticity $\sigma = 0.25$



Tables

Table 2.1: Top 10 service occupations by their 1980 employment share within services.

	Employment share within service occs.	
Occupation	1980	2005
Janitors	18.97	13.09
Health and nursing aides	13.72	15.38
Cooks	11.41	12.36
Waiters and waitresses	9.18	6.53
Housekeepers, maids, butlers, and cleaners	8.69	7.27
Misc. food preparation and service workers	6.46	5.83
Hairdressers and cosmetologists	4.92	4.64
Guards and police, except public	4.37	3.97
Child care workers	4.04	5.21
Gardeners and groundskeepers	3.05	6.89
Total	84.81	81.18

Note: Employment expressed in total hours workers, calculated as the product of a workers' usual weekly hours, weeks worked and sample weight. Source: Census 1980 and American Community Survey 2005.

Table 2.2: Top 10 routine occupations by their 1980 employment share, within routine occupations.

	Employment share within routine occs.	
Occupation	1980	2005
Secretaries and stenographers	12.38	8.90
Bookkeepers, accounting and auditing clerks	5.59	3.47
Assemblers of electrical equipment	5.26	3.45
General office clerks	4.72	2.37
Cashiers	3.91	4.93
Accountants and auditors	3.56	5.34
Production checkers, graders, and sorters	3.24	2.00
Office supervisors	2.94	4.57
Shipping and receiving clerks	2.29	1.61
Real estate sales occupations	2.19	2.61
Total	50.80	44.65

Note: Employment expressed in total hours workers, calculated as the product of a workers' usual weekly hours, weeks worked and sample weight. Source: Census 1980 and American Community Survey 2005.

Table 2.3: Expenditure categories in the CEX 2013-2019 classified as service expenditures.

Expenditure category	Examples
Repair services	Repair of tv, radio, and sound equipment Reupholstering, furniture repair
Babysitting	Babysitting and child care
Housekeeping services	Gardening, lawn care service Moving, storage, freight
Fitness and entertainment	Social, recreation, health club membership Play, theater, opera, concert
Food and drink away from home	Dinner at full service restaurants Dinner at fast food, take-out, delivery, ...

Note: Classification taken from Mazzolari and Ragusa (2013).

Table 2.4: Employment and wage data for workers without a college degree, by broad occupation groups in 1980 and 2005.

	Employment Share (%)			Hourly Wages		
Occupation Group	1980	2005	Change	1980	2005	Growth (%)
Service	13.1	19.1	6.0	9.6	10.2	5.8
Admin / Retail	24.2	22.7	-1.5	12.5	13.6	8.8
Operators	14.5	7.9	-6.6	14.0	13.6	-2.6
Transportation / Construction	25.8	28.0	2.2	16.2	15.2	-6.4
Management / Professional / Technical	16.0	17.8	1.8	17.1	19.6	13.6
Production	6.3	4.5	-1.8	17.7	15.9	-10.9

Note: Employment is calculated at total hours worked (product of weekly hours, weeks worked and individual sample weight). Hourly wages are calculated as annual wages divided by the product of weekly hours and weeks worked, and they exclude self-employed workers. Occupation groups follow standard Census definitions. Source: Census 1980 and American Community Survey 2005.

Table 2.5: Effect of personal computer adoption in businesses on college wages (column 1) and routine employment (column 2).

	College Wages	Routine Employment
Computer Penetration	0.402^{***}	-0.112^{***}
	(0.070)	(0.020)
Fixed Effects	Division × Year	Division × Year
Adj. R^2	0.576	0.601
Num. obs.	1335	1335

Note: Models are estimated on commuting zones for which computer penetration is available, and two time periods: 1980-1990 and 1990-2000. Computer penetration measure is from Beaudry, Doms, and Lewis (2006), aggregated at commuting zone level by Autor and Dorn (2013). College wages are expressed as mean college hourly wages within a commuting zone, computed by dividing annual wages by the product of weekly hours and weeks worked. Routine employment is expressed as total hours worked. Robust standard errors are clustered at the Census Division level. All models are weighted by beginning-of-period commuting zone share of population. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Source for college wages and routine employment: Census/ACS.

Table 2.6: Share of routine employment by broad occupational groups.

Occupation group	Share of routine occupations, 1980 (%)
Admin / Retail	74.62
Operators	50.10
Production	34.52
Service	27.55
Management / Professional / Technical	18.15
Transportation / Construction	8.44

Note: Employment is computed as total hours worked. Source: Census 1980.

Table 2.7: Regression estimates: Initial Routine Employment Share, College Wage Growth and Growth of Service Employment

	(1)	(2)	(3)	(4)	(5)
Initial share of routine occs	0.106***		0.092***	0.063**	0.081***
	(0.017)		(0.016)	(0.028)	(0.017)
Hourly wage growth: college		0.042***	0.026**	0.031**	0.018
		(0.013)	(0.013)	(0.014)	(0.012)
Unemployment rate				−0.438***	0.512***
				(0.084)	(0.108)
Share 65+				−0.107***	0.259***
				(0.020)	(0.068)
Share below min. wage				−0.046	
				(0.043)	
Share foreign born				0.025***	0.097***
				(0.008)	(0.029)
Share female emp.				−0.147***	0.134***
				(0.017)	(0.028)
Urban CZ				0.005**	
				(0.002)	
Share college workers					−0.020
					(0.046)
Fixed Effects	Division × Year	Division × Year	Division × Year	Division × Year	Division × Year
Controls	No	No	No	in levels	in changes
Adj. R^2	0.306	0.295	0.310	0.369	0.364
Num. obs.	2166	2166	2166	2166	2166

Note: Dependent variable: 10 × annual change in share of noncollege employment in service occupations. $N = 722$ commuting zones × 3 stacked time periods = 2166 in all models. Columns 1-3 do not include additional controls. Column 4 includes controls measured in levels at the beginning of each period. Column 5 includes controls measured in contemporaneous changes over each period. Robust standard errors are clustered at the Census Division level. All models are weighted by beginning-of-period commuting zone share of population. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Source for employment and wage data: Census/ACS.

Table 2.8: Regression estimates: Initial Routine Employment Share, Other Noncollege Wage Growth and Growth of Service Employment

	(1)	(2)	(3)	(4)	(5)
Initial share of routine occs	0.106***		0.106***	0.066**	0.084***
	(0.017)		(0.016)	(0.029)	(0.017)
Hourly wage growth: other NC		0.004	−0.001	0.009	0.012
		(0.017)	(0.016)	(0.016)	(0.014)
Unemployment rate				−0.445***	0.523***
				(0.085)	(0.112)
Share 65+				−0.108***	0.262***
				(0.021)	(0.071)
Share below min. wage				−0.044	
				(0.050)	
Share foreign born				0.027***	0.105***
				(0.009)	(0.028)
Share female emp.				−0.137***	0.133***
				(0.018)	(0.028)
Urban CZ				0.005**	
				(0.002)	
Share college workers					−0.011
					(0.044)
Fixed Effects	Division × Year	Division × Year	Division × Year	Division × Year	Division × Year
Controls	No	No	No	in levels	in changes
Adj. R^2	0.306	0.283	0.306	0.364	0.363
Num. obs.	2166	2166	2166	2166	2166

Note: Dependent variable: 10 × annual change in share of noncollege employment in service occupations. $N = 722$ commuting zones × 3 stacked time periods = 2166 in all models. Columns 1-3 do not include additional controls. Column 4 includes controls measured in levels at the beginning of each period. Column 5 includes controls measured in contemporaneous changes over each period. Robust standard errors are clustered at the Census Division level. All models are weighted by beginning-of-period commuting zone share of population. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Source for employment and wage data: Census/ACS.

Table 2.9: Regression estimates: Routine Employment Growth, College Wage Growth and Growth of Service Employment

	(1)	(2)	(3)	(4)	(5)
Hourly wage growth: college		0.042***	0.042***	0.038***	0.030**
		(0.013)	(0.011)	(0.011)	(0.012)
Routine employment growth	−0.039***		−0.039***	−0.039***	−0.046***
	(0.007)		(0.007)	(0.006)	(0.006)
Unemployment rate				−0.376***	0.382***
				(0.085)	(0.101)
Share 65+				−0.090***	0.075
				(0.015)	(0.072)
Share below min. wage				−0.011	
				(0.039)	
Share foreign born				0.032***	0.123***
				(0.009)	(0.026)
Share female emp.				−0.128***	0.205***
				(0.019)	(0.032)
Urban CZ				0.006***	
				(0.002)	
Share college workers					−0.107**
					(0.050)
Fixed Effects	Division × Year	Division × Year	Division × Year	Division × Year	Division × Year
Controls	No	No	No	in levels	in changes
Adj. R^2	0.335	0.295	0.346	0.412	0.403
Num. obs.	2166	2166	2166	2166	2166

Note: Dependent variable: 10 × annual change in share of noncollege employment in service occupations. $N = 722$ commuting zones × 3 stacked time periods = 2166 in all models. Columns 1-3 do not include additional controls. Column 4 includes controls measured in levels at the beginning of each period. Column 5 includes controls measured in contemporaneous changes over each period. Robust standard errors are clustered at the Census Division level. All models are weighted by beginning-of-period commuting zone share of population. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Source for employment and wage data: Census/ACS.

Table 2.10: Calibrated parameters for all models

Parameter	Description	Calibrated Value	Source
μ	Production elasticity	0.25	Krusell et al. (2000)
σ	Consumption elasticity	0.2	Barany and Siegel (2020)
α_s	Service production efficiency	1	θ is sufficient
δ	Routine skill distribution	1	Autor and Dorn (2013)
χ	College share	0.34	$\chi/(1-\chi) = 0.51$
path of p_k	Path of capital price 1980-2005	see figure 2.9	Penn World Tables

Note: All calibrated parameters are constant across locations. Robustness checks on the calibration choices for μ and σ are provided in section 2.6.5.

Table 2.11: GMM estimates for models with homothetic and non-homothetic preferences, without labor mobility

Parameter	Location	Estimate	Standard Error
Homothetic preferences:			
β	1	0.29^{***}	0.005
β	2	0.35^{***}	0.005
α_r	all	0.72^{***}	0.009
θ	all	2865.07^{***}	150.23
Non-homothetic preferences:			
β	1	0.28^{***}	0.005
β	2	0.35^{***}	0.005
α_r	all	0.72^{***}	0.009
θ	all	676.42^{***}	195.05
c_0	all	0.16^{***}	0.03

Bibliography

Acemoglu, Daron (1999). "Changes in Unemployment and Wage Inequality: An Alternative Theory and Some Evidence". In: *Am. Econ. Rev.* 89.5, pp. 1259–1278.

Autor, David H (2019). "Work of the Past, Work of the Future". In: *AEA Papers and Proceedings* 109, pp. 1–32.

Autor, David H and David Dorn (2013). "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market". In: *Am. Econ. Rev.* 103.5, pp. 1553–1597.

Autor, David H, David Dorn, and Gordon H Hanson (2013). "The China Syndrome: Local Labor Market Effects of Import Competition in the United States". In: *Am. Econ. Rev.* 103.6, pp. 2121–2168.

– (2015). "Untangling trade and technology: Evidence from local labour markets". en. In: *Econ. J.* 125.584, pp. 621–646.

Autor, David H, Frank Levy, and Richard J Murnane (2003). "The Skill Content of Recent Technological Change: An Empirical Exploration". en. In: *Q. J. Econ.* 118.4, pp. 1279–1333.

Barany, Zsofia L and Christian Siegel (2020). "Biased Technological Change and Employment Reallocation". In: *Labor Economics*.

Beaudry, Paul, Mark Doms, and Ethan Lewis (2006). "Endogenous Skill Bias in Technology Adoption: City-Level Evidence from the IT Revolution".

– (2010). "Should the Personal Computer Be Considered a Technological Revolution? Evidence from U.S. Metropolitan Areas". In: *J. Polit. Econ.* 118.5, pp. 988–1036.

Becker, Gary S (1965). “A Theory of the Allocation of Time”. In: *Econ. J. Nepal* 75.299, pp. 493–517.

Boppart, Timo (2014). “Structural change and the kaldor facts in a growth model with relative price effects and non-Gorman preferences”. en. In: *Econometrica* 82.6, pp. 2167–2196.

Card, David and Alan Krueger (1994). “Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania”. In: *American Economic Review* 84.4, pp. 772–93. URL: https://EconPapers.repec.org/RePEc:aea:aecrev:v:84:y:1994:i:4:p:772-93.

Casey, Marcus (2020). *Looking towards the future: Automation, training, and the middle class*. https://www.brookings.edu/blog/up-front/2020/07/22/looking-towards-the-future-automation-training-and-the-middle-class/. Accessed: 2021-11-7.

Chaussé, Pierre (2010). “Computing Generalized Method of Moments and Generalized Empirical Likelihood with R”. In: *Journal of Statistical Software* 34.11, pp. 1–35. DOI: 10.18637/jss.v034.i11. URL: https://www.jstatsoft.org/index.php/jss/article/view/v034i11.

