Essays on Local Labor Markets

Federica Daniele

TESI DOCTORAL UPF / Year 2018

THESIS SUPERVISOR
Jan Eeckhout and Giacomo Ponzetto
Department d'Economia i Empresa

A Maria Ester, Roberto e Costanza

Acknowledgments

When I sailed off for my PhD the global economy had just started recoverying from what had been the worst recession since the Great Depression. I was looking for definite answers: six years after, I could not be farther from the initial goal. Research is a collective effort: taken together the individual contributions expand the frontier, but none of them does so independently.

There are, nevertheless, private benefits of doing research: in my own experience, the two most fascinating aspects of research I discovered while doing it are the objectivity and sense of fairness that it grants, and the creativity that it sparks. By now these principles guide my everyday decisions and interaction with others. I would like therefore to spend a few words thanking here all those that helped me getting started on and finishing this journey.

I thank Francesca Salvati, whom I met during my undergraduate studies and who ignited the intellectual curiosity that led me to pursue a PhD.

I am deeply grateful to my supervisors, Jan Eeckhout and Giacomo Ponzetto. The discussions with Jan taught me the method and discipline necessary when thinking of an economic question and how to break it down for analysis. The conversations with Giacomo equipped me with creative thinking and intellectual honesty, which both accompany me in everyday life.

I benefited from the support and encouragement of the macroeconomics group at UPF, especially of Regis Barnichon, Julian di Giovanni, Jordi Galí, Manuel Garcia-Santana, Dávid Nagy, Edouard Schaal and Jaume Ventura. I am also indebted to Julia Faltermeier and Marta Santamaria, who provided loyal comments throughout the writing of this thesis. I gratefully acknowledge the cooperation of Heiko Stüber, to whom I owe the data on which most of the empirical analysis in this thesis is based.

I thank Elizaveta Konovalova and Marlene Rump, who have been family to me during these years in Barcelona, and Robbie Hill, who gave me the strength in the most difficult moments. Above all I thank my family, Maria Ester, Roberto and Costanza, who, in spite of the distance, never stopped feeding me with love and attention during these years.

Abstract

This thesis is composed of three essays in which I analyze how heterogeneity in productivity, either on the worker or on the firm side, interacts with the size of local labor markets and a set of outcomes of interest. In the first chapter, I analyze how the presence of firm-level uncertainty affects consumers and cities. I provide evidence supporting entrepreneurial risk-seeking in the non-tradable sector and that this has the strongest consequences for competition in large cities. I show how a reduction in uncertainty dampened entry and competition, and reduced the attractiveness of consumer cities. In the second chapter, I analyze the role of large firms for local labor market volatility. I provide empirical and narrative evidence supporting the existence of granularity- driven business cycles. I discuss the importance of size-dependent policies with respect to the systemic risk externality imposed by large firms on the economy. In the third chapter, I analyze how individual specialization shapes the urban wage premium. I investigate to what extent changes in specialization have accounted for the divergence in US workers location choices. I show that the evolution of specialization can explain the increase in between-cities wage inequality for high-skilled workers, while it counteracted the increase in the average skill premium.

Resum

Aquesta tesi està composta per tres assajos en els que estudio com l'heterogeneïtat en la productivitat, de part del treballador o de part de l'empresa, interactua amb la mida del mercat laboral local i en el conjunt de variables d'interès. En el primer capítol analitzo com la presència d'incertesa a nivell de l'empresa afecta als consumidors i a les ciutats. Aporto evidència que suporta la cerca de risc en el sector no comercial i que aquest té les conseqüències més forte per la competència entre grans ciutats. Demostro com una reducció en la incertesa redueix la entrada i la competència, i redueix l'atractiu de les ciutats consumidores. En el segon capítol analitzo el paper de les grans empreses i el seu paper en la volatilitat en el mercat de treball local. Aporto evidència empírica i narrativa recolzant l'existència de granularitat - cicles empresarials conduits. Discuteixo la importància de les polítiques relacionades amb la mida empresarial amb respecte a l'externalitat del risc sistemàtic imposada per les empreses grans en l'economia. En el tercer capítol analitzo com l'especialització individual dona forma al premi de salari urbà. Investigo fins a quin punt canvis en l'especialitzaciò expliquen la divergència en les decisions de localització dels treballadors del Estats Units. Demostro, que l'evolució de l'especialització pot explicar l'increment de la desigualtat salarial entre ciutats per treballadors amb alta habilitat, mentre que disminuïa l'increment en el premi per habilitat mitjà.

Preface

This thesis is composed of three essays in which I study how heterogeneity in productivity interacts with the size of local labor markets and a set of outcomes of interest. The motivation for each essay is empirical: economic theory plays nevertheless a fundamental role throughout, both ex ante in defining the relevant set of hypothesis to test in the data and ex post in providing a framework for gauging the economic significance of the results.

The notion that the dispersion of the fundamental shocks hitting the economy might change over time is an idea that has gained traction since the notion of the Great Moderation was introduced. While the Great Moderation refers to a protracted period of low macroeconomic volatility, there is evidence that microeconomic volatility has also declined substantially over the past decades, without showing a reversal in the most recent years as it has been the case for macroeconomic volatility. While this topic has recently attracted considerable attention, there has been little effort to put it into the bigger picture of the secular decline in firm entry, dynamism and overall changing nature of competition that is underway. In the first chapter, co-authored with Heiko Stüber, we take a first step towards filling this gap and we assess the welfare consequences from the point of view of consumers. We provide evidence that the decline in microeconomic volatility has been accompanied by a fall in uncertainty for companies and entrepreneurs. Exploiting administrative data for the population of German establishments we show that reduced idiosyncratic uncertainty is empirically consistent with fewer varieties available for consumption. Less idiosyncratic uncertainty reduces the likelihood that businesses grow large at the competitors' expenses, thus decreasing the value of running a firm and discouraging entry. We show that variety is more responsive to uncertainty in large cities, so that a fall in idiosyncratic uncertainty has hampered not only consumer surplus but also the extent of agglomeration. We build a tractable spatial equilibrium model of firm entry to quantify the negative implications of a decline in uncertainty for consumers and the spatial distribution of economic activity. We find that the 2.75 percent point decline in uncertainty observed during 1990-2014 in Germany has led to an average 9% decline in consumer surplus across cities, and a modest compression in the city size distribution.

The restructuring or opening of large plants and its consequences for the city where they are located is a hot topic for local policy-makers. A recent and famous example is the competition started among US cities when Amazon announced its intention to open a second headquarter. In the second chapter, also co-authored with Heiko Stüber, we investigate a negative consequence of sharing the labor market with a few very large firms: the systemic risk that arises when a few firms are too big relative to the average size of their competitors. Intuitively, if a large firm experiences difficulties in either production or demand and, as a consequence, it is forced to dismiss a large number of workers, its aggregate impact will not average out: slack in labor demand will drive down the wage and the economy will enter a recession that has a micro-origin. Since higher uncertainty reduces investment, large firms can represent a liability for the economy unless there exist proper policies in place. The existing evidence in favor of granularity-driven business cycles takes mostly a narrative approach. In this chapter, we exploit spatial heterogeneity in local business cycles and variation in the concentration of economic activity in the top local firms across cities and years to show that a deviation from steady state in concentration predicts the start of local recessions. Deeper recessions tend to be preceded by a substantial build-up in local concentration, and for recessions during which local concentration stays stubbornly high after the peak it takes longer to reach the trough that marks the start of the recovery. We supplement our findings with narrative evidence on the impact of large plants on local employment fluctuations in Germany over the past 25 years. We conclude with an overview of the policies adopted in four major European countries to limit the labor-market impact of shocks that affect individual firms. We discuss how in many of these countries more should be done to further bring down the systemic risk introduced by large firms.

The belief that individual specialization may matter for worker productivity dates back to Adam Smith, but it has received surprisingly scant empirical attention to date. In the third chapter, I construct a novel measure of specialization using data on the frequency-adjusted count of tasks involved in an occupation. I use this measure to test for the existence of a specialization wage premium and study whether large cities tend to attract more specialized workers, consistent with the Smithian argument that the division of labor is limited by the extent of the market. I provide evidence supporting the existence of a specialization wage premium, and - by showing how average specialization rises with market size - of a specialization urban wage premium. Next, I connect cross-city differences in specialization-driven productivity with the divergence in location choices between college and non-college US workers observed during 1980-2000. I find that specialization-driven productivity alone has been an important force behind the rising concentration of college workers in 1980 skill-abundant cities. At the same time, however, I also find that specialization growth over this period has been the highest for non-college workers, thus suggesting it may have counteracted the rise in the skill premium. The re-discovery of the importance of specialization as an engine of individual productivity has several interesting implications for future research, from the rising importance of between-firm wage dispersion to the complementarity with automation.

Contents

List of Figures

List of Tables

Chapter 1

DECLINING MICRO UNCERTAINTY: IMPLICATIONS FOR CONSUMERS AND CITIES

joint with Heiko Stüber

1.1 Introduction

Today there are fewer firms that feature very high or low growth rates than there were ten or twenty years ago (Davis, Haltiwanger, Jarmin and Miranda (2007)). Hence, the extent to which firms change their ranking in the firm size distribution due to random disturbances has gone down, which means that the net present value of holding a firm is increasingly less influenced by the opportunity to experience lucky draws in demand or technology and become over time a leader in a given market.[1] A long-run decline in the dispersion of idiosyncratic shocks to firms is thus easily reconcilable with the observed reduction in business dynamism, and it has the implication of reduced uncertainty for firms taking forward-looking decisions (e.g., investment, entry, hiring).[2] We use data on the population of establishments located in former West Germany to develop a proxy for idiosyncratic uncertainty, and we document that idiosyncratic uncertainty went down by 2.75 percent points over the past 25 years (Fig.1.1).[3]

[1]Another suggested interpretation behind the decline in business dynamism consists of the reduced responsiveness of firms to exogenous shocks (Decker, Haltiwanger, Jarmin and Miranda (2016)).