Comin, Diego, Danial Lashkari, and Mart Mestieri (2021). “Structural change with LongRun income and price effects”. en. In: *Econometrica* 89.1, pp. 311–374.

Comin, Diego, Mart Mestieri, and Ana Danieli (2020). “Income-Driven Labor Market Polarization”.

David Dorn (2021). https://www.ddorn.net/data.htm. Accessed: 2021-11-4.

Doms, Mark, Timothy Dunne, and Kenneth R Troske (1997). “Workers, Wages, and Technology”. In: *Q. J. Econ.* 112.1, pp. 253–290.

Goos, Maarten and Alan Manning (2007). “Lousy and lovely jobs: The rising polarization of work in Britain”. en. In: *Rev. Econ. Stat.* 89.1, pp. 118–133.

Goos, Maarten, Alan Manning, and Anna Salomons (2014). “Explaining Job Polarization: Routine-Biased Technological Change and Offshoring”. In: *Am. Econ. Rev.* 104.8, pp. 2509–2526.

Hansen, Lars Peter (1982). “Large Sample Properties of Generalized Method of Moments Estimators”. In: *Econometrica* 50.4, pp. 1029–1054.

Heston, Alan, Robert Summers, and Bettina Aten (2012). *Penn World Table Version 7.1*. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania. URL: http://pwt.econ.upenn.edu/.

Houseman, Susan N (2018). *Understanding the Decline of U.S. Manufacturing Employment*. Tech. rep.

Jardim, Ekaterina et al. (forthcoming). “Minimum Wage Increases and Low-Wage Employment: Evidence from Seattle”. In: *American Economic Journal: Economic Policy*.

Krusell, Per et al. (2000). “Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis”. In: *Econometrica* 68.5, pp. 1029–1053.

Lee, Keyoung (2019). “Benefits of Working Near College Workers: Human Capital Spillovers Across Neighborhoods”. PhD thesis. UCLA.

Liu, Shimeng and Xi Yang (2021). “Human capital externalities or consumption spillovers? The effect of high-skill human capital across low-skill labor markets”. en. In: *Reg. Sci. Urban Econ.* 87.103620, p. 103620.

Manning, Alan (2004). “We can work it out: The impact of technological change on the demand for low-skill workers”. en. In: *Scott. J. Polit. Econ.* 51.5, pp. 581–608.

Mazzolari, Francesca and Giuseppe Ragusa (2013). “Spillovers from High-Skill Consumption to Low-Skill Labor Markets”. In: *Rev. Econ. Stat.* 95.1, pp. 74–86.

Michaels, Guy, Ashwini Natraj, and John Van Reenen (2014). “Has ICT polarized skill demand? Evidence from eleven countries over twenty-five years”. en. In: *Rev. Econ. Stat.* 96.1, pp. 60–77.

Moretti, Enrico (2004). “Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data”. In: *J. Econom.* 121.1, pp. 175–212.

Nedelkoska, Ljubica and Glenda Quintini (2018). *Automation, skills use and training*. Tech. rep.

Newey, Whitney K and Kenneth D West (1987). "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix". In: *Econometrica* 55.3, pp. 703–708.

Ruggles, Steven et al. (2020). *IPUMS USA: Version 10.0*.

Tolbert, Charles M and Molly Sizer (1996). *U.S. Commuting Zones and Labor Market Areas: A 1990 Update*. Tech. rep. 278812.

United States Employment Service (1977). *Dictionary of Occupational Titles*. The Administration. URL: https://books.google.com/books?id=NWL0ilZKvg8C.

Chapter 3

Deregulated Temporary Contracts and Workers' Outcomes: Evidence from a 2003 French Reform

3.1 Introduction

This paper examines the impact of a 2003 judicial decision on labor market outcomes. This decision — hereafter called the "reform" — liberalized the use of a specific type of temporary contract in France, the CDDU (*contrat à durée déterminée d'usage*). This contract is allowed in 16 service sectors whose activities are in good part temporary by nature (such as, e.g., hotels). CDDU sectors are very diverse, ranging from tourism to entertainment and healthcare services. The CDDU gives the opportunity for the employer to hire on temporary contracts in a very unrestricted way as compared to the classic temporary contract (CDD, *contrat à durée déterminée*). This standard form of temporary work entails several major restrictions for the employer: the total duration and the number of renewals of a CDD are strongly limited; after the end of a CDD, the employer cannot hire a new CDD at the same position for a significant window period; at the end of a CDD, the employer must pay a severance payment equal to 10% of all paid wages to the employee. The CDDU is exempt from all these restrictions, allowing the employer to hire temporarily, but for whatever length and number of renewals, while not having to pay any severance payment at the end.

Prior to 2003, the hiring under a CDDU is conditioned to essentially the same rules as a CDD: the

job that is filled has to be temporary by nature. In November 2003, the French Civil Supreme Court (*Cour de Cassation*) rules in favor of employers in several cases where employees complain about being abusively hired under a CDDU to perform tasks that pertain to the regular activity of their company, and are hence not temporary. More specifically, the court does not retain the "temporary need" requirement that used to prevail in order to use a CDDU contract. In essence, the court starts allowing employers in CDDU-authorized sectors to use unrestricted temporary contracts for needs that are not momentary by nature. This notably creates an attractive alternative to the highly protected French permanent contract (CDI, *contrat à durée indéterminée*). I use this policy change as a natural experiment to study the impact of extending employers' access to temporary contracts on workers' outcomes.

I use data from the French Employment Survey from the French National Statistical Institute (INSEE). I use this data as a yearly repeated cross-section of the French population and I restrict myself to individuals employed in service sectors, which include all CDDU sectors. My identification strategy lies in the fact that prior to 2003, CDDU sectors show similar characteristics and trends as other service sectors, notably in terms of contract type distribution and wages. After 2003, I observe significant changes in CDDU sectors while other service sectors are much less affected. I interpret these variations as being a consequence of the 2003 reforms. I provide several robustness checks, notably by grouping CDDU and other services sectors in more coherent subcategories, and I can still observe the differences found in aggregate.

I find that this reform is associated with two main effects. First, CDDU employers start relying more on temporary contracts, benefiting from the opening of CDDUs to non-momentary needs. This kind of effect has already been documented in the literature on employment protection and temporary contracts (see below for a review). I find that the share of CDDs in CDDU sectors increases from an average of 12% over 1998-2004 to 14.9% over 2005-2008. Over the same time, the share of CDDs in other service sector's goes from 6.6% to 7.3%, only about 25% of the relative increase in CDDU sectors. This phenomenon especially makes sense as opening the CDDU to non temporary jobs enables employers to circumvent the restrictive employment protection provisions that come with a CDI. This finding is shown to be robust across CDDU sectors: it is a pattern common to most of these sectors rather than being driven by only a few of them, which could have

otherwise been the result of a confounding phenomenon.

The second effect that I find is novel. While most of the literature focuses on the fact that substituting CDDs for CDIs entails more temporary workers that are typically less paid, trained and satisfied with their job, I find evidence that permanent workers' wages are also impacted by the expanding access to temporary contracts. More specifically, I find that CDDU-sector young permanent workers experience a fall in wages of about 2% per year over 2004 and 2005. This is striking as there is normally very strong downward wage rigidity in France. From 2006 on, their wages start growing at their pre-2003 rate again (roughly the same growth rate as in other service sectors) but they do not catch up to their pre-2003 level, at least not until 2008. I show that these findings are also consistent across CDDU sectors and not the result of a single of few sector-specific variations. This finer sectoral decomposition shows even more striking results, where some CDDU sectors experience a prolonged change in the growth rate of their wages as compared to other related service sectors. Turning to the source of these effects, I find some preliminary evidence that the relationship between employers and employees has changed in CDDU sectors because of the 2003 reform. More specifically, I run Oaxaca decompositions to split the difference in wages between CDDU and other service sectors into a composition component (occupational type or educational level) and a within-group return component. I find that the return to education/occupation explains the most part of the post-2003 increase in the wage gap between these sectors. This is suggestive that the same workers are now treated differently in CDDU sectors. The opening of the CDDU contract's availability for non-temporary jobs seems to have decreased the attractiveness of the CDI contract for CDDU-sector employers, as well as having given them the actual possibility (and a credible threat) not to offer CDIs while keeping their employee indefinitely. They are now less inclined to offer CDIs, which weighs on the wages of young workers who try to obtain them.

This article relates most to the extensive literature studying the effects of tightening or widening the employment protection gap between temporary and permanent contracts. While most articles have, in theory or in the data, studied variations in the actual levels of protection of either temporary or permanent contracts, my research exploits a increase in the scope of jobs that temporary contracts can be used for. One can however argue that the 2003 French reform puts more workers at risk of being hired under an unprotected contract (namely, the employer starts to use more CDDUs

in place of CDIs) and is thus expected to have effects that are correlated with the ones coming from changes in the employment protection gap. In theory, there are ambiguities around the effects that employment protection reforms can have. The most well-known ambiguity is the one surrounding aggregate employment effects of changes in permanent jobs' employment protection: while reducing the stringency of employment protection increases the incentives for firms to hire, it also makes it directly less costly to fire (S. Bentolila and Bertola, 1990; Bertola, 1990; Bertola, 1994; Saint-Paul and Bentolila, 1994). A similar ambiguity arises around reforms making it easier to use temporary contracts (Blanchard and Landier, 2002; Cahuc and Postel-Vinay, 2002): firms may hire more because of reduced costs of labor, but there are also more terminations of temporary contracts, making the net effect on employment ambiguous.

On the other side, the empirical literature has either used variations (often relaxations) in the regulatory framework around temporary contracts, or variations in the stringency of employment protection for permanent contracts. Most of these articles focus on one country, most often in Europe. Some exploit targeted reforms, whereby a specific group receives a different treatment, and where the identification of effects is enabled through the existence of a natural control group; others exploit country-wide reforms and show that workers' outcomes correlate with the expected impacts of the policy changes. There is substantial evidence that facilitating the access to or relaxing the rules of temporary contracts increases labor turnover (Blanchard and Landier, 2002; Kugler, Jimeno, and Hernanz, 2005; Kugler and Pica, 2008; Centeno and Novo, 2012). This is mostly due to the substitution of temporary contracts for permanent ones, which mechanically increases terminations (programmed in advanced for temp workers) and encourages rehires, instead of having continuous employment phases. The temporary for permanent substitution phenomenon has received international evidence too (Kahn, 2010). However, some of this evidence has been questioned by Bassanini and Garnero, 2013 who find no significant evidence of increased turnover when employment protection of permanent contract becomes more stringent. Blanchard and Landier, 2002 find that following the expansion of temporary contracts in France over the 1980's and 1990's, young workers also suffered in three other dimensions: the probability to become unemployed after a temporary contract increased over that period; the probability to find a permanent contract after a temporary contract decreased; wages decreased. Politicians behind reforms that make it easier to

hire temp workers often argue that this will raise employment, but these effects appear to only be short-lived (S. Bentolila and G. Saint-Paul, 1992; Boeri and Garibaldi, 2007) and they are often followed by a reduction in aggregate productivity due to the lack of human capital investment into new temporary workers (Boeri and Garibaldi, 2007). The literature has thus rather focused on the direct effect of having more temporary (which are also poorer, more precarious) workers. This paper shows that permanent workers can also suffer from the deregulation of temporary jobs. In the face of numerous negative results surrounding the expansion or deregulation of temporary contracts, some economists started arguing against their existence altogeether, advocating for a unique contract that could fulfill both the required employment protection needs while giving enough flexibility to firms (Cahuc, 2012; Ignacio García Pérez and Osuna, 2014).

The paper is organized as follows. Section 3.2 provides some elements of context about the French labor contract system and a detailed description of the 2003 reform. Section 3.3 describes the data. Section 3.4 presents the main findings. Section 3.5 briefly concludes and opens to future research.

3.2 Context

3.2.1 Overview of the French labor contract system

In France, in the service sector, there are essentially two kinds of labor contracts: the CDI and the CDD. The CDI (*contrat à durée indéterminée*) is an open-ended contract that offers substantial protection to the worker against termination. Dismissal costs are imposed whenever an employee is fired for a reason other than serious misconduct. These dismissal costs increase with job tenure. Yet, the legal minimum increase is only about 1/10 of the most recent monthly wage per year of tenure. However, direct individual firing costs are not what makes it costly and dissuasive for firms to fire permanent workers (Abowd and Kramarz (2003) and Kramarz and Michaud (2010)). The main source of costs is either coming from collective dismissals (more than 10 terminations within 30 days) which entail a specific process whereby the firm has to care about finding the terminated workers a new job (*plan social*); or from the possibility for the dismissed worker(s) to go to court and request significant severance payments. Nowadays, over 90% of legal proceedings at the Labor

Affairs Court (*Conseil des Prud'hommes*) are related to the termination of a CDI contract, and the verdict is in favor of the plaintiff (in such cases, the worker) over 70% of the time (IPP, 2017). The average severance payment required in this case is around 25,000 euros, or 10 monthly wages. This credible threat leads companies to compensate their permanent employees way more than the legal requirement in case of a termination, in a move to agree not to go to court. Importantly, the law also states that the CDI is also the default type of contract: signing any other type of contract requires the firm to justify special circumstances.