[2]Uncertainty can also be of the *Knightian* type (Kozeniauskas et al. (2016)): in this case agents do not know the distribution that underlying shocks are drawn from.

[3]Additionally, idiosyncratic uncertainty peaks at the onset of recessions and quickly declines as the recession unfolds (Bloom (2014)).

The macroeconomic literature concerned with the effect of an increase in uncertainty on economic activity agrees that, while this being overall negative, changes in uncertainty affects firm-level decisions through many and often opposing channels (Bloom (2014)).[4]

The evidence provided in this paper is consistent with the view that time-invariant idiosyncratic uncertainty raises the value of holding a business by making room for the chance for any given firm to grow large upon receiving a sequence of favorable, idiosyncratic shocks. In other words, lower idiosyncratic uncertainty maps into a more compressed stationary productivity distribution (Luttmer (2007)), and it harms consumers by discouraging firm entry, and therefore variety and competition.[5] This paper quantifies the extent of such long-run adjustments and the ensuing consumer surplus losses due to a permanent fall in uncertainty: in doing so, it fills a void left by the literature that has been so far concerned with the macroeconomic impact of short-run changes in uncertainty.

Conditional on consumption variety traditionally being an important advantage offered by large cities (Krugman, 1991), the fall in idiosyncratic uncertainty has the potential to either reinforce or run against spatial concentration.[6] An additional contribution of this paper is to provide evidence consistent with the view that a fall in idiosyncratic uncertainty not only has hampered product diversity, but it has done so especially in large cities. This paper quantifies the ensuing compression in the city size distribution, and therefore the extent to which declining uncertainty has opposed alternative forces that favored over the same years an increase in agglomeration and the productivity gains coming along with it.[7]

The magnitude of the decline in idiosyncratic uncertainty depicted in Fig.1.1 suggests that the consequences on consumer surplus and geographical concentration are not going to be trivial.

We construct a proxy for idiosyncratic uncertainty at the finest industry-level

[4]Among the mechanisms behind the negative impact of uncertainty on economic activity are, for example, real options effects: given that investment decisions are often partially irreversible, an increase in uncertainty induces firms to postpone them to a time when uncertainty has reverted to the initial level.

[5]Oi (1961), Hartman (1972) and Abel (1983) are the first ones to highlight the asymmetry associated with a change in uncertainty. Provided that firms can expand to exploit good outcomes and contract to insure against bad ones, a mean-preserving spread raises the return of a given investment, e.g., the decision of entry. In recent work, Schaal (2017) finds that increases in dispersion of the cross-sectional productivity distribution are positively associated with firm entry over the business cycle.

[6]A number of studies have recently documented the empirical validity of consumption variety as agglomeration mechanism (e.g., Couture (2015), Handbury and Weinstein (2015)). Benefits from variety can accrue also to local producers (Ethier (1982)), and, through the effect of trade, propagate beyond the location of producers.

[7]See, for example, de la Roca and Puga (2017) on the productivity gains for workers; Combes et al. (2012) for firms; Lucas (1988), Desmet and Rossi-Hansberg (2014), for aggregate gains.

of disaggregation available (5-digit) for non-traded industries using data on the population of German establishments.[8] The empirical hurdle is analogous to the one faced by the extant literature, i.e., volatility is observed, while uncertainty is not. While for this reason volatility and uncertainty end up often being juxtaposed concepts in the empirical macro literature, there are circumstances under which the difference between the two may be substantial. For example, changes in firm-level volatility can be due to variation in industry or regional-level volatility, or to variation in the degree of firm heterogeneity, such as the one implied by the decreasing share of young, fast growing firms in the economy (Haltiwanger, Jarmin and Miranda, (2013), Pugsley and Sahin (2015)).[9] We rely on the richness of the data to develop a proxy that is immune from the above mentioned criticisms and several others. Specifically, We decompose establishment employment growth rates into 1) co-movement among establishments located in the same city, 5-industry and year, 2) predictable growth based on establishment characteristics (i.e. size and age) and 3) a residual.[10] We interpret the within-industry dispersion of residuals as the baseline proxy for idiosyncratic uncertainty. We compare this measure to within-industry dispersion of firms forecast error (Bachmann et al. (2013)), and we find the correlation between these two measures to be high, which confirms that the baseline proxy is capturing idiosyncratic uncertainty to a good extent. Next, we link this measure to the distribution of establishments across industries and Metropolitan Areas, and find that an increase of 1 percent point in idiosyncratic uncertainty is associated with $+.5$ more establishments in small markets, and that the correlation is six times larger in big cities.[11] We also find that an increase of 1 percent point in idiosyncratic uncertainty reduces average establishment size by 1.5% in all cities. This result is robust to controlling for several industry characteristics and composition, different subsamples and definitions of idiosyncratic uncertainty, and heterogenous loadings on aggregate shocks.

Idiosyncratic uncertainty thus measured provides a proxy for the dispersion of innovations to firm productivity emphasized by the literature on random growth, where a well-defined steady-state productivity distribution follows from the as-

[8]These industries amount to 30% of aggregate employment and they on average feature low capital intensity, widespread presence of limited liability companies, etc., thus suggesting that mechanisms through which uncertainty reduces the value of the firm may be second-order.

[9]Jurado, Ludvigson and Ng (2015) also discuss the points of departure between empirical proxies and macroeconomic uncertainty.

[10]Controlling for co-movement among establishments is important for idiosyncratic uncertainty not to capture differences in aggregate volatility across industries (Gabaix (2011)). The proper geographical dimension to measure such co-movement is the city given the focus on non-traded goods. We borrow the second methodology from the empirical literature on firm dynamics that stresses the importance of mean reversion and deterministic growth based on age (Haltiwanger, Jarmin and Miranda (2013)).

[11]The median count of establishments across cities in the industries considered is 20.

sumption of proportionality in the growth process.[12] The insight from this literature is that the mean and dispersion of the steady-state productivity distribution are both increasing in the dispersion of innovations to firm productivity: when the dispersion falls, the number of lucky firms in the cross-section that have received a sequence of large, favorable shocks also goes down. We implement a calibration of Luttmer (2007) and back-out the parameters of the Pareto productivity distribution associated with estimated average and dispersion of firm growth in each industry.

What are the consequences of a decline in idiosyncratic uncertainty, and therefore a thinner right tail of the productivity distribution, on the endogenous measure of firms? While a potential entrant faces a “worse” productivity distribution, so do all competitors in general equilibrium. The proposed evidence on the nexus between uncertainty and variety suggests that the first channel is dominant.

To examine quantitatively the effects of declining idiosyncratic uncertainty on consumer surplus and the spatial distribution of economic activity, we extend the closed economy in Melitz and Ottaviano (2008) to allow for multiple industries and cities of endogenous size. Consumers/workers trade-off higher housing costs for a greater variety of goods at lower prices. Firms, on the other hand, are compensated for the “worse” productivity distribution through a less competitive environment. By the linearity of the demand system the positive effect of a decline in uncertainty through lower competitors average productivity is outweighed by the negative effect through lower own expected productivity. Since expected profits unambiguously drop as the productivity distribution becomes more compressed, competition must decline to compensate, and *especially so* in markets that are otherwise more competitive. Intuitively, moving from two to one competitor entails a larger compensation in terms of increasing market share compared to moving from 101 to 100 competitors.[13] Finally, we bring the model closer to the data by adding location-specific worker productivities, or wages: this feature allows me to test the relative importance of the consumption advantages of agglomeration vis à vis the traditional advantages in production.

Next, we estimate the parameters of the structural model by simulated method of moments. The model captures well a number of features in the data: for example, the model gets close to replicate the Zipf’s Law in the data, and underestimates the size of only the two largest cities. While the estimation does not use any information on prices, the elasticity of the price level with respect to city size is estimated to be 2%, close to the point estimate of 1.1% obtained by Handbury and

[12] Otherwise known as Gibrat’s Law, see for example Gabaix (1999), Eeckhout (2004), Luttmer (2007).

[13] The pro-competitive gain of higher idiosyncratic uncertainty would be larger in big cities also if firms competed à la Cournot. On the other hand, the magnitude of the gain would not depend on market size if firms competed à la Bertrand.

Weinstein (2015), thus validating the ensuing analysis of pro-competitive losses form declining uncertainty.

We use the estimated parameters to quantify the effects of a decline of 2.75% points in idiosyncratic uncertainty. Consumer surplus declines by 5% (13%) in the largest (smallest) city in the sample. While losses from variety account for 75% of total consumer surplus losses in the smallest city, with the remaining 25% consisting of losses due to less intense competition, the partition is roughly 50%-50% for the largest city. These numbers are fairly close to the 66%-33% partition estimated in Feenstra and Weinstein (2017) in the context of gains from globalization: the disproportionate importance of variety for smaller markets underscores the relevance of asymmetries between gains from variety vs. pro-competitive gains based on product market size. Furthermore, a decline in idiosyncratic uncertainty reduces agglomeration: the largest city shrinks by .22% and the smallest one grows by .56%. Given the invariance in this analysis of worker productivities to changes in firm-level uncertainty, the explanation behind the small impact on geographical concentration is that traditional advantages in production turn out to have better explanatory power with respect to the city size distribution compared to the advantages in consumption studied in this paper.