The second type of contract is the CDD (*contrat à durée déterminée*). It is France's main temporary contract. The CDD can be signed in a very limited number of circumstances, where the job is *by nature* temporary. These cases include the replacement of workers on temporary leave; temporarily filling a position recently vacated and that will be suppressed in the near future; filling a position before a permanent worker starts; responding to a temporary increase in activity[1]; filling a position that is seasonal by nature (for instance in the tourism sector). Other very specific circumstances allow using a CDD, for instance to recruit a professionnal athlete for a game season. The goal of these regulations are to avoid making the CDD the standard contract that employers would choose for flexibility, which would defeat all the employment protection that comes with a CDI. In particular, a company cannot rely on repeated CDD to avoid the protective burden of a CDI, even in times where business is uncertain. On top of these hiring motive restrictions, additional regulations ensure that the CDD is not used abusively. We can summarize them into four rules, leaving aside some very specific and minor cases:

- a CDD can be renewed twice; sector-specific agreements can increase the number of authorized renewals, but they remain very limited in practice
- a CDD can last 18 to 24 months in total, including all allowed renewals; again, the maximum duration can be extended by certain agreements, but does not exceed 36 months in practice
- when a CDD ends and is not legally renewed with the same employee, the company cannot re-hire a CDD for the same job for some time; while the exact time window may depend on sectoral agreements, it is generally between a third and half of the total duration of the CDD

[1]This reason is probably the most vague one and could lead to abuses.

that just ended

- at the end of a CDD, the employer must pay the employee a severance payment (*prime de précarité*) equal to 10% of the total pay that was perceived during the contract, to compensate for the precarious situation of the worker who just lost his job.

The main threat to a company who breaks any of the above rules is the right for the employee to require his CDD to be turned into a CDI, which can mean years of salaries for the company to pay even in the case the employee is not needed anymore. All companies must file a document to the French administration each time they hire an employee (including CDD renewals) and frequent controls are performed.

Figure 3.1 shows how these two contracts' shares have evolved over the last two decades. As mentionned before, the CDI is by law the standard type of contract. As a matter of fact, it is by far the most prevalent contract in France, accounting for 92.1% of the total number of CDIs and CDDs in the service sector, on average over 1998-2008. The CDD, while being of frequent recourse for firms, only accounts for 7.9% of the stock of contracts. These contract shares are remarkably stable over time.

3.2.2 The specific case of CDDU

A specific type of CDD allows a company to bypass all the aforementioned rules. This contract, called the CDDU (*contrat à durée déterminée d'usage*) is allowed in a well-defined set of sectors, either according to article D. 1242-1 of the Labor Code or to a sector-specific agreement (*convention collective*). These sectors have an important share of their activity that is temporary by nature, hence the exception to the standard CDD rules. The sectors' description is reported in table 3.1. For instance, hotels and restaurants have important variations in their business intensity over the year, requiring them to adjust their workforce considerably. Entertainment sectors, such as the movie, exhibitions, sports, etc. industries have a mostly project- or event-based activity, that also requires frequent and important changes in the size of the workforce. It is important to emphasize that with CDDUs, companies in these sectors can freely re-hire temporary workers, for any period of time (no lower or upper bound), as many times as they want, and without ever paying a severance

payment for the end of such a contract. All sectors allowed to rely on CDDU are within the service sector. As shown by table 3.1, they are however very diverse. This diversity will be an important argument for the causal identification of the effect of the reform studied here.

3.2.3 The change in the CDDU jurisprudence

While belonging to the list of CDDU sectors (table 3.1) is a necessary condition to hire under a CDDU contract, it is not (and has never been) sufficient. Prior to 2003, conditions to use a CDDU were threefold:

- the company must belong to the CDDU list
- there must be a frequent use of temporary contracts for the same type of job, in other companies of the same sector
- the job itself has to be temporary by nature

These conditions can of course not be verified by the administration for each and every hire under CDDU. They are evaluated by a judge when a worker goes to court about an abusive CDDU (for instance if the worker is re-hired many times, over a long period, to perform tasks that are part of the permanent business of the company). Just as for abusive CDD complaints, if an employer is found guilty of abusing CDDUs, they have great chance to be forced to turn the contract into a CDI, which means important costs for the company.

In November 2003, the court of final appeal for civil matters (*Cour de Cassation*) changed its jurisprudence regarding the aforementioned conditions. In a judgement involving workers sueing their company for having hired them under CDDU while they were not expected to perform temporary tasks but permanent ones, the court did not retain the "temporary nature" requirement. This change in the jurisprudence means that, in principle, employers in the CDDU sectors can now hire workers on a succession of very unrestricted temporary contracts in order to fill a position that is actually permanent. The reach of such a change is potentially very important in a country like France. Indeed, this change provides CDDU-sector employers an essentially perfect and less costly alternative to the CDI. In particular, this could enable them to avoid all the employment protection provisions that come with a CDI. While employers could already use this contract for temporary

jobs, CDDU-sector companies have a significant amount of regular activity that is performed by permanent workers. As a consequence, the expansion of CDDU to jobs that are not temporary by nature could concern a dramatic amount of people in these sectors.

3.3 Data

3.3.1 Data sources and sample selection

The main source of data used in this paper is the French Employment Survey (*Enquête Emploi*) which is produced by the French National Statistical Institute (INSEE). I use this survey from 1998 to 2008. Despite having data up to 2016, I want to avoid using the financial crisis period, which may have effects correlated with the 2003 reform and that would thus confound when doing time-aggregated analyses. The survey covers a representative sample of the French population above 15 years old and asks questions at both individual and household levels. The sample is a draw of *residences*, not individuals. When a residence is surveyed, all members of the corresponding household who are above 15 years of age are surveyed. Prior to 2003, the survey happens only in March of every year and households are surveyed 3 years in a row. From 2003 on, the survey is conducted every quarter and households are surveyed 6 quarters in a row. I do not (yet) use the panel dimesion of this survey and treat it as a repeated cross-section. I focus on individual observations that satisfy the following criteria: the person must be employed under a contract, in the private service sector and their wage must be reported. Being employed under a contract rules out self-employed persons as well as other minor forms of employment. Restricting the analysis to the service sectors is done to form a better control group in the natural experiment that the 2003 reform constitutes. CDDU sectors are indeed all service sectors. It is thus a straightforward choice to compare these sectors to other service sectors than to the entire rest of the economy. In particular, industrial sectors have very different characteristics over time and in the cross-section, so it is better to rule them out. Even within services, comparison between CDDU and other sectors is debatable. I thus also provide evidence using finer sectoral divisions where some CDDU and non-CDDU sectors that are more likely comparable are grouped. Wage data is available for every wave (i.e. every time

the same individual is surveyed) of the survey prior to 2003. Starting 2003, it is only available for two waves out of six (the first and last ones). I restrict the sample to the waves were individuals are asked about their wages. Responding to the survey is mandatory. Restricting the survey to waves where wage data are available also happens to restrict it to waves where an employee of INSEE comes in person to ask the questions and collect the answers in the household's residence. This is a good guarantee against missing data and some kinds of misreporting (e.g. misreporting due to misunderstanding the question). Since the survey sample is a sample of residences, respondents who move between two waves of the survey are not followed. The data made available by INSEE already provides correction for this bias (INSEE, n.d.). In total, the sample that I use contains 237,860 observations. There are approximately 25,000 observations per year prior to 2003. From 2003 on, this number drops to 17,000 on average. The yearly number of observations actually increased after 2003, but this is counteracted by the fact that only one third of waves now contains wage data. Accounting for the fact that my sample contains several waves for some persons, the sample holds a total of 164,684 unique individuals.

The main variables that I use are:

- sector affiliation: whether or not the individual works in a CDDU sector, and some other sectoral grouping variables;
- wages: the survey reports net monthly wages (that is after payroll taxes, and including potential benefits/bonus given by the firm) for everyone who is employed;
- number of hours worked: the survey reports hours worked in the reference week, as well as "usual" amount of hours worked weekly;
- contract duration: for fixed-term contracts, the survey reports a duration in weeks prior to 2003, and in days after 2003[2];
- age;

[2]I have concerns that this variable might be misreported, especially by people employed in a sequence of very short temporary contracts, which is a phenomenon that the 2003 reform seems to have encouraged. When someone who agreed off the record with their employer on a 1-month job that will be subdivided into four 1-week contracts, they might report a duration of one month instead of the actual one week. This should be checked against administrative data.

- education level: the survey reports very detailed information about degrees earned and fields of study;
- occupation type: the survey reports very detailed information about the type of job;
- gender;

Unfortunately with these data, the exact identification of people under CDDU is impossible. This information has not been collected in any French statistical dataset until 2014[3]. The best proxy to being hired under an actual CDDU contract that the data can provide is to be hired under a CDD in a sector that is allowed to use CDDU. For such a sector, the requirements to hire under a CDDU are not stricter than for a standard CDD (and they become weaker after 2003, when using a CDDU for a non-temporary task becomes allowed) and the CDDU is a more flexible and less costly option for the employer. There is thus good reason to think that most CDDs are in fact CDDUs in CDDU sectors. Identifying sectors that are allowed to use CDDUs is done by using the NAF (*Nomenclature d'activités française*) codes of sectors in table 3.1. NAF is a French sectoral classification system with over 700 classes. It is finer that the CDDU sectors, in the sense that most sectors in table 3.1 consist of several NAF codes. NAF codes thus enable an almost perfect identification of CDDU sectors. It is important to note that the absence of an exact identification of people under a CDDU contract will not threaten the validity of most of the analysis presented here. Most findings indeed feature comparisons at the sector level (which are well-identified) and not, for instance, comparisons between individuals in CDDU contracts and individuals in standard CDD contracts.

Finally, I also use data from the statistics branch of the French Social Security (ACOSS) to which every employer has to report when they hire an employee. The public data they provide is restricted to the number of hires by sector (88-class sector classification), contract type (CDD under one month, CDD over one month and CDI) and by year. The fairly coarse sectoral aggregation as well as the absence of any individual information makes these data very aggregate. I use them to support one of my findings on the substitution between permanent and temporary contracts following the

[3]I requested funding to access administrative data from the statistics branch of the French Social Security that will provide detailed information on where CDDU contracts are used from 2014. Despite not being contemporaneous to the 2003 reform, these data can provide a valuable picture of the CDDU usage.

2003 reform.

3.3.2 Weighting and seasonal corrections

The INSEE Employment Survey experiences important changes in 2003. It goes from a first-quarter-only yearly survey, to a continuous survey performed every quarter. In terms of unique individuals, the sample size is roughly divided by two after 2003. The frequency weights that come with the survey are thus naturally multiplied by about two, so that the sample remains representative of the French population of 15 years and older. There is also a special set of weights to make waves 1 and 6 (which are the only waves to report some data, e.g. wages) representative of the full population. Since these two waves put together form about one third of the whole post-2003 sample, their weights are about three times as large as the full-sample weights. But keeping waves 1 and 6, which are four quarters apart, does not change the number of individuals that we have in the sample (unless they dropped out of the survey, but non-response bias already has some correction embedded in the weights provided by INSEE). So we end up with a sample for which the yearly sum of weights triples after 2003. To have consistent sums of weights across years, I thus divide the post-2003 weights by 3. Despite this correction, some breaks in the series between 2002 and 2003 survive and appear on the graphs I show. They do not however threaten my findings, which rely on post-2003 variations.

Another potential issue with the 2003 survey changes is that it now adds data from other quarters than the first one (prior to 2003, the survey only asked question about employment and wages in March of each year). I check that the introduction of these new data does not affect CDDU sectors in a different way than other service sectors, which could be a potential confounding factor. Table 3.2 shows the results of a regression of log wages on a dummy variable for working in a CDDU-sector company (CDDU), binary variables for quarters (Q_t, $t = 2, 3, 4$) and the interaction of CDDU and quarter variables:

$$\log \text{wage}_{it} = a_0 \text{CDDU}_{it} + \sum_{t=2}^{4} b_t Q_t + \sum_{t=2}^{4} c_t \left[Q_t \times \text{CDDU}_{it} \right] + \epsilon_{it}$$

The sample used is the one described section 3.3.1 restricted to years 2003-2008, where quarterly data is available. The interaction coefficients show that there are no significant differences in wages

between CDDU and other service sectors that come from using specific quarter data. If anything, the estimates are actually positive, and we will later find that CDDU-sector wages drop relative to other services' wages after 2003. As will be clear in the results section, findings other than wage variations tend to kick in 2004 and later, which limits the possibility that they are due to adding data from other quarters.