In spite of the effects on the city size distribution being small, the consequences for consumer surplus are large. Therefore, we finally consider the problem faced by a central government who observes the reallocation of economic activity caused by the exogenous decline in idiosyncratic uncertainty, and who is interested in minimizing consumer surplus losses. Provided that consumer surplus is not observed, We study the case in which the government attempts to revert to the status quo by implementing a proportional subsidy on the sunk cost chosen so as to minimize observed changes in the city size distribution. My calculations suggest this subsidy to be 6% of the sunk cost, and that if it were to be financed through a lump-sum tax on consumers, the tax would amount to 8% of total production in the non-traded sector.

Related Literature This paper contributes to different strands of literature. The first is the economic geography literature. As Duranton and Puga (2003) point out non-traded intermediate inputs variety is a source of increasing returns to scale conducive to urban specialization as in Henderson (1974). In this paper, we focus on the benefits from variety for consumers, with respect to which there has been a recent resurgence of interest (Glaeser er al. (2001), Handbury and Weinstein (2015), Couture (2015), Hottman (2016)). This paper is closely related to the findings in Handbury and Weinstein (2015), according to which retail price indices tend to be lower in large cities due to more variety, and those in Hottman (2016), who provides evidence that prices tend to be lower in large cities by a pro-competitive channel. These studies typically provide informative evidence in partial equilibrium settings, while the structural estimation of a spatial equilibrium model provided by this paper allows quantifying the contribution of amenities in consumption to observed agglomeration patterns. This paper is also connected to recent work on agglomeration economies with heterogeneous firms (Combes et al. (2012), Gaubert (2018). In contrast with these papers, this study focuses on ex-ante uncertainty as opposed to ex-post heterogeneity, and it analyzes how this contributes to generating differences in consumption opportunities across cities, as opposed to nominal wages.

The analysis conducted in this paper also touches different areas of investigation in the macroeconomics literature. First and foremost, it speaks to the literature on declining business dynamism (Davis et al. (2007), Decker et al. (2014)). While this strand of research has paid special attention to dissecting the empirical patterns underlying the secular decline in business dynamism, to the best of our knowledge it has so far abstracted from quantifying the implications for consumers, and the current paper takes a step in this direction. Additionally, this paper is related to the literature concerned with the relevance of random and non-random firm growth for business cycle fluctuations and trade patterns (Gabaix (2011), di Giovanni et al. (2014)), and aggregate growth (Haltiwanger et al. (2013), Foster et al. (2006)). This paper also bears a connection with the literature on random growth and equilibrium size distribution (Gabaix (1999), Eeckhout (2004), Luttmer (2007)), by highlighting for the first time the implications of declining firm dynamism for the cross-sectional productivity distribution of firms. On a more theoretical side, this paper also contributes to the debate on the effects of changes in uncertainty on economic activity (Bloom (2014)). By emphasizing the positive effect of higher idiosyncratic uncertainty on firm entry decisions, this paper is connected to the work of Schaal (2017), who finds that upswings in uncertainty lead to an increase in unemployment but also to an expansion along the extensive margin over the business cycle, and the traditional work of Oi (1961), Hartman (1972) and Abel (1983), who study the consequences of idiosyncratic uncertainty on firm investment decisions.

Finally, this paper is related to the fast-growing literature on monopolistic competition and endogenous market power, both theoretical (Kokovin et al. (2012), Parenti et al. (2017)), and applied to cross-country trade (Melitz and Ottaviano (2008), Simonovska (2015)). The contribution of the current paper to this literature is to provide an estimate concerning the magnitude of pro-competitive gains associated with an expansion in market size in a spatial equilibrium setup.

The rest of the paper is organized as follows: Section 1.2 describes the empirical approach and the novel findings on the connection between idiosyncratic uncertainty and variety by market size, Section 1.3 lays out a model of endogenous product market competition and market size, Section 1.4 presents the results of the model estimation and the counterfactual analysis, Section 1.5 concludes.

1.2 Empirical Analysis

1.2.1 Methodology

The empirical analysis is conducted in two stages: in Section 1.2.1 we describe how we investigate the association between idiosyncratic uncertainty and variety and the interaction effect with market size; in Section 1.2.1 we illustrate the approach we follow to construct a proxy for idiosyncratic uncertainty in narrowly defined industries.

Idiosyncratic Uncertainty and Variety

Given a measure of idiosyncratic uncertainty, from here on also referred to as risk, we run the following OLS regression:

$$n_{mk} = \gamma_m + \beta_1 \text{Uncertainty}_k + \beta_2 \text{Uncertainty}_k \times \text{LARGE}_m + \epsilon_{mk} \tag{1.1}$$

The outcome variable, n_{mk}, consists of the ratio between the number of establishments observed in Metropolitan Area m (or city), 5-digit industry k in a given year, and a normalizing factor, $\overline{N} = 250$[14]; γ_m are city fixed effects, Uncertainty_k is a proxy for industry-specific idiosyncratic uncertainty, and LARGE_m is a dummy variable that takes value 0 if the city is small, 1 if it is large.[15] In a few instances

[14] $\overline{N} = 250$ corresponds to the 97.5^{th} percentile of the cross-city/industry distribution of establishment counts. This normalization is adopted to ease comparability with the theoretical model: given the continuous measure of firms in the model, we will match the relative number of establishments in the data to the relative measure of firms in the model, with both values being rescaled by $\overline{N} = 250$.

[15] The full list of Metropolitan Areas for all European countries is available at http://ec.europa.eu/eurostat/. A Metropolitan Area is considered large according to the following definition: we calculate total employment in the sample we use per each Metropolitan Area (*Arbeitsmarktregion* or AMR from here onwards); we next rank AMR by size and calculate the cumulative share out of aggregate employment: all AMR situated above the .50 cumulative share threshold classify as large, all the ones situated below are considered as small. According to this definition there are 9 large Metropolitan Areas and 63 small, where the large ones are: 1) Hamburg, 2) München, 3) Köln-Leverkusen-Bergisch Gladbach, 4) Frankfurt am Main-Offenbach am Main-Hanau, 5) Stuttgart-Sindelfingen-Esslingen am Nemkar-Ludwigsburg, 6) Düsseldorf-Neuss, 7) Hannover, 8) Erlangen-Fürth-Nürnberg, and 9) Duisburg-Essen-Mülheim an der Ruhr-Oberhausen-Moers-Bottrop-Gelsenkirchen-Remklinghausen-Bochum-Dortmund-Hagen-Hamm-Herne-Witten. This grouping favors a stricter definition for a large city when it is compared to a grouping scheme based on median

we also run the same specification on the log of average establishment size in a given city/industry. Fig.2.1 provides a map of Metropolitan Areas identified by Eurostat and located in former West Germany.

Idiosyncratic Uncertainty Estimation

We use employment data on the population of establishments located in the former West Germany Metropolitan Areas to construct a proxy for idiosyncratic uncertainty.[16]

Backing-out Idiosyncratic Shocks It is useful to describe under what assumptions concerning preferences, market structure and the production function data on employment alone can be used to extract a measure of establishment-specific residual shock. Suppose that establishment n, producing a variety of good k (5-digit industry) and selling to market m in year t solves the profit maximization problem:

$$\begin{array}{c} \max_{P_{nmkt}} \Pi_{nmkt} = \left(P_{nmkt} - MC_{nmkt}\right) Q_{nmkt} \\ s.t.: \quad Q_{nmkt} = f\left(P_{mkt}, P_{nmkt}, Y_{mt}\right) \end{array} \tag{1.2}$$

where P_{mkt} equals a price index for varieties of good k sold to market m in year t, Y_{mt} is the income of consumers located in market m in year t. Given optimal output choice labor demand is pinned down by the production function by $Q_{nmkt} = g(Z_{nmkt}, L_{nmkt}, M_{nmkt})$, where Z_{nmkt} is efficiency of producer n and M_{nmkt} is an input that is combined with labor to produce. If producers face isoelastic demand, i.e. $P_{nmkt} = \mathcal{M} \times MC_{nmkt}$, and unit elasticity of substitution between inputs, i.e., $\frac{\partial \ln(M_{nmkt}/L_{nmkt})}{\partial \ln(MPM_{nmkt}/MPL_{nmkt})} = 1$, then, information on employment can be used to extract a measure of establishment-specific residual shock. Under CES demand with elasticity of substitution $\rho > 1$ and labor as the only input for simplicity, labor demand of establishment n is:

$$L_{nmkt} = \text{constant} \times \left(Z_{nmkt}^{\rho-1} \widetilde{Y}_{mkt}\right) \tag{1.3}$$

city size. On the other hand, it is desirable as it allows to have a similar sized sample of establishments on both sides of the threshold, while the alternative method would yield a smaller sample size for establishments located in small cities.