3.3.3 Descriptive statistics on the Employment Survey

In this section, I provide some basic summary statistics from the French Employment Survey. Table 3.3 summarizes them. It is split between the period before the reform, 1998-2003, and after the reform, 2004-2008. These statistics are split along two sector groups: the ones allowed to use CDDU (a.k.a the treatment group) other services (control group). This table shows that CDDU and other service sectors are highly comparable along some demographic dimensions of their workers, namely age, the ratio of women in the workforce and the number of years of schooling. The evolution of these characteristics between the period 1998-2003 and 2004-2008 seem comparable as well, with very little time variation in total. Differences arise when looking at labor market dimensions, here the wage, the ratio of CDD contracts and the duration of these contracts. Average wages are generally lower in CDDU sectors. They grow by about 11% between 1998-2003 and 2004-2008, while other services' wages grow by nearly 15%. As expected, the ratio of CDDs is significantly higher in CDDU sectors, with 12% against 6% in other service sectors. Between the two time periods, this ratio gains 2.7 percentage points within CDDU sectors, while only 0.7 points in other service sectors. This hints at some ongoing process of raising the share of temporary workers in CDDU sectors over the period 1998-2008, as compared to other service sectors. Finally, CDD contracts are typically shorter in CDDU sectors, by about a month on average. We can see that the duration of CDD contracts decreased by 20% for both sector groups. While there is no difference between CDDU and other service sectors when looking at the average duration changes, I will later provide evidence that CDDU sectors made a more aggressive move towards very short temporary contracts following the reform.

3.4 Results

3.4.1 Substitution between temporary and permanent contracts

CDDU sectors constitute a substantial and quite stable share of employment in France's service economy, around 15%, as shown in figure 3.2. We do see a slight decrease around 2003, the year of the reform, which is most likely due to the survey weighting changes. Overall, the employment share of CDDU sectors does not seem to vary significantly aroung 2003.

The 2003 jurisprudence change increases the opportunity to substitute temporary for permanent workers in companies within the CDDU sectors. It extends the set of jobs for which a CDDU contract is allowed to jobs that are not temporary by nature. Employers now have a way to hire workers for an indefinite period of time, on a wide range of tasks, without having to rely on a CDI. In particular, the reform enables them to avoid the protection given to CDI workers, which has been blamed for preventing firms to adjust their labor force flexibly enough and for refraining them to hire in the first place.

Figure 3.3 shows the evolution of the shares of CDD and CDI contracts within the CDDU sectors and within other service sectors (CDD/(CDD+CDI) or CDI/(CDD+CDI)). Before the reform, CDDU sectors and other service sectors show similar flat trends in both contract types' share. CDDU sectors experience a sharp increase in the use of CDD contracts, whose share went up from 12% over 1998-2004 to 14.9% after 2004, on average. Other service sectors experience a much slighter increase in the share of CDDs, from 6.6% to 7.3%. A symmetric pattern is observed for CDI contracts. In CDDU sectors, their share falls sharply from an average of 88% before 2004 to 85.1% after the reform, while other service sectors experience a drop of about a fourth the size.

This contract composition changes might be driven by confounding factors affecting only one or some of the CDDU sectors. One could still observe important aggregate effects if these particular sectors account for a great enough share of employment in CDDU sectors. To argue against this possibility, I provide two pieces of evidence. First, figure 3.4 shows the employment share of different NAF36[4] groups within CDDU sectors, both over the period 1998-2003 and over 2004-

[4]NAF36 is a 36-class sector classification. While most NAF36 groups contain both CDDU and non-CDDU sectors,

2008. This figure shows that there is only two important sector composition changes within CDDU sectors, namely the shrinkage of educational services and the rise of healthcare and social services. Second, figure 3.5 shows the share of CDDs in CDD+CDI both over 1998-2003 and over 2004-2008, for CDDU sectors, split by NAF36 groups. We see that most groups exhibit the common pattern of a rise in the share of CDDs, with the exception of consulting services and personal and home services. The sharpest drop happens in personal and home services, but they account for less than 2% of employment in CDDU sectors. Finally, educational services and healthcare services both exhibit a rise in temporary employment, so that we can't attribute the aggregate CDD/CDI substitution pattern shown in figure 3.3 to these sector composition changes appearing in figure 3.4.

Another piece of evidence for the substitution of temporary contracts for permanent ones is the hiring patterns observed. Figure 3.6 shows the total number of hires in CDDU and other services sectors, split by three contract types: short CDD (less than 1 month), long CDD (over 1 month) and CDI. We see that the evolution of hires are similar across sectors both the long CDD contracts and the CDI ones. The difference is however striking when it comes to short CDDs: the trend in their number breaks clearly in CDDU sectors, while it increases only slightly in other service sectors[5]. This, combined with the increase in the stock of CDDs within CDDU sectors, reflects two things. First, CDDU sectors simply more unique individuals in CDDs than prior to the reform. Second, the average duration of a CDD in this sector has decreased, a finding that is supported by the summary statistics in the Employment Survey. In all, these first pieces of evidence suggest that CDDU-sector employer now rely more on temporary contracts and make them shorter, exploiting

I filtered the non-CDDU ones out of the data both for figures 3.4 and 3.5.

[5]This could seem at odds with the summary statistics presented in the Data section, which showed very similar evolutions of average temporary contract durations in CDDU and other service sector. However, a simple story could explain this apparent discrepancy: short CDDs might be more frequently used ion CDDU sectors that in other service sectors. When looking at the average duration of contracts existing at a given point in time, one could see little difference. But if the people under a short CDD contract get more systematically renewed in CDDU sectors than they are in other service sectors, this creates a difference in the number of hires on short temp contracts. This is consistent with the spirit of the reform, which allowed CDDU employers to use sequences of short CDDUs instead of longer temp or permanent contracts for an extended population of workers.

the full extent of the CDDU contract's flexibility by renewing the contract of the same individual many times rather than committing to a long CDD or a CDI.

Overall, the reform seems to have little to no effect on the employment share of CDDU sectors. However, within the mass of people who work for CDDU sectors, there are substantial changes. First, there is a clear substitution of temporary for permanent contracts showing up in stock. Second, there is also flow evidence of a change. A much greater number of hires under temporary contract start happening in the CDDU sectors after the reform. This in part translates the increase in the use of CDDs, but this also likely reflects a shortening of temporary contracts and their multiplication for the same individuals.

3.4.2 Wages

In this section, I present evidence that the reform has weighted on wages in the CDDU sector, especially those of young permanent workers. There are several reasons why wages may have been impacted by the reform. Since there is no minimum duration for a CDDU, the employer can hire workers by the task. Intermittent working patterns may arise, where workers are paid only some days or weeks within a month, reducing their total income compared to a continuous employment situation. While a standard CDD cannot be renewed indefinitely, CDDU give the possibility for an employer to keep the same worker under this precarious situation for a long time, while being able to terminate the worker at any time and no cost. This may reduce both the employer's incentive to train or invest in the worker and the worker's ability to acquire professional skills. Limited accumulation of human capital likely weighs on wage growth. Another reason why wages may be impacted is a bargaining power story. Provided with a less costly alternative to the CDI, that they can now even apply to workers on non-temporary positions, CDDU employers become less inclined to offer CDI contracts. Since offering a CDDU is much more flexible than a CDI and exempt of termination costs for the employer, they may only want to offer CDI at reduced wages, to compensate for picking the worse option. This is likely to go through as the bargaining power of the employer increased: when negotiating the terms of a CDI, they can now wave the threat of transforming the offer into a sequence of CDDUs. Being employed under a sequence of temporary

contracts rather than a CDI has dramatic consequences for French workers. Beyond the usually lower level of worker satisfaction associated with temporary jobs (see e.g. Blanchard and Landier, 2002) a temporary contract is often not enough to access credit or rentals at reasonable prices (Decreuse and Ypersele, 2011). Workers are thus likely to accept a CDI for lower pay.

Young workers' wages suffer from the reform

To analyze the effect of the reform on wages, I use a difference-in-differences strategy. I choose the non-CDDU service sectors as a control group. This setup calls for several caveats. First, the control group is made of different, mostly non-overlapping sectors. There may be a number of confounding effects that impact either the treatment group (CDDU sectors) or the control group (other services) differently and would compromise causal inference on the effect of the reform. A argument against this possibility is the fact that both treatment and control groups contain very diverse sectors. In particular, CDDU sectors comprise the movie and TV industry, healthcare services, consulting services, hotels and restaurants, etc. It is unlikely to see an effect on all these sectors but on no other service sector, except for changes in the CDDU rules that ties them together. Second, the effects on CDDU sectors could be driven by changes in a only a few sectors that could be unrelated to the reform (e.g. a drop in the tourism business caused by outside factors). Figure 3.5 already showed that the CDI/CDD substitution patterns were common to most CDDU sectors. In this section, I provide further evidence that wage effects are observed in several CDDU sectors and not driven by just one or two of them.

The difference-in-differences setup specifically estimates the following equation:

$$\log \text{wage}_{it} = \alpha \text{CDDU}_{it} + \gamma \text{Post}_t + \beta \text{CDDU}_{it} \times \text{Post}_t + X'_{it}\delta + \epsilon_{it}$$

where Post_t is a dummy variable for begin after 2003; CDDU_{it} is a dummy variable for being in a CDDU sector, versus being in another service sector; X_{it} are additional demographic controls, such as occupation category, education level, gender and 5-year age bracket. Table 3.4 shows the results of this estimation for three age brackets: 15 to 34 ("young" hereafter), 35-49 (mid-aged) and 50-64 (old). For each age bracket, I present the simple specification results without controls, as well as the specifications adding controls X_{it}. Although I present results for all age brackets, figure 3.7

suggests that the parallel trend assumption is most likely to be verified for the young people group (and maybe fine for the mid-age group). The simple specification results are consistently negative and highly significant, while adding controls seems to add enough correlation to lose significant results. The results for the young group are the ones that survive best the addition of demographic controls. I focus on this group from now on.

The interpretation of the regression coefficients for the 15-34 years old group is a little tricky, since it does not simply come from a trend change, but from a trough and a reversion back to the previous trend, approximately. What we can say is that these results suggest that the reform is associated with a drop in wages for the young of about 2% per year between 2003-2005, in CDDU sectors. The growth rate of wages then reverts back to its previous value.

To reinforce the results on the young, we run another difference-in-differences by accounting for every year fixed effects:

$$\text{log wage}_{it} = \alpha \text{CDDU}_{it} + \sum_t \gamma_t \text{Year}_t + \sum_t \beta_t \text{CDDU}_{it} \times \text{Year}_t + \epsilon_{it}$$

where Year_t is now a categorical variable, whose reference is the reform year 2003. We plot the coefficients of interest β_t in figure 3.8, for the sample of ages 15-34. Prior to 2003, the estimated coefficients are mostly not significantly different from zero. We see that after 2003, there is a fall in log wages, leading to significantly negative coefficients in 2004 and 2005. After 2005, the wage growth differences between CDDU sectors and other service sectors are reduced back to pre-2003 levels, emphasizing an important short-term effect that is not persistent.

As mentioned earlier, I also perform a check to verify that the fall in young workers' wage growth within CDDU sectors is not driven by an idiosyncratic shock to a particular sector that would be unrelated to the reform. For example, an unrelated negative shock on the tourism industry could have impacted hotels and restaurants, which account for almost 40% of employment in CDDU sectors and could thus have triggered the more aggregate effects we observe. To rule out this possibility, I examine four sector subgroups that have a mix of CDDU-authorized sectors as well as other service sectors. These subgroups are: transportation services, services to business, home and personal services, and educational, healthcare and social services. Table 3.5 summarizes the results. We see that the estimates of interest are consistently negative. We have significant estimates

for the personal and educational-health-social service sectors. In all we observe consistent drops in wages in different subgroups of CDDU sectors, relative to other related services sectors. This is important evidence for two reasons. First, the fact that we see a change in very diverse sectors is suggestive of the effect of the reform, which is probably the only thing they have in common. Second, comparing CDDU sectors with non-CDDU ones falling under the same business category and finding differences is stronger evidence that something is happening precisely with the CDDU sectors. In other words, if something unrelated to CDDU contracts was happening, we would really expect other service sectors within the same category to have similar behavior.

The reform impacts permanent workers most negatively

Focusing on young workers, we perform additional tests by contract type. Table 3.6 shows the effects on young CDI workers and on young CDD workers. We see that the entire effect we observed so far is driven by permanent workers. There is no significant evidence of a change in CDD workers' wages. This constitutes an interesting a rather new finding. Most studies have shown that facilitating the access to or relaxing the rules of temporary employment fail to increase employment in a sustainable way, rather increase worker turnover and raise the share of temporary jobs, which are associated with lower wages, human capital and level of satisfaction. However, there is to my knowledge no study of the side effects of such reforms on permanent workers. The 2003 French reform has likely affected the attractiveness of the CDI (which was, until 2003, the only way to keep a worker indefinitely with non-temporary duties) for employers in CDDU sectors, which may have become reluctant to pay these permanent workers as much as before. The fact that most of the effect happens on young generations is also suggestive of two things. First, employers likely adjusted the wage levels of CDI workers at the moment of a hire, which is an event that happens more to young workers. Indeed, French law prohibits an employer from reducing the salaries of their current employees (without a significant job change). This leaves the hiring margin as the main margin for wage adjustment. Second, young workers are likely to have less bargaining power than older workers. It is thus easier for employers to negotiate lower wages with this population, here by waiving the new threat of hiring under CDDU instead of CDI.

Finally, I provide some additional evidence that something in the employer-employee relationship as changed with the 2003 reform, as suggested by the line of thought just developed. I perform Oaxaca decompositions between CDDU sectors' and other service sectors' wages along two dimensions: the occupation type and the education level of workers. The occupation type is summarized in 4 categories: lead management, management, employee and blue-collar worker. The education level is summarized in 7 categories, based on the French schooling system. They roughly correspond to different lengths of schooling. The Oaxaca decomposition enables us to examine how much of the difference between wages in CDDU and other service sectors comes from compositional difference in either occupation types or education levels, and how much of it comes from the differential returns to occupation or education across these sectors. I perform standard Oaxaca decompositions within each year of data and I plot the results in figure 3.10. The sample used is restriced to workers aged 15-34 and in a permanent contract. The "coefficient" line denotes the contribution of group-specific (occupation or education) return differentials to the overall difference in wages across CDDU and other service sectors. The "endowment" line denotes the part of the variation that comes from compositional effects. Finally, the interation term is a residual that completes the contributions to the overall variation. It is, as expected, mostly negligible in front of the coefficient and endowment effects.

These preliminary results show that the coefficient effects generally dominates, and especially so in the rise of the wage gap that follows the reform. This suggests that rather than a composition effect, where the reform would have triggered movements of specific types of population in or out of the CDDU sectors, something has changed in the returns to occupation and education in CDDU sectors. Given the nature of the reform, it is likely that the employer-employee relationship has changed, with additional bargaining power given to the employer.

3.5 Conclusion and future research

We find significant effects of a reform extending the opportunities of hiring under a very unrestricted temporary contract, the CDDU, for some sectors of the French economy. These effects are twofold. First, there is a substitution of temporary contracts for permanent contracts after the reform,

reflecting the desire of employers for more flexibility in their workforce adjustment. Observing this effect is welcome as is has already been highlighted by the literature on the deregulation of temporary contracts. Second, the reform also seems to weigh on permanent workers' wages. This novel finding has not received much attention. The literature has often focused on the relatively worse situation of temporary workers as compared to permanent ones and neglected potential spillover effects. My findings are backed by checks within sectoral subgroups. The mechanism that seems to come out of these results is a bargaining power story: by becoming widely allowed to hire under a very flexible temporary contract, employers receive a perfect, less costly alternative to the permanent contract. This makes the permanent contract less attractive, especially provided the employment protection provisions that it comes with in France. To obtain a permanent contract — which remains in high demand, being the key to housing and credit markets among other things — workers now need to accept lower wages.

The findings in this article require deeper analysis. Further work will be done to understand the change in the employer-employee relationship better. Part of this work can be done using the data that I have currently. I notably intend to use the panel dimension of the survey to learn more about workers trajectories and their evolution around the reform in CDDU and other service sectors. Some information can be obtained on whether CDD workers have been re-hired by the same company (sometimes termed "recurrence") by combining information on their current CDD contract's duration as well as on their tenure length in their current company. This will enable us to see whether CDDU employers have indeed started providing more sequences of short-term CDDs in place of long term ones or permanent contracts.

The relatively short-term nature and the small size of the wage effects makes it likely that some unobservable confounding effect remains at work and drives the results partly. To that purpose, I requested funding to access French administrative employment data, coming from the *Déclaration Annuelle de Données Sociales* (DADS). This data comes from the requirement for an employer to file a report both when signing a contract with an employee and yearly on every employee that works in the company. These data thus contain the universe of French employees, with information concerning contract types, duration of contract, wages, number of hours worked, etc. Data at both firm and establishment levels are also reported, such as NAF codes (necessary to identify

companies falling under CDDU sectors) and number of employees. These data could help back my findings more. It would be especially useful to provide enough observations to do comparisons between CDDU companies and non-CDDU ones within tighter sectoral subgroups, to improve the comparability of the treatment and control groups.

Figures

Figure 3.1: Shares of the three main contract types in total employment under contract.

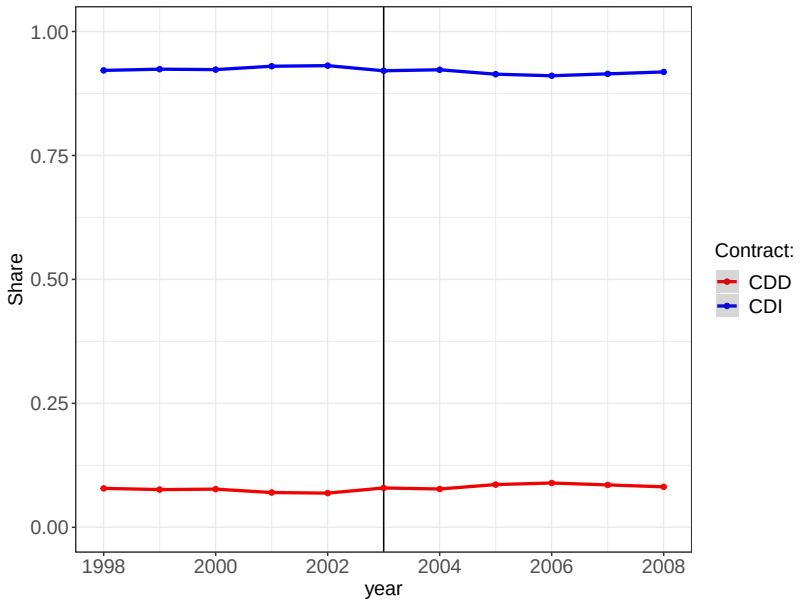


Figure 3.2: Employment in CDDU and non-CDDU sectors, as a share of total service sector employment.

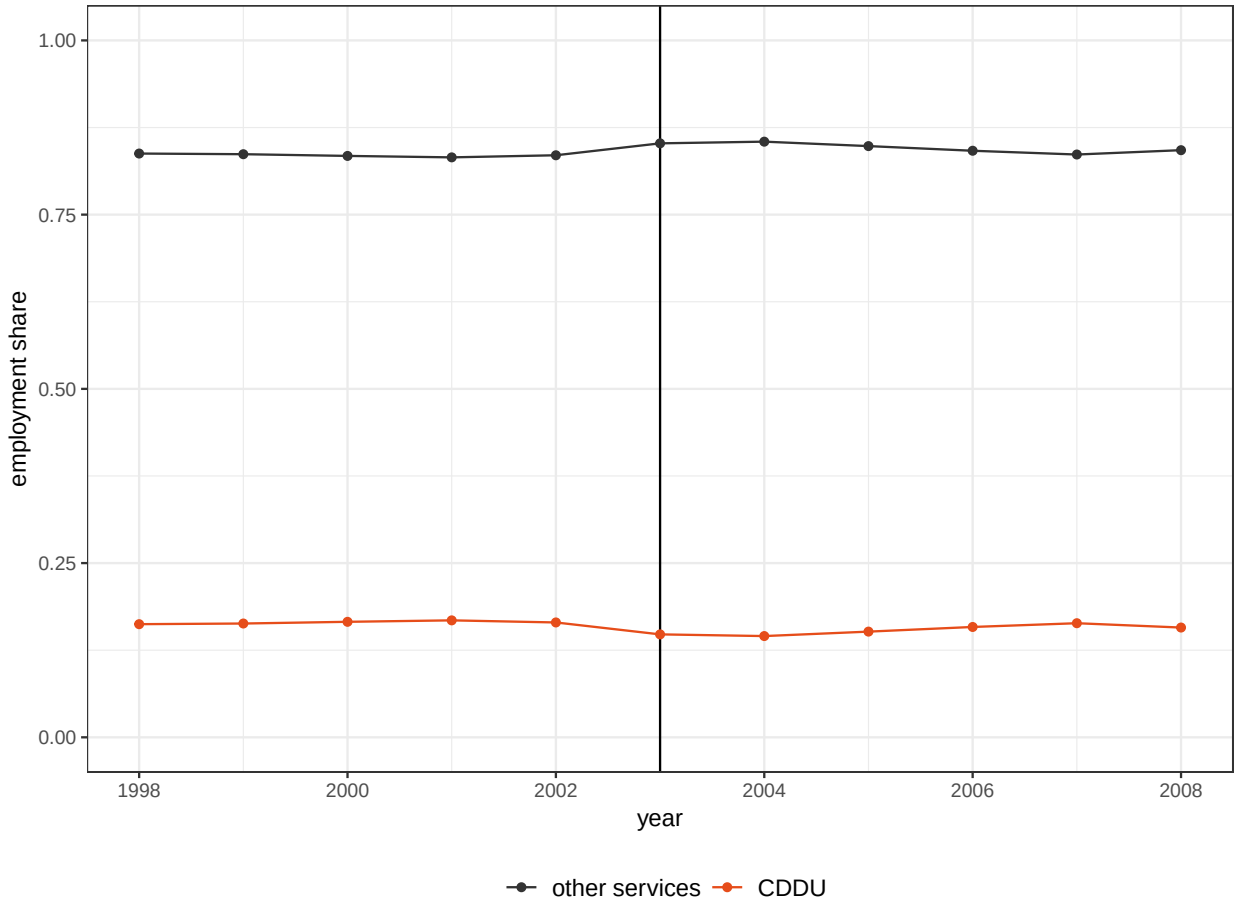


Figure 3.3: Shares of CDD (left panel) and CDI (right panel) in CDD+CDI total, in CDDU sectors and in other service sectors. Gray areas represent 95% confidence bands.

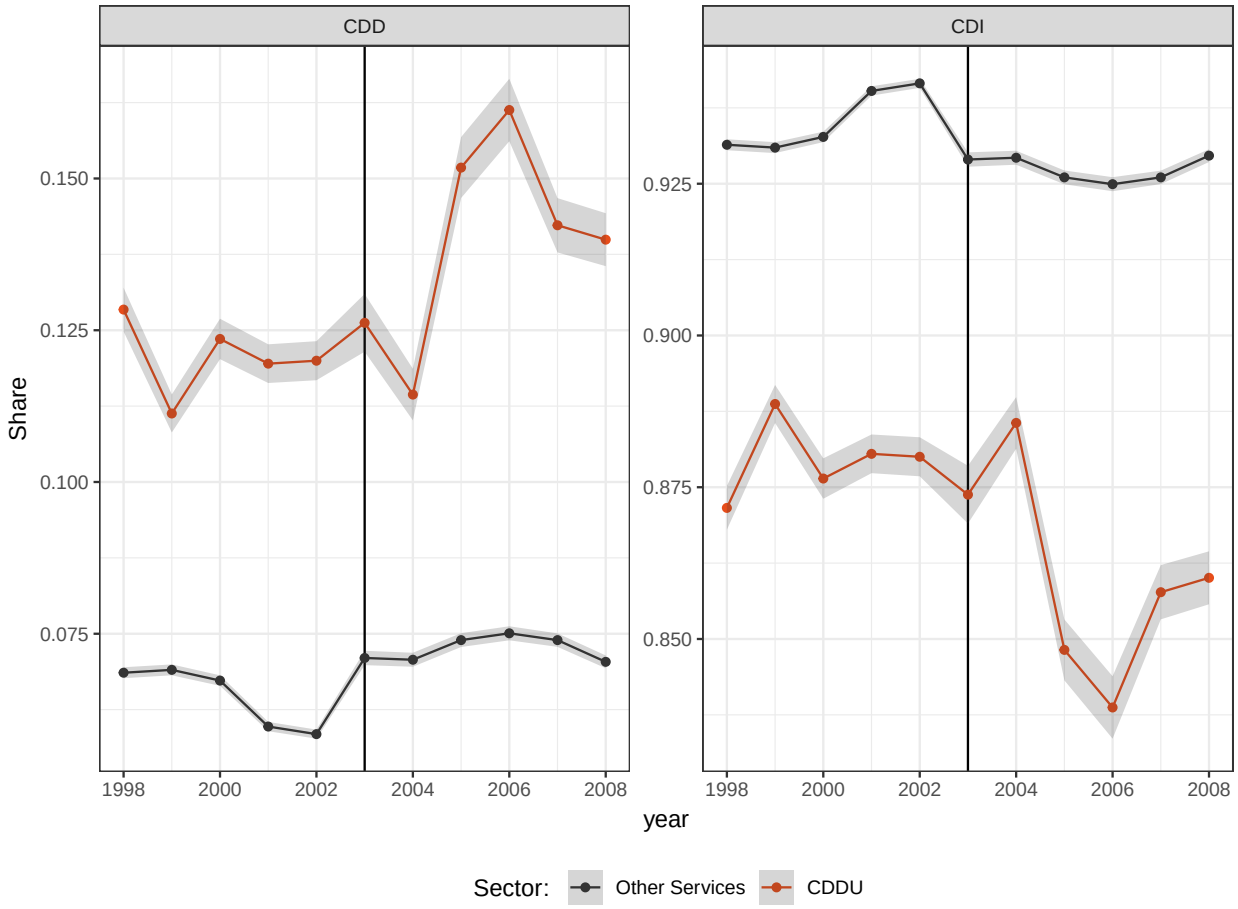


Figure 3.4: Average share of employment in different CDDU sectors over 1998-2003 and over 2004-2008, as part of total CDDU employment.