[16]Previous papers have focused on other measures of firm performance, such as sales or value added (e.g., di Giovanni et al. (2014)). These measures have the advantage of a closer tie with firm profits. Suppose that a firm product receives a positive demand shock, but it is not able to adjust its productive capacity immediately. In this case, employment would be an uninformative metric while sales would still be informative about the shock that would be absorbed by a change in prices.

with $\widetilde{Y}_{mkt}$ capturing both the cost of labor and the expenditure in market m on good k at time t.[17]

Log-linearizing and first-differencing eq.1.3 yields:

$$g_{nmkt} = \ln \widetilde{Y}_{mk,t} - \ln \widetilde{Y}_{mk,t-1} + (\rho - 1)\left[\ln\left(Z_{nmk,t}\right) - \ln\left(Z_{nmk,t-1}\right)\right] \tag{1.4}$$

A regression of establishment growth rates on industry/market/year fixed effects (e.g., Gabaix (2011), di Giovanni et al., (2014)) is not sufficient for idiosyncratic fluctuations to have a close interpretation to the structural shocks: deterministic growth based on establishment age (Haltiwanger et al. (2013)) and rejection of the Gibrat's Law among small establishments (Hall (1987)) are consolidated facts in the applied literature on firm dynamics. For example, an entrepreneur pondering whether to open a new plant is aware that her growth perspectives are higher during the first years from the investment: for this reason, higher volatility does not qualify as higher uncertainty. Suppose that the data generating process for establishment productivity consists of a mixture between a AR(1) term featuring a plant-specific intercept, an age-dependent term and a disturbance, $\nu_{nmk,t}$:

$$z_{nmk,t} = (1 - \theta_1)\, z_{nmk} + \theta_1 z_{nmk,t-1} + \theta_2 \text{age}_{nmk,t} + \nu_{nmk,t} \tag{1.5}$$

with $0 < \theta_1 < 1$ and $\theta_2 > 0$. Approximating the first-difference of eq.1.5 non-parametrically yields:

$$z_{nmk,t} - z_{nmk,t-1} = \sum_{s=1}^{4} \widetilde{\beta}_s D_{s,nmkt} + \sum_{a=1}^{4} \widetilde{\beta}_a D_{a,nmkt} + \widetilde{\varepsilon}_{nmk,t} \tag{1.6}$$

where $D_{s,nmkt}$ and $D_{a,nmkt}$ are dummy variables for plant size and age at time t.[18] We combine eq.1.4-1.6 and estimate:

$$g_{nmkt} = \gamma_{mkt} + \sum_{s=1}^{4} \beta_s D_{s,nmkt} + \sum_{a=1}^{4} \beta_a D_{a,nmkt} + \varepsilon_{nmk,t} \tag{1.7}$$

[17] Giroud and Mueller (2017) have recently shown the importance of geographical risk-sharing for multi-establishments firms. Adoption of the establishment as the unit of analysis provides therefore an incomplete picture of the strategy employed by firms to handle risk. The concern that especially service industries have undergone in recent years a shift towards multi-unit firms (e.g. Foster et al. (2006)) is attenuated if plants belonging to the same firm share similar characteristics, as we do control for differences in *both* average *and* dispersion of growth rates based on observable plant characteristics.

[18] Age $\in \{0, 1-4, 5-9, 10-15, 16+\}$ years and size $\in \{1-7, 8-24, 25-99, 100+\}$ employees. These thresholds correspond to 4 quartiles of the aggregate age and size distribution. Plant size is calculated according to Davis et al. (1998) as the mean between plant size at time t and time $t-1$.

where γ_{mkt} are Metropolitan Area/5-digit industry/year fixed effects and $\varepsilon_{nmk,t} \sim (0, \sigma_k^2)$. The objective is to back-out an estimate of σ_k, the standard deviation of unforeseeable proportional changes in establishments' idiosyncratic productivity (or demand) component.

We estimate eq.1.7 separately for each sector of the economy on the sample of continuing establishments after having replaced the standard growth rate with the definition provided by Davis et al. (1998) more robust to outliers, i.e. $g_{nmkt} = \frac{2(l_{nmkt} - l_{nmkt-1})}{(l_{nmkt} + l_{nmkt-1})}$.[19]

Idiosyncratic Uncertainty Measurement The baseline proxy for idiosyncratic uncertainty in industry k is given by:

$$\text{Uncertainty}_k = \hat{\sigma}_k = \sqrt{\frac{1}{N_k}\sum_{i=1}^{N_k} \hat{\varepsilon}_{nmkt}^2} \tag{1.8}$$

where N_k is the number of observations in industry k during 1995-2005, and $\hat{\varepsilon}_{nmkt}$ are the shocks estimated in the previous stage.[20] One concern is that the estimator in eq.1.8 may capture differences across industries regarding age and size distribution of establishments when different groups are subject to shocks having different variance on average (e.g., young establishments tend to experience very dispersed growth rates in all industries). Alternatively, it may reflect differences across industries regarding the variance of shocks to which different groups of establishments are subject (e.g., young establishments tend to experience very dispersed growth rates especially in manufacturing). We therefore also perform a shift/share analysis and construct:

$$\hat{\sigma}_k|_{\text{Fixed Share}} = \sqrt{\textstyle\sum_g \text{share}_g \hat{\sigma}_{k,g}^2} \quad \hat{\sigma}_k|_{\text{Fixed Uncertainty}} = \sqrt{\textstyle\sum_g \text{share}_{k,g} \hat{\sigma}_g^2} \tag{1.9}$$

$$\text{share}_g = \frac{\frac{1}{K}\sum_k N_{k,g}}{\sum_g' \left(\frac{1}{K}\sum_k N_{k,g'}\right)} \quad \text{share}_{k,g} = \frac{N_{k,g}}{\sum_g' N_{k,g'}}$$

[19] These sectors are: "Manufacturing", "Construction", "Wholesale and Retail Trade", "Hospitality", "Finance", "Real Estate", "Business Services", "Other Services". We estimate several variants of eq.1.7. These include: a saturated regression including the interaction between all age and size dummies; a regression including the interaction of age and size dummies in eq.1.7 with a dummy indicating whether the market is large or small; the combination of the two. We have also estimated eq.1.7 separately for each 3-digit industry in the economy. Being the results mostly unchanged, we retained the most parsimonious specification.

[20] Cross-sectional dispersion in the growth rate of firm-level performance indicators is often used in the finance and macroeconomics literature as a proxy for idiosyncratic uncertainty (Bloom (2009)). Davis et al. (2007) construct a time-series of both cross-sectional dispersion and firm-level volatility for U.S. firms and they find the two measures to closely track each other.

$$\hat{\sigma}^2_{k,g} = \sqrt{\frac{1}{N_{k,g}} \sum_{i=1}^{N_{k,g}} \hat{\varepsilon}^2_{nmkt,g}} \quad \hat{\sigma}^2_g = \frac{1}{K} \sum_k \hat{\sigma}^2_{k,g}$$

Variation by $\hat{\sigma}_k|_{\text{Fixed Share}}$ is driven by differences in risk across groups weighted according to a common scheme; conversely, variation by $\hat{\sigma}_k|_{\text{Fixed Uncertainty}}$ is explained by compositional differences. We consider $g \in \{Y/S, Y/L, O/S, O/L\}$, where a plant is classified as small (S) if it has less than 5 employees and large (L) otherwise, young (Y) if it is below 10 years of age and old (O) otherwise [21]

One might be concerned with the consistency of the estimator in eq.1.8: the unrestricted sample consists of all city/industry/year groups, including those characterized by as few as only 2 establishments. In these instances it is impossible to accurately estimate the common shock γ_{mkt}, and the estimator in eq.1.8 would likely cause an upward bias in the coefficients of eq.1.1. A remedy would consist of adopting a pooled variance estimator that applies a standard degree of freedom correction to account for estimation of within-group sample means. In Appendix 1.5 we describe in detail the pooled variance estimator. Estimation results show that a simpler correction consisting of the elimination of groups with less than 30 observations from the sample seems if anything to over-correct the small sample bias. Therefore, the estimates of the coefficients in eq.1.1 are derived based on a measure of risk calculated for a restricted sample consisting of city/industry/year groups featuring at least 30 establishments.

1.2.2 Data

Data Description

We use the Administrative Wage and Labor Market Flow Panel (AWFP) to estimate eq.1.7 (Seth and Stüber (2017)). These data comprise of the population of establishments filing for social security payments with at least one registered regular employee. We focus on continuing establishments that are active in districts located in former West-Germany between 1995 and 2005 and that never relocated.[22] Additionally, we source information on industry-specific characteristics from AMADEUS firm-level data (Wharton Research Data Services), e.g., labor share, fixed assets to output ratio; and the Expectations Business Panel (2016) distributed by the Economics Business and Data Centre at the CESifo, e.g., predominant ownership type of firms.

[21] We restrict throughout the attention to 5-digit industries for which we observe $N_{mkt,g} > 100$ establishments in each group.

[22] The data are collected by the Institute for Employment Research (IAB, Institut für Arbeitsmarkt- und Berufsforschung). A description of the public release data of the AWFP is available in Seth and Stüber (2017).

We report in Fig.1.2 the growth rate of national employment according to the National Accounts (DESTATIS) vs. the growth rate of total employment in the selected IAB sample after having excluded the following sectors: Agriculture and Mining, Utilities and Transportation, Public Administration. The correlation is .79 and strongly significant.

Industry Selection

We use the German Classification of Economic Activities, Edition 1993, at the 5-digit level of disaggregation to define a good. Since there exists no harmonized classification at the 5-digit level over the entire sample period considered, we use the 5-digit industry classification available between 1998 and 2003 and impute the industry category for establishments operating in the remaining years.[23] The final sample consists of 6.043.000 establishment/year observations located in urban areas in sectors other than Agriculture and Mining, Utilities and Transportation and Public Administration.

Next, we define the subset of non-tradable 5-digit industries we use to estimate eq.1.7. We construct an index of geographical concentration for each 5-digit industry as in Ellison and Glaeser (1997):[24]

$$\begin{gathered}\text{Geog.Conc.}_k^{EG} = \frac{G_k - \left(1 - \sum_m x_m^2\right) H_k}{\left(1 - \sum_m x_m^2\right)(1 - H_k)} \\ G_k = \sum_m \left(\frac{empl_{mk}}{\sum_{m'} empl_{m'k}} - \frac{\sum_{k'} empl_{mk'}}{\sum_{m'} \sum_{k'} empl_{m'k'}} \right)^2 \\ H_k = \sum_{n=1}^{N_k} x_{nk}^2 \end{gathered} \tag{1.10}$$

where $m \in \{1, 2, ..., 72\}$ and $k \in \{1, 2, ..., 862\}$, all 5-digit industries in the sample with the exclusion of Education and Health, x_m being Metropolitan Area m share of total employment, and x_{nk} plant n share of industry k total employment. Next, we define a 5-digit industry as non-tradable if it satisfies two criteria: 1) Geog.Conc.$_k^{EG}$ is below the median value of the cross-industry distribution, 2) the industry features positive establishment count for each city in the sample in year 2000. The set of industries that fulfils these criteria and the one on minimum establishment count described in the previous section comprises of 73 industries. The full list is reported in Appendix 2.6.

As a robustness check, we also consider the index of geographical concentration suggested in Jensen and Kletzer (2005). The index described in eq.1.11 aims at avoiding the mistake of classifying a given industry as geographically

[23] Details on the imputation procedure can be found in Appendix 1.5.

[24] This index corrects for geographical concentration induced by lumpiness in establishment employment shares.

concentrated while, in fact, it provides non-tradable intermediate goods to a geographically concentrated downstream industry:

$$\begin{aligned} \text{Geog.Conc.}_k^{JK} &= \sum_m \left(\frac{empl_{mk}}{\sum_{m'} empl_{m'k}} - IDS_{mk} \right)^2 \\ IDS_{mk} &= \sum_{k'} \frac{Y_{k,k'}}{Y_k} \frac{empl_{mk'}}{empl_{k'}} \end{aligned} \tag{1.11}$$

where k' is the downstream industry, $\frac{Y_{k,k'}}{Y_k}$ is the fraction of output in industry k being consumed by industry k', and $\frac{empl_{mk'}}{empl_{k'}}$ is employment of industry k' located in city m. IDS_{mk} captures Metropolitan Area m demand of industry k, and it is based on Metropolitan Area m employment share across all downstream industries rescaled by intermediate input usage. We can construct this index only at the 2-digit industry-level, the classification featured in the input-output tables provided the German National Statistical Office (DESTATIS). We calculate Geog.Conc.$_k^{EG}$ and Geog.Conc.$_k^{JK}$ for all 2-digit industries and restrict the attention to those such that the employment share in previously defined non-tradable 5-digit industries is $\geq 50\%$. The correlation stands at around 47% and statistically significant.

We therefore retain the classification based on eq.1.10 and show throughout the analysis how the results change when we restrict the focus to industries located exclusively in the retail/restaurants and bars business (Mian and Sufi (2014)).

1.2.3 Results

Estimates of Idiosyncratic Uncertainty

We calculate residual dispersion by estimating eq.1.7 separately for each macro-sector of economic activity. We first estimate a variant of eq.1.7 that excludes city/5-digit industry/year dummies (Measure#1) and report in Table 1.1 the minimum, average and maximum estimated coefficient on each age/size dummy together with the fraction of sector-level regressions (8 in total) for which each coefficient, in turn, is significant. It is evident that young establishments tend to grow faster on average (upper panel), and that there is a substantial amount of mean reversion for small establishments (lower panel).

Next, we estimate a variant of eq.1.7 that excludes age/size dummies (Measure#2). Finally, we estimate eq.1.7 (Measure#3). We report in Table 1.2 the R-squared from each specification for each macro-sector of economic activity. We compare the selected vs. the excluded sample of industries along several industry-wide characteristics, e.g., labor/intermediate input intensity, capital/output ratio (Table 1.3). Not surprisingly, the selected industries are more labor intensive and feature lower levels of investment than the excluded ones.

We then move to the set of predominantly non-tradable industries. We report in Table 1.4 descriptive statistics of estimated idiosyncratic uncertainty for all vs. specific groups of establishments. Once again not surprisingly, within-group dispersion of residuals is higher for young and small establishments. Furthermore, cross-industry variation is also the highest for this group of establishments.

Finally, we compare the baseline measure of uncertainty with the one constructed starting from information on business-specific short-term forecast error on business volume provided by the Expectations Panel distributed by the Economics and Business Data Centre (2016) for Retail and Wholesale Trade and Business Services.[25] The Pearson coefficient of correlation is .45 and statistically significant (Fig.1.3).

Regression Results

Table 1.6 contains the main results. An increase by 1 basis point in residual dispersion translates into $+.5$ establishments in small and $+3$ establishments in large markets. The non-interacted coefficient is 6 times larger when narrowing down the focus to industries located in retail and restaurants. At the same time, an increase by 1 basis point in residual dispersion translates into establishments being on average 1.5% smaller, 4.7% in retail and restaurants. The interaction coefficient is positive for the selected sample of industries and negative for industries located in retail and restaurants, and not significant in either case.

Given that idiosyncratic uncertainty can vary during the life-cycle of the productive unit, we next ask during what stage of the life-cycle is an establishment less negatively affected by idiosyncratic uncertainty. We report in Table 1.7 the estimated coefficients in eq.1.1 based on the estimator described in eq.1.9. The positive correlation between the number of establishments and idiosyncratic uncertainty is driven cross-industry differences in uncertainty when the establishment is either young and small or old and large. This result is intuitive: on the one hand, higher uncertainty raises the payoff associated with starting a new business, but on the other, it also increases the probability of exiting the market. Small and young establishments are the ones for which the first margin is the strongest, large and old establishments are the ones for which the last margin is the weakest, thus making these two groups the most sensitive to changes in uncertainty.

We next test whether the baseline result highlighted in Table 1.6 is driven by differences across industries in the composition of the establishment pool or by differences in uncertainty for different groups of establishments. We consider a

[25]The construction of industry-specific idiosyncratic uncertainty follows Bachmann et al. (2013) and it is only reliable at the 3-digit level of analysis due to limited sample size. We thus collapse the baseline uncertainty measures at the 3-digit level of detail by using 5-digit industry employment shares as weights.

shift/share analysis as described in eq.1.9 and report in Table 1.8 the estimated coefficients in eq.1.1 based on $\hat{\sigma}_k^2$ (column 1), $\hat{\sigma}_k^2|_{\text{Fixed Uncertainty}}$ (column 2) and $\hat{\sigma}_k^2|_{\text{Fixed Share}}$ (column 3), for both all selected non-tradable industries and industries in the retail or restaurants business only. While for the extended sample of industries the positive and significant association between idiosyncratic uncertainty and number of establishments seems to be driven mostly by compositional differences (column 2, upper panel), for industries in the retail or restaurants business differences in uncertainty have explanatory power even after having conditioned on industry composition.

We also consider an alternative estimator for uncertainty based on establishment-level *residual* volatility. To this end, we follow Davis et al. (2007) and calculate establishment-level volatility considering either all establishments (column 1 Table 1.10) or only establishments surviving throughout 1995-2005 (column 2 Table 1.10).[26] The results are robust to this alternative measure of uncertainty. We finally consider different years for 1) the construction of residual dispersion (column 1-2 of Table 1.9), 2) the cross-section of establishment counts employed in eq.1.1 (column 3-4 of Table 1.9). The results are robust in both cases.

Finally, we plot in Fig.1.4 the growth rate in idiosyncratic uncertainty between 2000-2002 against the growth rate in the number of establishments between 1998-2000 after having taken sectoral fixed effects. First, the correlation is positive and significant with the exclusion of an outlier, thus implying that the industries where uncertainty has declined the most are also the ones that have seen the most sluggish growth in the number of establishments. Second, changes in uncertainty lag changes in the number of establishments: a plot of the change between 1998-2000 in uncertainty against the change between 1998-2000 in the number of establishments (not reported here) delivers a zero correlation between the two variables. In other words, entrepreneurs deciding whether to enter today are more concerned with the evolution of future uncertainty rather than with its current level. The same analysis repeated only on the sample of establishments located in large cities delivers even a stronger correlation, which is in line with the interaction effect found when working with levels. Finally, the evidence presented in Fig.1.4 is consistent with the idea that the changing distribution of establishments across industries over time reflects the changing distribution of entrants.

[26]We construct a measure of idiosyncratic volatility by 1) weighting the squared deviation for establishment n at time t in proportion to its size at t relative to its average size during the period considered, 2) applying a degrees of freedom correction to take into account differences in the length of establishment-specific panels.

1.2.4 Robustness

Heterogeneity across Industries

Economic theory suggests that the sign and magnitude of the correlation between number of establishments and idiosyncratic uncertainty across industries and city/industries should depend on a battery of industry-specific characteristics. In this section, we consider two of them, the sunk entry cost and the predominant ownership type. Industry-specific sunk cost is constructed as in Syverson (2004): we take the ratio of fixed assets to output in the industry and multiply it by the employment share of a median-sized establishment. We use data from the Expectations Business Panel (2016) distributed by the Economics Business Data Centre at the CESifo to construct a proxy for predominant ownership type. We classify industries into "prevalence of limited liability ownership types" vs. "prevalence of public owned". Specifically, we calculate for each industry the fraction of publicly owned firms in a given year (2016) and divide industries into PUBLIC = 1 if the fraction is above the cross-industry median and PUBLIC = 0 otherwise.[27]

Results are reported in Table 1.12. While the non-interacted coefficient is always positive and statistically significant across all specifications, the interacted one loses significance when we control for the sunk entry cost. As expected, idiosyncratic uncertainty has a negative impact on the number of establishments when it is coupled with a high sunk entry cost (column 2) or public ownership structure type (column 4).