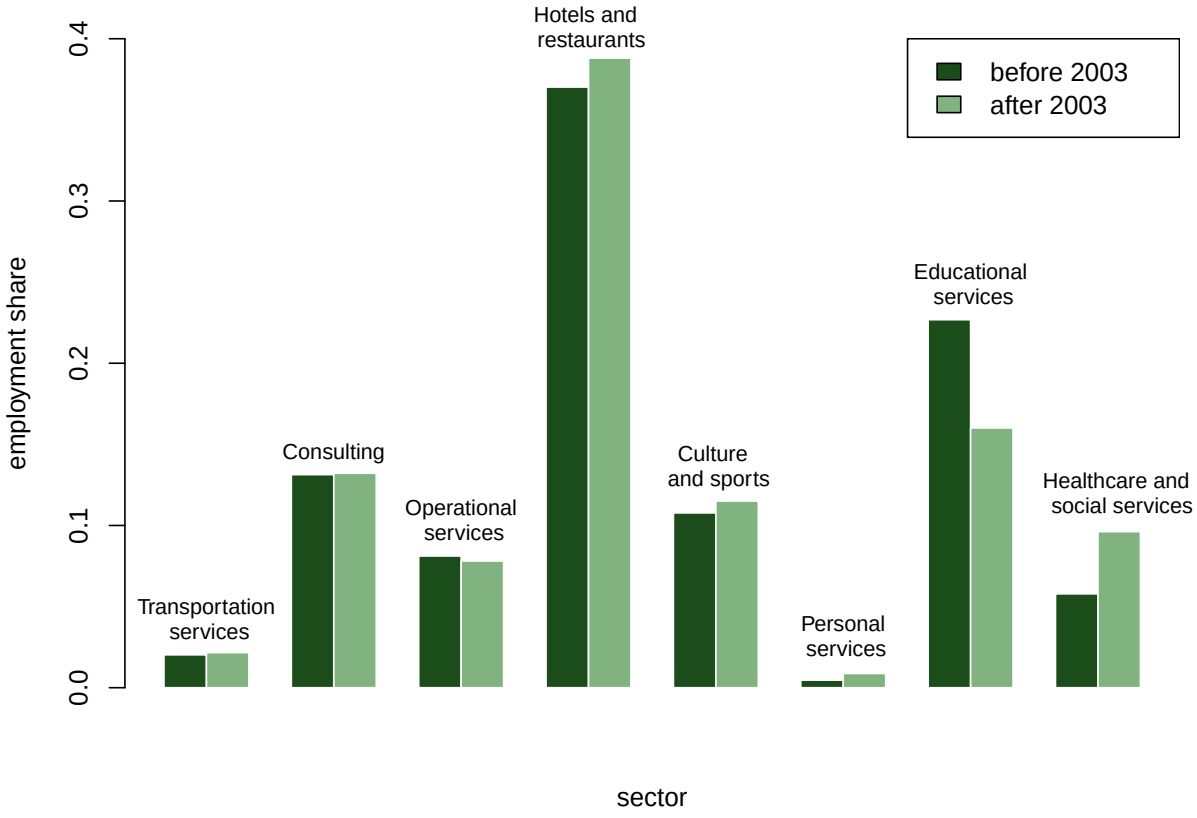


Figure 3.5: Shares of CDD in CDD+CDI, in CDDU sectors, before and after the reform. Vertical bars represent 95% confidence intervals.

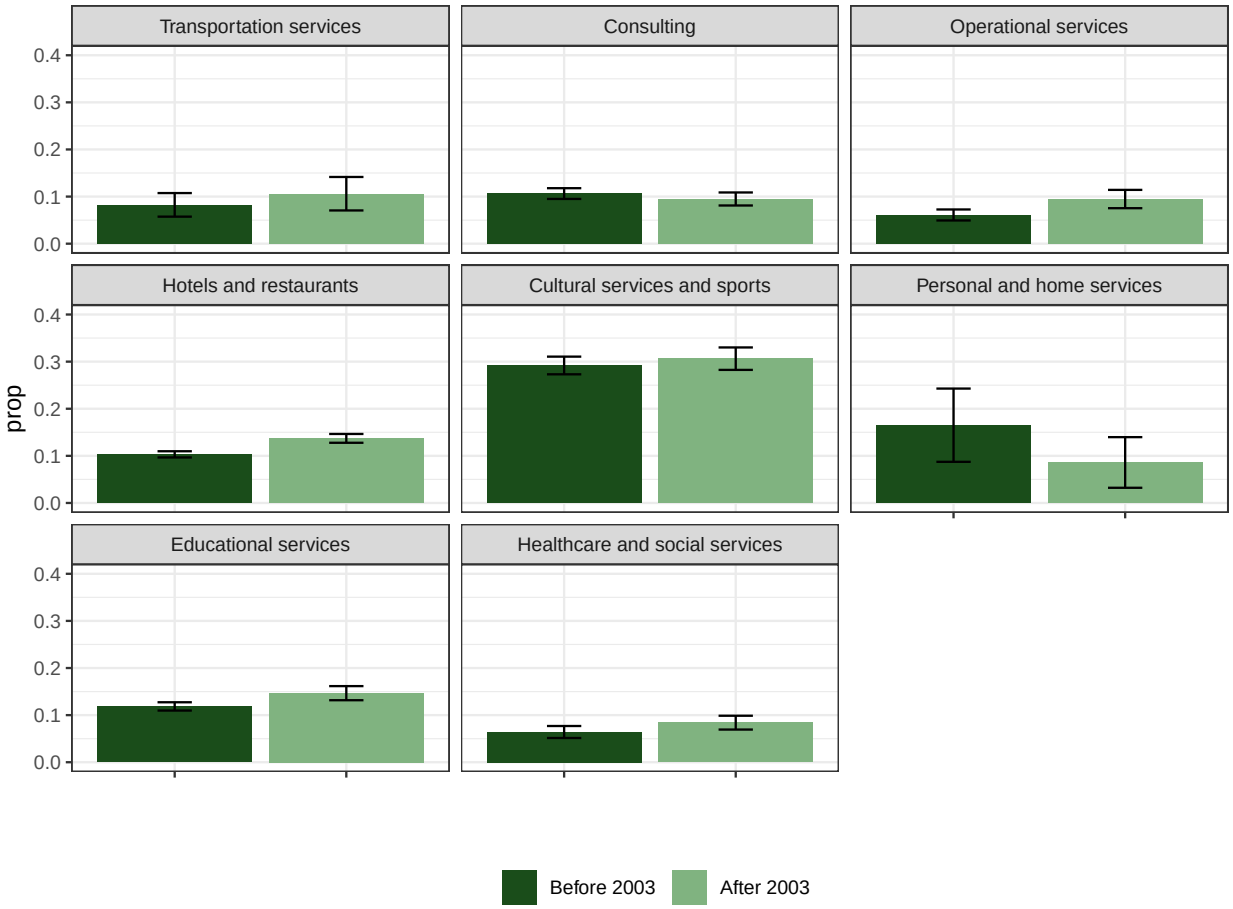


Figure 3.6: Number of hires by contract type (short CDD < 1 month, long CDD > 1 month, CDI, total) and by sector group (CDDU sectors and other service sectors).

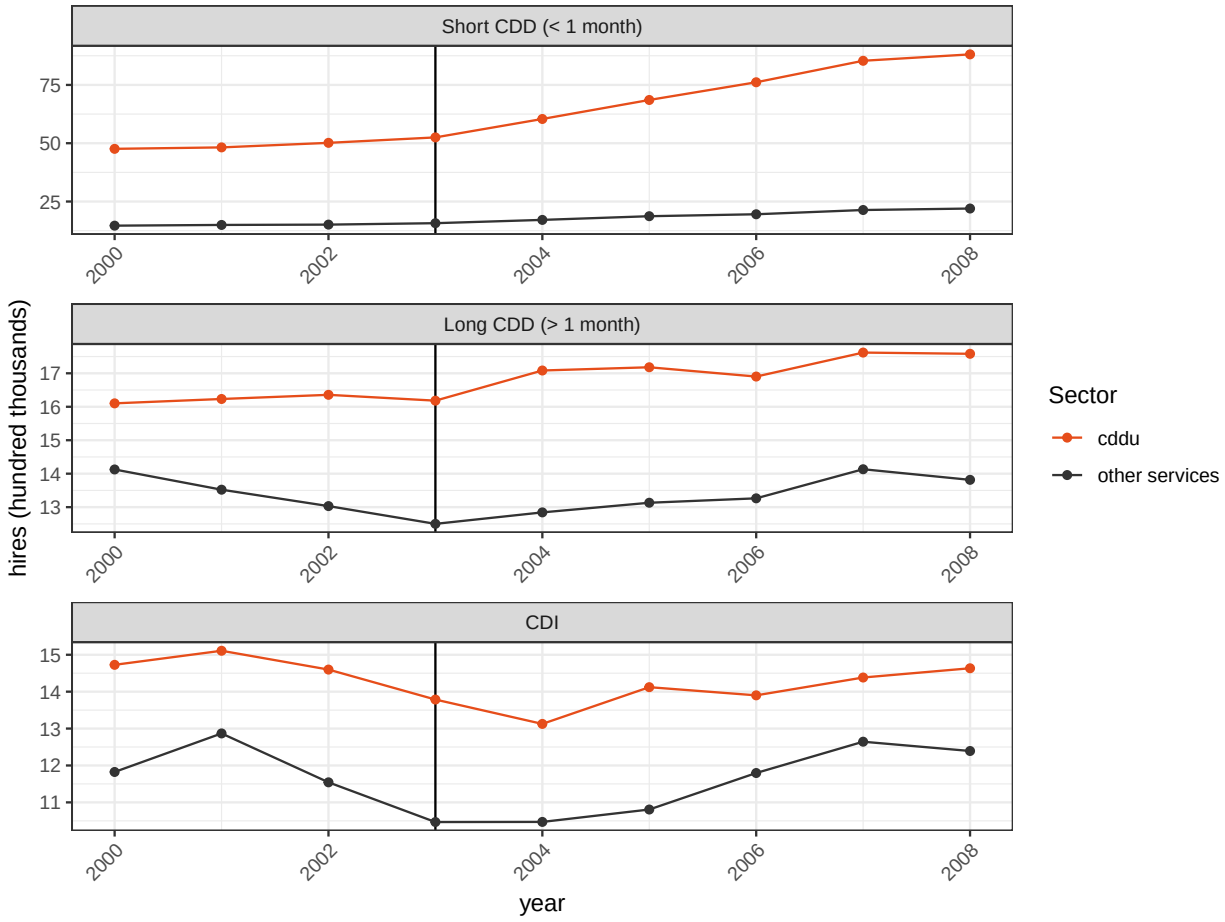


Figure 3.7: Evolution of log wages by age bracket and by sector group. Gray areas represent 95% confidence bands.

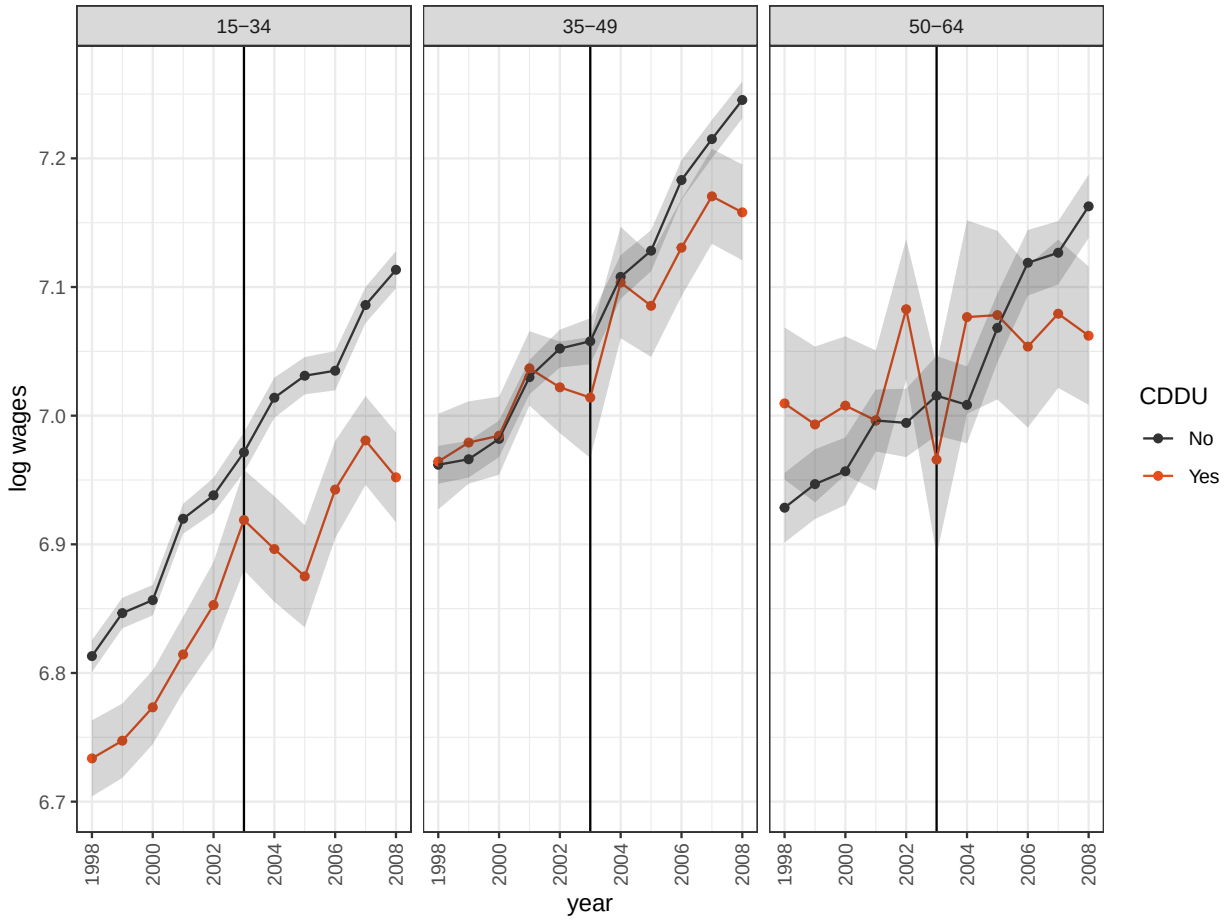


Figure 3.8: Coefficients β_t from the year-fixed-effects difference-in-differences regression: $\text{wage}_{it} = \alpha \text{CDDU}_{it} + \sum_t \gamma_t \text{Year}_t + \sum_t \beta_t \text{CDDU}_{it} \times \text{Year}_t + \epsilon_{it}$. Vertical bars represent 95% confidence intervals.

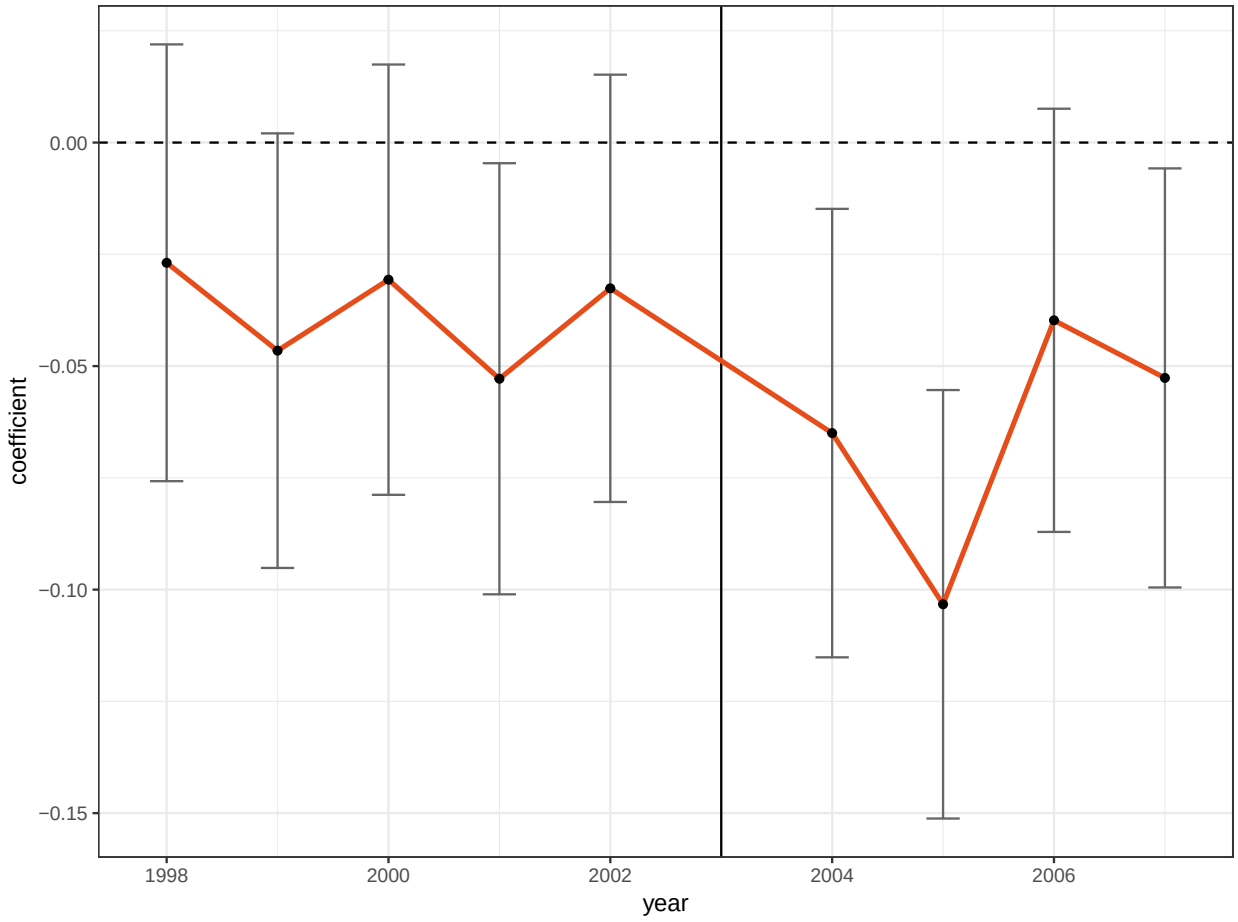


Figure 3.9: Evolution of log wages by sectoral subgroups that contain both CDDU and non-CDDU service sectors. Gray areas represent 95% confidence bands.

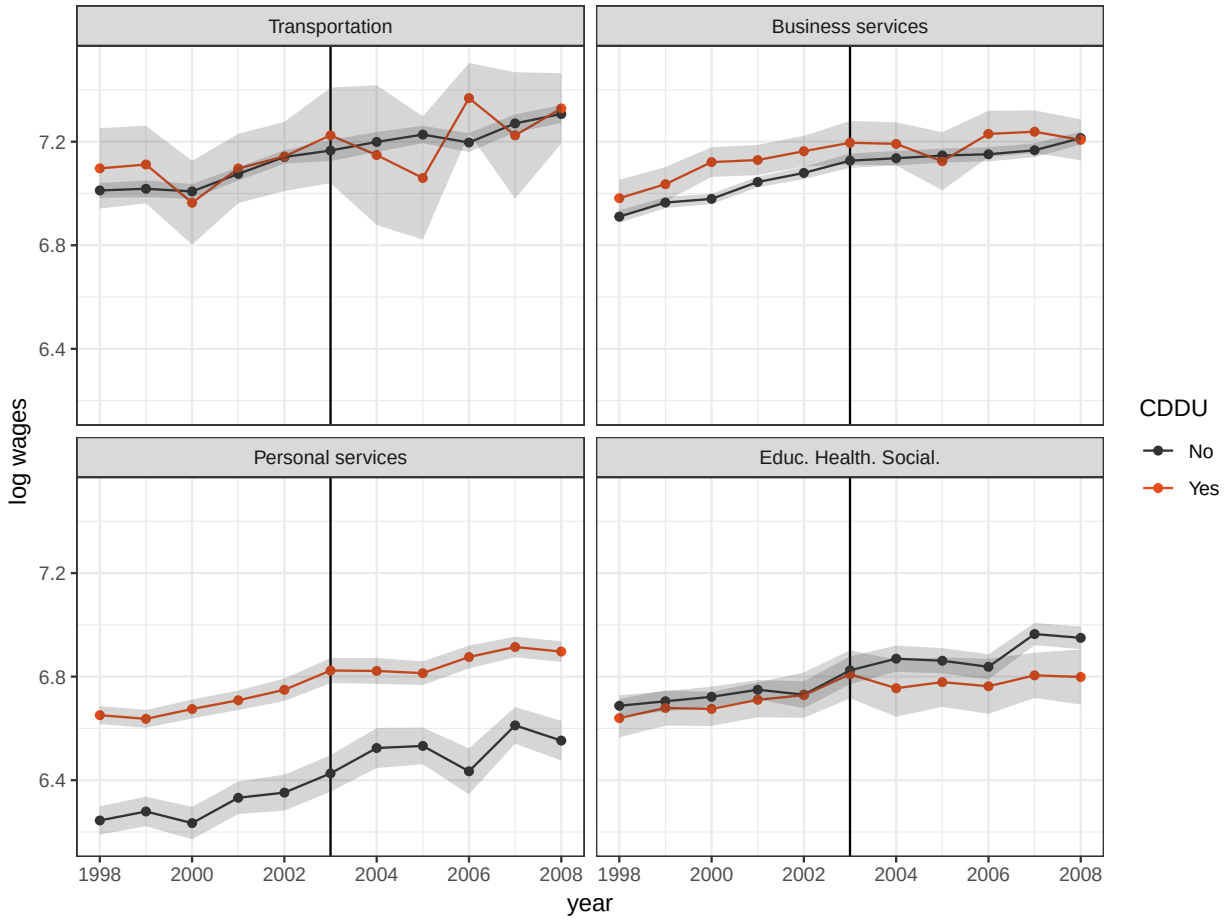
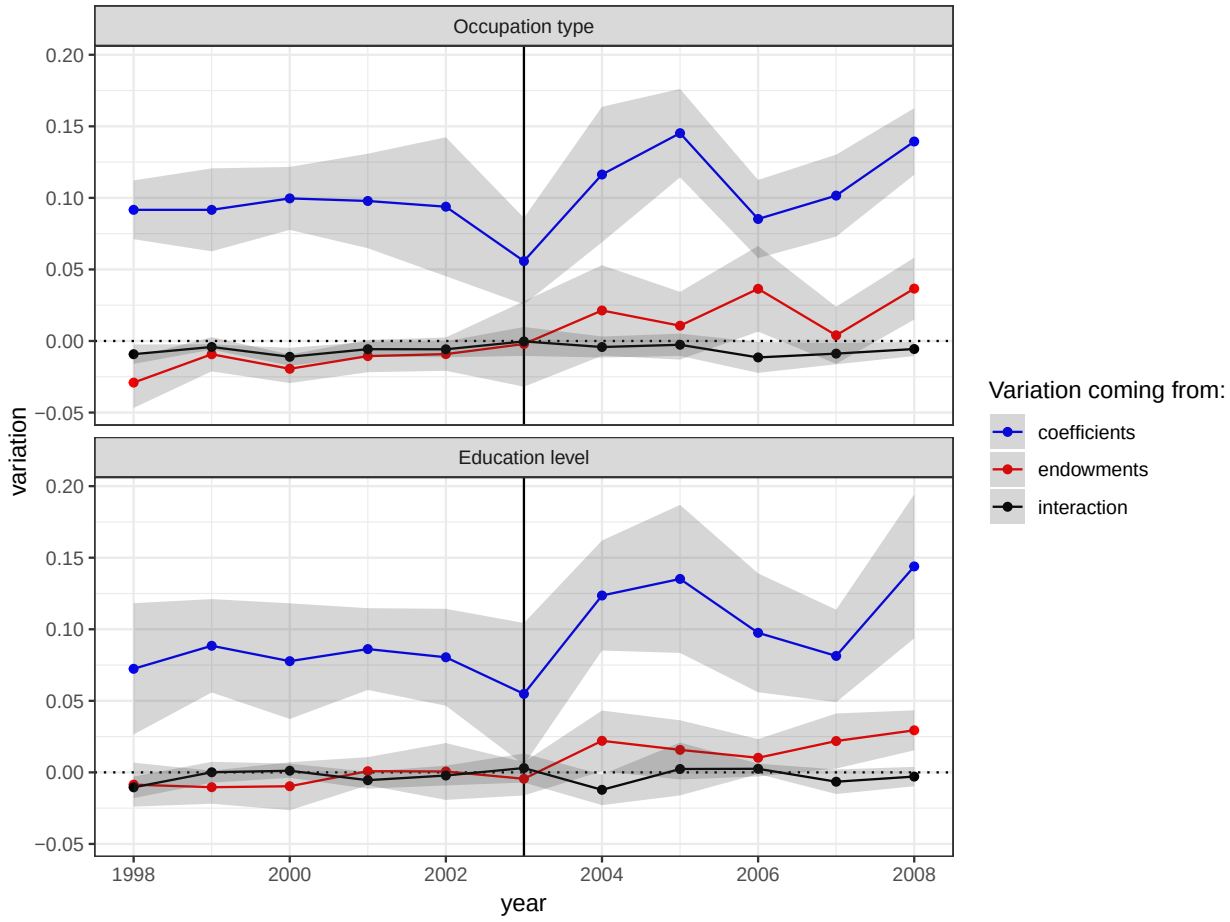


Figure 3.10: Oaxaca decompositions for occupation type (top panel) and education level (bottom panel) for young workers in permanent contracts. Gray areas represent 95% confidence bands.