Heterogenous and Endogenous Loadings on Aggregate Shocks

The empirical analysis hinges on the assumption of employment being isoelastic with respect to productivity. This assumption is violated in models featuring endogenous markups (e.g., Melitz and Ottaviano (2008), Atkenson and Burstein, (2008). We run the following test aimed at detecting whether the empirical findings of this paper are predominantly attributable to endogenous market power. We

[27] We define as "publicly owned" a firm that has ownership type in either of the four categories: "Public Utilities", "Aktiengesellschaft (AG)" or corporation, "Aktiengesellschaft & Co. Kommanditgesellschaft auf Aktien (AG & Co. KGaA)" or limited partnership by shares with a corporation as general partner, "AG & Co. KGaA" with public limited partnership. All other types, primarily Gesellschaft mit beschränkter Haftung (GmbH) or company with limited liability, are classified as non-public.

consider a variant of eq.1.7 as suggested in di Giovanni et al. (2014):

$$g_{nmkt} = \gamma_{mkt} + \sum_{s=1}^{4} \beta_s D_{s,nmkt} \times \gamma_{mkt} + \sum_{a=1}^{4} \beta_a D_{a,nmkt} \times \gamma_{mkt} + \\ + \sum_{s=1}^{4} \beta_s D_{s,nmkt} + \sum_{a=1}^{4} \beta_a D_{a,nmkt} + \varepsilon_{nmk,t} \quad (1.12)$$

The point of departure consists of interacting city/industry/time fixed effects with age/size dummies, in order to account for endogenous responsiveness to aggregate shocks based on observable characteristics. Since the responsiveness to shocks decreases in the establishment market share we further drop all city/industry/cells where we observe less than $N < 500$ establishments. Results are reported in Table 1.13. The non-interacted coefficient is now halved but still significant at the 10% level.

1.3 Model

1.3.1 Geography and Population

There exists a continuum of locations denoted by $m \in M$ and characterized by some exogenous amenities A_m. Total population is equal to $\overline{L}$ and consumers/workers are perfectly mobile. Each worker receives upon choosing a location h_m location-specific efficiency units of labor.

1.3.2 Preferences and Demand

Preferences are defined over $K + 1$ goods. One of them is homogenous and it is fully tradable while the other K goods are non-tradable and horizontally differentiated. There exists an endogenously determined measure of varieties produced in equilibrium for each differentiated good k denoted by N_{mk}. Utility in location m can be expressed in the generic form:

$$U_m^i = g\left(A_m; \widetilde{U_m^i}\right) \quad (1.13)$$

where $\widetilde{U_m^i}$ denotes the utility accruing from the consumption of commodities.

$\widetilde{U_m^i}$ is additive in two lower utility nests: one that is linear in consumption of the homogenous good and one that is non-additive and quadratic in consumption of the differentiated goods as in Ottaviano, Tabuchi and Thisse (2002). Consumer i choosing to locate in city m solves the utility maximization problem:

$$\max_{q_0^i, q_{mkn}^i \forall k,n} \widetilde{U_m^i} = q_0^i + \sum_{k \in K} \left[\alpha \int_{N_{mk}} q_{mkn}^i dn - \frac{\gamma}{2} \int_{N_{mk}} (q_{mkn}^i)^2 dn - \frac{\eta}{2} \left(\int_{N_{mk}} q_{mkn}^i dn \right)^2 \right] \tag{1.14}$$

subject to the budget constraint:

$$\omega_m + \tau = p_0 q_0^i + \sum_{k \in K} \int_{N_{mk}} p_{mkn} q_{mkn}^i dn + r_m \tag{1.15}$$

where τ denotes the symmetric share of each individual into a mutual fund owning the totality of the housing stock in the economy, and $\omega_m = \omega_0 h_m$ corresponds to location-specific labor income, with ω_0 being the wage per efficiency unit of labor. In order to move to location m the consumer must pay rent r_m.

The inverse demand for variety n of good k is:

$$\alpha - \gamma q_{mkn}^i - \eta Q_{mk}^i = \lambda p_{mkn} \qquad \forall m \in M, k \in K, n \in N_{mk} \tag{1.16}$$

with $Q_{mk}^i = \int_{N_{mk}} q_{mkn}^i dn$ denoting total demand of good k by consumer i living in city m. Demand of the homogenous good is:

$$1 = \lambda p_0 \tag{1.17}$$

The homogenous good is chosen to be the numeraire and $\lambda = 1$ follows.

The parameters governing this utility specification have the following interpretation. Both α and η affect the substitutability between homogenous and differentiated goods. An increase in α raises demand for each variety of the differentiated goods, while an increase in η reduces demand for each variety of the differentiated goods in proportion to total individual consumption across varieties of each differentiated good: intuitively the consumer prefers consuming less of any given good as her total level of consumption rises. Finally, γ parametrizes the degree of differentiation of varieties of good k: an increase in γ causes the consumer to care more about the individual consumption level of each variety and less about the total number of varieties she is consuming.

The crucial difference between this demand system and the classic CES one is the variable price elasticity of demand:

$$|\varepsilon_{mkn}| = \frac{p_{mkn}}{\frac{\alpha\gamma + \eta N_{mk} \overline{p}_{mk}}{\gamma + \eta N_{mk}} - p_{mkn}} \tag{1.18}$$

The absolute value of ε_{mkn} increases (decreases) and at a convex rate in the price charged for each variety. Additionally, it increases as total individual consumption rises, hence as $\overline{p}_{mk}$ declines and N_{mk} increases. When the average price in the market is lower and/or there exists a larger measure of varieties for a given good, firms producing the same good must charge lower markups.

Aggregation across all goods yields the indirect utility from consumption of all commodities:

$$\widetilde{V_m^i} = \omega_m + \tau - r_m + \sum_{k \in K} \left\{ \frac{N_{mk}}{2} \left[\frac{1}{\gamma + \eta N_{mk}} \left(\alpha - E(p_{mkn})\right)^2 + \frac{1}{\gamma} \sigma^2(p_{mkn}) \right] \right\} \tag{1.19}$$

with $E(p_{mkn})$ and $\sigma^2(p_{mkn})$ being respectively the mean and the variance of prices for varieties of good k sold in market m. Consumer surplus is therefore increasing and concave in the measure of varieties since the consumer gets increasingly satiated as the variety set expands. Moreover, it is decreasing and convex in the average price for the varieties of each good as a lower price triggers a more than proportional response in consumption due to the endogeneity of demand elasticity. Finally, it is increasing in the dispersion of prices.

For consumption of the homogenous good to be strictly positive it must be:

$$\omega_m + \tau > \sum_{k \in K} \int_{N_{mk}} p_{mkn} q_{mkn}^i dn + r_m \tag{1.20}$$

in each local product market $m \in M$.

1.3.3 Production and Firm Behavior

The homogenous good The homogenous good is produced by a representative firm that operates in a perfectly competitive environment with CRTS technology:

$$q_{0m} = l_{0m} \tag{1.21}$$

It absorbs residual units of labor not demanded by producers operating in the differentiated goods sector:

$$l_{m0}^d = h_m L_m - \sum_{k \in K} l_{mk}^d \tag{1.22}$$

and charges:

$$p_0 = mc_0 \tag{1.23}$$

with $mc_0 = \omega_0$ after normalization of the marginal product of labor to 1 without loss of generality. The choice of the homogenous good as the numeraire good implies $\omega_0 = 1$ and $\omega_m = h_m$ in city m.

The differentiated goods Firms operating in the differentiated sector are monopolistically competitive and heterogenous. They must pay a sunk cost f in order to start producing: once they have entered they operate according to a linear technology and constant marginal cost c_{mkn}. The problem of a firm n producing a variety of good k in city m consists of choosing the price that maximizes profits subject to total consumer demand:

$$\max_{p_{mkn}} \pi_{mkn} = (p_{mkn} - c_{mkn})\, q_{mkn} \tag{1.24}$$

subject to the demand function in eq.1.16.

The marginal cost consists of a location-specific component, $\widetilde{\mu_m}$, and a stochastic component, ν_{mkn}, $c_{mkn} = \widetilde{\mu_m} + \nu_{mkn}$, where ν_{mkn} is independently drawn across firms from a k-specific distribution characterized by first and second moment $\nu_{mkn} \sim (\mu_k, \sigma_k^2)$. We define the mean of marginal cost for a firm operating in market m and good k as $\mu_{mk} = \widetilde{\mu_m} + \mu_k$.[28]

Using the first order condition:

$$q_{mkn} + (p_{mkn} - c_{mkn}) \frac{\partial q_{mkn}}{\partial p_{mkn}} = 0 \tag{1.26}$$

into consumer demand yields the following solution to the price charged, the quantity produced, labor demand in terms of efficiency units of labor, profits made by producer n:

$$p_{mkn} = \tfrac{1}{2}\left(\tfrac{\alpha\gamma+\eta N_{mk}\overline{p}_{mk}}{\gamma+\eta N_{mk}} + c_{mkn}\right) \qquad q_{mkn} = \tfrac{L_m}{2\gamma}\left(\tfrac{\alpha\gamma+\eta N_{mk}\overline{p}_{mk}}{\gamma+\eta N_{mk}} - c_{mkn}\right)$$

$$l_{mkn} = \tfrac{L_m}{2\gamma}\left(\tfrac{\alpha\gamma+\eta N_{mk}\overline{p}_{mk}}{\gamma+\eta N_{mk}} - c_{mkn}\right) c_{mkn} \qquad \pi_{mkn} = \tfrac{L_m}{4\gamma}\left(\tfrac{\alpha\gamma+\eta N_{mk}\overline{p}_{mk}}{\gamma+\eta N_{mk}} - c_{mkn}\right)^2$$

with $\overline{p}_{mk} = \frac{\alpha\gamma+(\gamma+\eta N_{mk})\mu_{mk}}{2\gamma+\eta N_{mk}}$.

1.3.4 Free-Entry

Linearity of demand implies that for given regions of the parameter space demand for an individual variety producer can be negative (Melitz and Ottaviano (2008)).

[28] Idiosyncratic uncertainty in this economy concerns the technology with which a firm operates. It is straightforward to include demand-side uncertainty as in Foster, Haltiwanger and Syverson (2008). In this case the firm-optimality condition becomes:

$$p_{mkn} = \tfrac{1}{2}\left(\tfrac{\alpha\gamma+\eta N_{mk}\overline{p}_{mk}}{\gamma+\eta N_{mk}} + \epsilon_{mkn} + c_{mkn}\right) \quad q_{mkn} = \tfrac{L_m}{2\gamma}\left(\tfrac{\alpha\gamma+\eta N_{mk}\overline{p}_{mk}}{\gamma+\eta N_{mk}} + \epsilon_{mkn} - c_{mkn}\right) \tag{1.25}$$

with ϵ_{mkn} being idiosyncratic taste for variety n of good k sold to market m. When $E(\epsilon_{mkn}) = 0$ and $\sigma^2_{\epsilon,k} + \sigma^2_{c,k} - 2\sigma_{\epsilon c,k} = \sigma^2_k$, then $c_{mkn} - \epsilon_{mkn} \sim \left(\mu_k, \sigma_k^2\right)$, so that in the model, as in the data, demand and supply-driven idiosyncratic uncertainty are perfectly confounded.

Conditional on the distribution for the marginal cost draws in industry k being inverse Pareto:

$$1/\nu_{mkn} \sim Pareto(1/b_k, \theta_k) \tag{1.27}$$

where b_k is the maximal marginal cost draw and θ_k is the shape parameter of the associated productivity distribution. The solution to the free-entry equilibrium is given by:

$$\begin{aligned} p^*_{mk} \leq b_k \Leftrightarrow (c^*_{mk}, N) &: \begin{cases} \frac{L_m}{4\gamma}\int_0^{b_k}\left(p^*_{mk} - c_{mkn}\right)^2 dG_k(c) = f \\ p^*_{mk} = c^*_{mk} = \frac{2\alpha\gamma + \eta N_{mk} E\left(c_{mkn} | c_{mkn} \leq c^*_{mk}\right)}{2\gamma + \eta N_{mk}} \end{cases} \\ p^*_{mk} > b_k \Leftrightarrow N &: \begin{cases} \frac{L_m}{4\gamma}\int_0^{b_k}\left(p^*_{mk} - c_{mkn}\right)^2 dG_k(c) = f \\ p^*_{mk} = \frac{2\alpha\gamma + \eta N_{mk} E(c_{mkn})}{2\gamma + \eta N_{mk}} \end{cases} \end{aligned} \tag{1.28}$$

where p^*_{mk} is the choke-price above which consumer demand is negative.

In the absence of first-hand evidence on the endogenous selection cutoff, we focus on a region of the parameter space for which selection does not arise. Notice, however, allowing for an endogenous selection cutoff would nevertheless be compatible with a positive relationship between variety and idiosyncratic uncertainty. The free-entry equilibrium is standard: potential entrants base their decision upon the option value of entry, defined by the expectations of profits, and compare it against a sunk cost paid in terms of efficiency units of labor and potentially location-specific (Arkolakis (2010)):

$$V^n_{mk} = \frac{L_m}{4\gamma}\left\{\left[\frac{2\gamma\left(\alpha - \mu_{mk}\right)}{2\gamma + \eta N_{mk}}\right]^2 + \sigma^2_k\right\} - \widetilde{f_m} \tag{1.29}$$

where V^n_{mk} denotes the expected value of holding a business in market m for good k (the superscript n is meant to distinguish it from the indirect utility attached to location m for consumer i).

Under free-entry in the market for good k in city m the measure of producers operating in equilibrium satisfies:

$$E\left(\pi_{mk}\right) = \widetilde{f_m} \tag{1.30}$$

The assumption of ex post heterogeneity implies that firms make positive profits in equilibrium: by free-entry cumulative profits are $\sum_{k\in K}\int_{N_{mk}}\pi_{mkn}dn = \sum_{k\in K} N_{mk}\widetilde{f_m}$. By eq.1.18 an increase in the number of competitors causes the elasticity of demand for an individual variety to increase. Thus, firms expected profits net of the sunk cost are dissipated through the effect of increased competition on the markups charged by firms. A larger customer base induces more entrepreneurs to enter. A decrease in firm-level uncertainty entails an increase in average marginal cost and a decrease in marginal cost dispersion, as explained in Section 1.4. The deterioration in the productivity distribution is matched by a decline in variety. The decline is more sizeable in larger markets due to the diminishing marginal impact of competition on firm profits.

1.3.5 Partial Equilibrium

For the purpose of comparative statics, we abstain from spatial equilibrium and take the distribution of city sizes as given.

Measure of Varieties Consider the solution to the equilibrium measure of varieties in industry k and market m:

$$N_{mk} = \frac{2\gamma}{\eta}\left[\frac{(\alpha - \mu_{mk})}{\sqrt{4\gamma f_m - \sigma_k^2}} - 1\right] \tag{1.31}$$

after having redefined $f_m = \frac{\widetilde{f_m}}{L_m}$.[29]

An increase in α is associated with higher expected profits, and therefore a larger measure of firms, since this parameter shifts demand towards the differentiated good. An increase in γ lowers expected profits since more product differentiation grants low-cost producers larger profits and high-cost producers lower profits (see also Syverson (2004)). By profit convexity, the impact on expected profits of an increase in γ is negative. Finally, an increase in η reduces expected profits since it magnifies the impact of competition on consumer demand elasticity. The comparative static with the sunk cost is as standard. It follows that the negative effect of less idiosyncratic uncertainty on the measure of varieties is more pronounced when α (γ, η and f_m) is high (low).

Average Firm Size Consider the solution for the log of equilibrium average firm size in industry k and market m:

$$\ln E(L_{mkn}) = \ln \frac{L_m}{\eta h_m} + \ln\left(\mu_{mk} - \sigma_k^2 \lambda_{mk}\right) - \ln \lambda_{mk} \tag{1.33}$$

[29] In equilibrium it is ensured that $4\gamma f_m - \sigma_k^2 > 0$ and:

$$\frac{2\gamma}{\eta}\left[\frac{(\alpha - \mu_{mk})}{\sqrt{4\gamma f_m - \sigma_k^2}} - 1\right] < \frac{2\gamma}{\eta}\frac{\alpha - \widetilde{\mu_m} - b_k}{b_k - \mu_k} \tag{1.32}$$

$\forall m$ and k for no endogenous selection to arise. Section 1.4 contains more details on the parametrization chosen to ensure that these restrictions are satisfied.

with $\lambda_{mk} = (4\gamma f_m - \sigma_k^2)^{-1/2}$ summarizing the intensity of competition.

Firm size can either increase or decrease in marginal cost: while higher marginal cost lowers demand, it also increases labor requirement per unit of production. Hence, the effect on firm size of variation in marginal cost need not be monotonous. In equilibrium the second effect is always dominating and firms are on average larger when they are on average less productive (high μ_{mk}). Furthermore, firms are on average smaller when a) there is high product differentiation (low γ) by the positive effect on competition, b) the pro-competitive channel is stronger (high η), c) the sunk cost is lower, d) they operate in an industry characterized by high dispersion of marginal cost draws (high σ_k^2) due to ensuing fiercer competition and the concavity of firm size in marginal cost. It follows that the positive effect of less idiosyncratic uncertainty on average firm size is more pronounced when γ, f_m (η) is low (high).[30]

Average Price Consider the solution for the equilibrium average price in industry k and market m:

$$p_{mk} = \frac{1}{2}\sqrt{4\gamma f_m - \sigma_k^2} + \mu_{mk} \tag{1.34}$$

Hence, the average price depends exclusively on market-specific expected profitability f_m $(+)$, the differentiation parameter γ, whereby higher γ induces firms to charge higher prices by reducing consumer responsiveness to price changes $(+)$, the average marginal cost μ_{mk} $(+)$ and the dispersion of the marginal cost σ_k^2 $(-)$. Losses from less competition due to reduced uncertainty are more sizeable in larger markets. The explanation behind this at first counterintuitive result is provided by the combination of consumer and producer first order conditions:

$$\begin{aligned} |\varepsilon_{mkn}| &= \frac{p_{mkn}}{\frac{\alpha\gamma+\eta N_{mk}\bar{p}_{mk}}{\gamma+\eta N_{mk}} - p_{mkn}} \\ p_{mkn} &= \left(\frac{|\varepsilon_{mkn}|}{|\varepsilon_{mkn}|-1}\right) c_{mkn} \end{aligned} \tag{1.35}$$

As competition decreases due to reduced uncertainty, markups go up and more than proportionally in smaller markets owing to the convexity of $\frac{\alpha\gamma+\eta N_{mk}\bar{p}_{mk}}{\gamma+\eta N_{mk}}$ in N_{mk}. As markups go up, the price feedback on the demand elasticity offers a counteracting force, which is also more pronounced in smaller markets. The net effect is more pronounced in larger cities.