Tables

Table 3.1: Sectors allowed to use the CDDU contract

Sector	Legal Source
Creative activities, shows	Labor Code Article D. 1242-1
Restaurants	Labor Code Article D. 1242-1
Advertising	Sectoral agreement
Hotels	Labor Code Article D. 1242-1
Movie industry	Labor Code Article D. 1242-1
Other professional support activities	Sectoral agreement
Teaching and tutoring	Labor Code Article D. 1242-1
Sports	Labor Code Article D. 1242-1
Exhibition organization	Sectoral agreement
Home care services	Labor Code Article D. 1242-1
Polls	Labor Code Article D. 1242-1
Professional insertion services	Labor Code Article D. 1242-1
TV industry	Labor Code Article D. 1242-1
Movers	Labor Code Article D. 1242-1
Cargo handling	Sectoral agreement
Casinos	Sectoral agreement

Table 3.2: Testing the effect of adding quarters 2-4 data on the wage gap between CDDU and other service sectors

	Dependent variable:
	log(salred)
Q2	0.006
	(0.006)
Q3	0.026^{***}
	(0.006)
Q4	0.026^{***}
	(0.006)
CDDU	-0.084^{***}
	(0.011)
Q2 × CDDU	0.006
	(0.016)
Q3 × CDDU	0.019
	(0.015)
Q4 × CDDU	0.004
	(0.016)
Constant	7.086^{***}
	(0.004)
Observations	105,707
Adjusted R^2	0.002
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 3.3: Summary statistics from the French Employment Survey, by period (pre/post reform) and by sector group

	1998-2003		2004-2008	
	CDDU	Other Services	CDDU	Other Services
Age	37.4	38.1	37.9	38.9
	(10.9)	(10.6)	(11.3)	(11)
Women Ratio	0.56	0.53	0.58	0.54
	(0.5)	(0.5)	(0.49)	(0.5)
Years of Schooling	10.5	10.6	10.3	10.3
	(2.95)	(2.9)	(2.88)	(2.92)
Monthly wage (euros)	1229	1263	1366	1450
	(765)	(755)	(814)	(794)
CDD/(CDD + CDI) ratio	0.12	0.066	0.149	0.073
	(0.318)	(0.237)	(0.343)	(0.249)
CDD duration (days)	286	315	228	251
	(337)	(379)	(288)	(342)

Table 3.4: Log wages difference-in-differences estimates for different age-brackets

	Dependent variable:					
	Log wages					
Age bracket	15-34	15-34	35-49	35-49	50-64	50-64
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	6.891^{***}	6.378^{***}	7.009^{***}	7.898^{***}	6.979^{***}	8.010^{***}
	(0.003)	(0.415)	(0.003)	(0.007)	(0.006)	(0.013)
CDDU	-0.086^{***}	-0.088^{***}	-0.008	-0.086^{***}	0.037^{**}	-0.064^{***}
	(0.007)	(0.006)	(0.008)	(0.006)	(0.014)	(0.011)
Post 2003	0.151^{***}	0.143^{***}	0.152^{***}	0.118^{***}	0.108^{***}	0.104^{***}
	(0.004)	(0.004)	(0.005)	(0.004)	(0.009)	(0.007)
CDDU × Post 2003	-0.030^{***}	-0.010	-0.029^{**}	0.017^{*}	-0.054^{**}	0.006
	(0.011)	(0.009)	(0.013)	(0.010)	(0.022)	(0.017)
Additional controls		✓		✓		✓
Observations	89,099	89,087	89,305	89,301	39,814	39,814
Adjusted R^2	0.018	0.332	0.012	0.394	0.004	0.398

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3.5: Log wages difference-in-differences estimates by sectoral subgroups containing both CDDU and non-CDDU sectors

	Dependent variable:			
	Log wages			
	Transportation	Business services	Personal services	Educ. Health. Social
	(1)	(2)	(3)	(4)
CDDU	0.042	0.085***	0.392***	−0.032*
	(0.031)	(0.014)	(0.016)	(0.018)
Post 2003	0.155***	0.131***	0.214***	0.149***
	(0.011)	(0.008)	(0.020)	(0.015)
CDDU × Post 2003	−0.052	−0.036	−0.060**	−0.076***
	(0.045)	(0.022)	(0.025)	(0.029)
Constant	7.070***	7.020***	6.314***	6.735***
	(0.007)	(0.005)	(0.013)	(0.010)
Observations	5,666	24,550	13,842	9,538
Adjusted R^2	0.037	0.013	0.077	0.013
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 3.6: Log wages difference-in-differences by contract type for workers aged 15-34 y.o.

	Dependent variable:		
	Log wages		
Contract type	CDD	CDI	CDI
	(1)	(2)	(3)
Constant	6.609***	6.975***	7.102***
	(0.010)	(0.003)	(0.018)
CDDU	−0.011	−0.066***	−0.055***
	(0.020)	(0.008)	(0.006)
Post 2003	0.120***	0.165***	0.143***
	(0.016)	(0.005)	(0.004)
CDDU × Post 2003	0.001	−0.049***	−0.011
	(0.030)	(0.012)	(0.010)
Additional controls			✓
Observations	9,813	65,683	65,677
Adjusted R^2	0.008	0.021	0.363
Note:			*p<0.1; **p<0.05; ***p<0.01

Bibliography

Abowd, John M and Francis Kramarz (2003). "The costs of hiring and separations". en. In: *Labour Economics* 10.5, pp. 499–530. ISSN: 09275371. DOI: 10.1016/S0927-5371(03)00017-4. URL: http://linkinghub.elsevier.com/retrieve/pii/S0927537103000174 (visited on 12/07/2017).

Bassanini, Andrea and Andrea Garnero (2013). "Dismissal protection and worker flows in OECD countries: Evidence from cross-country/cross-industry data". In: *Labour Economics* 21, pp. 25–41. ISSN: 0927-5371. DOI: 10.1016/j.labeco.2012.12.003. URL: http://www.sciencedirect.com/science/article/pii/S0927537112001273 (visited on 05/15/2018).

Bentolila, Samuel and Giuseppe Bertola (1990). "Firing Costs and Labour Demand: How Bad is Eurosclerosis?" In: *The Review of Economic Studies* 57.3, pp. 381–402. ISSN: 0034-6527. DOI: 10.2307/2298020. URL: http://www.jstor.org/stable/2298020 (visited on 10/15/2017).

Bentolila, Samuel and Gilles Saint-Paul (1992). "The macroeconomic impact of flexible labor contracts, with an application to Spain". In: *European Economic Review* 36.5, pp. 1013–1047. ISSN: 0014-2921. DOI: 10.1016/0014-2921(92)90043-V. URL: http://www.sciencedirect.com/science/article/pii/001429219290043V (visited on 04/10/2018).

Bertola, Giuseppe (1990). "Job security, employment and wages". In: *European Economic Review* 34.4, pp. 851–879. URL: http://www.sciencedirect.com/science/article/pii/0014292190900668 (visited on 10/15/2017).

Bertola, Giuseppe (1994). "Flexibility, Investment, and Growth". en. In: *Journal of Monetary Economics* 34, pp. 215–238. DOI: 10.3386/w3864. URL: http://www.nber.org/papers/w3864.pdf (visited on 03/28/2018).

Blanchard, Olivier and Augustin Landier (2002). "The Perverse Effects of Partial Labour Market Reform: Fixed-Term Contracts in France". In: *The Economic Journal* 112.480, F214–F244. ISSN: 0013-0133. URL: http://www.jstor.org/stable/798373 (visited on 05/01/2018).

Boeri, Tito and Pietro Garibaldi (2007). "Two Tier Reforms of Employment Protection: A Honeymoon Effect?" In: *The Economic Journal* 117.521, F357–F385. ISSN: 0013-0133. URL: http://www.jstor.org/stable/4625528 (visited on 05/02/2018).

Cahuc, Pierre (2012). "For a Unified Contract". en. In: *European Labour Law Journal* 3.3, pp. 191–206. ISSN: 2031-9525, 2399-5556. DOI: 10.1177/201395251200300302. URL: http://journals.sagepub.com/doi/10.1177/201395251200300302 (visited on 05/01/2018).

Cahuc, Pierre and Fabien Postel-Vinay (2002). "Temporary jobs, employment protection and labor market performance". In: *Labour Economics* 9.1, pp. 63–91. ISSN: 0927-5371. DOI: 10.1016/S0927-5371(01)00051-3. URL: http://www.sciencedirect.com/science/article/pii/S0927537101000513 (visited on 09/07/2018).

Centeno, Mário and Álvaro A. Novo (2012). "Excess worker turnover and fixed-term contracts: Causal evidence in a two-tier system". In: *Labour Economics* 19.3, pp. 320–328. ISSN: 0927-5371. DOI: 10.1016/j.labeco.2012.02.006. URL: http://www.sciencedirect.com/science/article/pii/S0927537112000176 (visited on 05/01/2018).

Decreuse, Bruno and Tanguy van Ypersele (2011). "Housing market regulation and the social demand for job protection". In: *Journal of Public Economics*. Special Issue: International Seminar for Public Economics on Normative Tax Theory 95.11, pp. 1397–1409. ISSN: 0047-2727. DOI: 10.1016/j.jpubeco.2011.06.001. URL: http://www.sciencedirect.com/science/article/pii/S0047272711000892 (visited on 06/25/2018).

Ignacio García Pérez, J. and Victoria Osuna (2014). "Dual labour markets and the tenure distribution: Reducing severance pay or introducing a single contract". In: *Labour Economics*

29, pp. 1–13. ISSN: 0927-5371. DOI: 10.1016/j.labeco.2014.05.001. URL: http://www.sciencedirect.com/science/article/pii/S0927537114000554 (visited on 05/02/2018).

INSEE (n.d.). *Correction des ruptures de série suite au passage en 2003 de l'enquête Emploi annuelle à l'enquête Emploi en continu*. fr. Tech. rep., p. 10.

IPP (2017). *Prudhommes: Peut-on Expliquer La Disparité Des Décisions ?* Tech. rep.

Kahn, Lawrence M. (2010). "Employment protection reforms, employment and the incidence of temporary jobs in Europe: 19962001". In: *Labour Economics* 17.1, pp. 1–15. ISSN: 0927-5371. DOI: 10.1016/j.labeco.2009.05.001. URL: http://www.sciencedirect.com/science/article/pii/S0927537109000591 (visited on 08/30/2018).

Kramarz, Francis and Marie-Laure Michaud (2010). "The shape of hiring and separation costs in France". en. In: *Labour Economics* 17.1, pp. 27–37. ISSN: 09275371. DOI: 10.1016/j.labeco.2009.07.005. URL: http://linkinghub.elsevier.com/retrieve/pii/S0927537109000761 (visited on 12/20/2017).

Kugler, Adriana, Juan F Jimeno, and Virginia Hernanz (2005). *Employment Consequences of Restrictive Permanent Contracts: Evidence from Spanish Labor Market Reforms*. en. Tech. rep., p. 41.

Kugler, Adriana and Giovanni Pica (2008). "Effects of employment protection on worker and job flows: Evidence from the 1990 Italian reform". In: *Labour Economics* 15.1, pp. 78–95. ISSN: 0927-5371. DOI: 10.1016/j.labeco.2006.11.002. URL: http://www.sciencedirect.com/science/article/pii/S092753710600090X (visited on 03/31/2018).

Saint-Paul and Bentolila (1994). "A model of labor demand with linear adjustment costs". In: *Labour Economics*. URL: https://ac.els-cdn.com/0927537194900159/1-s2.0-0927537194900159-main.pdf?_tid=fcd1cb82-e03c-11e7-bbba-00000aab0f02&acdnat=1513194000_b13d4d52c3bbfb15695fb4c19ca1bb4c (visited on 12/13/2017).