[30] It is a priori unclear whether larger markets are characterized by smaller firms on average across industries. On the one hand, they feature more competition that tends to reduce firm size on average; on the other hand, firms in larger markets also face a larger customer base. Finally, markets differ also based on how many efficiency units of labor is each worker endowed with: higher efficiency tends to reduce average firm size as measured by per capita labor requirement.

Consumer Surplus Unlike Krugman (1979), the measure of firms is a concave function of market size: the benefits associated with a larger market decline in the size of the market as consumer demand decreases in aggregate consumption. Hence, the *relative* measure of firms rises more than proportionately in larger markets as uncertainty rises only if the market is sufficiently large. On the other hand, the *relative* average price always declines more than proportionately in larger markets as uncertainty rises: the more subdued rise in competition in larger markets attenuates the response of the demand elasticity to competition and the feedback of price changes on it, thus letting prices fall more than proportionately in larger markets.

By eq.1.19 consumer surplus increases through an expansion in the set of available varieties and a reduction in prices and it is therefore unambiguously higher in absolute terms in larger markets. Furthermore, it rises in response to an increase in idiosyncratic uncertainty by 1) a *direct* effect mediated by the associated increase in marginal cost dispersion, since consumers can reoptimize their consumption basket by shifting expenditure more than proportionally towards less pricey commodities, and 2) an *indirect* effect mediated by the impact of the associated decrease in average marginal cost and increase in marginal cost dispersion on the measure of varieties and prices.

The direct effect is trivially more pronounced in a large market by the presence of a larger set of varieties over which to reoptimize expenditure. The indirect effect can be decomposed into changes in the consumer surplus holding prices fixed (variety channel), and changes in the consumer surplus holding the set of varieties fixed (price channel).

Proposition 1. *The effect of an increase in idiosyncratic uncertainty on consumer surplus is more pronounced in a large market:*

$$\frac{\partial^2 CS_{mk}}{\partial Uncertainty_k \partial L_m} > 0$$

The proposition provides a direct intuition on the consequences of a decline in uncertainty: so long as consumer surplus is more negatively affected in larger markets, a decline in uncertainty would trigger spatial reallocation of workers from large to small cities. While the effect of a decline in uncertainty on consumer surplus through the extensive margin of production, N_{mk}, is not necessarily more sizeable in larger markets due to the decreasing marginal gain from variety given the indirect utility function, the effect of a decline in uncertainty on consumer surplus through the price level, p_{mk}, is always more sizeable in larger markets due to the convexity of the indirect utility function in prices. The second effect is stronger, so that the combined *negative* effect is always unambiguously more sizeable in larger markets.

1.3.6 Spatial Equilibrium

Definition 1. *Given parameters α, γ, η, f, a function $r_m = R(L_m)$ such that $R_L > 0$, a distribution of amenities $\mathbf{A}$ and efficiency units $\mathbf{H}$ for each location $m \in M$, average marginal cost $\mathbf{C}$ for each location $m \in M$ and industry $k \in K$, dispersion of marginal cost $\mathbf{S}$ for each industry $k \in K$, and total population in the economy $\overline{L}$, the competitive equilibrium of the economy is defined by a subset of available locations $\Omega \subseteq M$ such that:*

- *the number of firms, $N_{mk} \in \mathbb{R}$, choosing to operate in industry $k \in K$ and location $m \in \Omega$ is such that:*

$$E\left(\pi_{mk}\right) = f$$

- *the price, $p_{mkn} \in \mathbb{R}$, charged for each variety produced in each industry $k \in K$ and location $m \in \Omega$ such that:*

1. *consumers in each location$m \in \Omega$ maximize their utility, taking as given N_{mk} and p_{mk},*
2. *firms in each industry $k \in K$ and location $m \in \Omega$ maximize their profits, taking as given N_{mk} and p_{mk},*
3. *the market for each variety clears.*

- *the measure of consumers L_m choosing to locate in location $m \in \Omega$ and indirect utility $\overline{V}$ are such that:*

$$V_m^i = \overline{V} \quad \vee \quad \textstyle\int_\Omega L_m dm = \overline{L}$$

- *the demand for the homogenous good is positive in each location $m \in \Omega$:*

$$\omega_m + \tau > \sum_{k \in K} \int_{N_{mk}} p_{mkn} q_{mkn}^i dn + r_m$$

- *the supply for the homogenous good is positive in each location $m \in \Omega$:*

$$L_m h_m > \sum_{k \in K} \int_{N_{mk}} \left(l_{mkn} + f\right) dn$$

- *the aggregate resource constraint holds:*

$$\int_{\Omega} (\omega_m + \tau)\, L_m dm = \int_{\Omega} \left(q_0^i + \sum_{k \in K} \int_{N_{mk}} p_{mkn} q_{mkn}^i dn + r_m \right) L_m dm$$

Definition 1 establishes the existence of the equilibrium for this economy. There are several features that can characterize such equilibrium, some of them particularly amenable from the point of view of the quantitative evaluation of the model. Allen and Arkolakis (2014) emphasize three of them: regularity, uniqueness and stability. A spatial equilibrium is defined as regular if all available locations are inhabited. This requirement is appealing in quantitative models that have a pronounced geographical component, absent in this paper. We therefore choose not to enforce this property and focus on the remaining two. Point-wise stability entails that $\frac{\partial V_m^i}{L_m} < 0$ for $0 < L_m < \overline{L}$ and any A_m: in a stable spatial equilibrium it must not be possible for any arbitrarily small number of individuals to improve on their utility by moving to another location. So long as this property is satisfied, any equilibrium of this economy given the set of primitives must also be unique.

Moreover, such equilibrium must feature the following property by the implicit function theorem:

$$\frac{\partial L_m}{\partial A_m} > 0 \quad \vee \quad \forall m \in (\underline{m}, M) \tag{1.36}$$

where any location with amenities $A_m < A_{\underline{m}}$ will not be populated.

The following proposition establishes point-wise stability and uniqueness of the spatial equilibrium.

Proposition 2. *If:*

$$\frac{\partial V_m^i}{L_m} < 0$$

for $0 < L_m < \overline{L}$ *and any* A_m*, then the equilibrium given in Definition 1 is unique and stable.*

1.4 Structural Estimation

1.4.1 Parametric Assumptions

Structural estimation of the model requires to take a stand on: a) $r_m = R(L_m)$, or the housing cost function; b) the distribution of amenities $\mathbf{A}$; c) the distribution of efficiency units of labor endowment $\mathbf{H}$; d) the joint distribution of city/industry first and second moments, $\mathbf{C}$ and $\mathbf{S}$.

Assumption 1. *The housing cost function is defined by:*

$$r_m = \beta L_m^{\xi} \tag{1.37}$$

We estimate β and calibrate $\xi = .8$ using data on housing prices for municipalities (*Gemeinden*) collected from the 2000, 2005 and 2010 edition of the IVD-Wohnimmobilienpreisspiegel, which offers the longitudinally longest record on housing prices in German cities. This value is close to the average of .66 found in Saiz (2010) on U.S. Metropolitan Areas.[31]

Assumption 2. *Suppose that each location is indexed based on the rank in terms of the distribution of amenities. Overall utility of location m is:*

$$U_m^i = A_m \widetilde{U_m^i} \tag{1.38}$$

with:

$$A_m = m^{\psi} \tag{1.39}$$

and $\psi < 0$.

We estimate ψ: by $\psi < 0$ low index locations are characterized by better amenities. The parametrization for amenities has two alternative interpretations: $|\psi|$ can either represent the value of the exponent characterizing the amenity distribution or the elasticity of utility to amenities.
The functional form in eq.1.39 generates a power law distribution for amenities if $\psi = -1$. The assumption of a multiplicative rather than additive interaction between the exogenous (A_m) and the endogenous ($\widetilde{U^i}_m$) agglomeration component is motivated by the goal of estimating a city size distribution that follows as closely as possible the Zipf's Law for German cities.

[31] These are not microdata and we select two series corresponding to the rent per square meter of a 70 m^2 - on average - flat, newly built in either good or medium conditions, being the rent of new housing stock the most responsive to changes in housing demand. We multiply the quote per 70 (square meter) times 12 (months). We aggregate the data up to the Metropolitan Area-level using weights based on data on *Gemeinde* population for 2000, 2005 and 2010 to obtain $P_{AMR,t}$ and $L_{AMR,t}$. Next, we run the regression:

$$\Delta P_{AMR,t} = c_t + \xi \Delta L_{AMR,t} + e_{AMR,t}$$

on 2010-2005, 2005-2000 first differences and obtain $\hat{\xi} = .8$.

Allowing amenities to be *potentially* power law distributed along with the assumption of multiplicative utility aggregator enables the model solution to flexibly inherit the distributional properties of the amenity vector according to the relative strength of agglomeration/congestion parameters featured in $\widetilde{U_m^i}$.[32]

In the model wage per efficiency unit is fixed across locations, and we account for differences in total labor income in the following way. Large cities tend to specialize in the production of the non-traded good. The efficiency units labor requirement per firm tends to increase in firm-specific productivity and decrease in the extent of local competition in the same production of the same good. At the same time, total demand of efficiency units of labor tends to increase in city size by the extensive margin. The effect of competition through the extensive margin dominates over the effect of competition through the intensive margin: aggregation across firms of per capita demand of efficiency units yields that, assuming efficiency units were constant across cities, large cities would specialize in the non-tradable - differentiated - sector.

This model prediction would be counterfactual as non-tradable employment as a fraction of local employment is constant. We use this piece of evidence to motivate our calibration of efficiency units of labor. If workers receive a large endowment of efficiency units upon moving to a given location and firms labor demand is in terms of *efficiency units* rather than *per capita units* of labor, then, as long as, the endowment is sufficiently increasing in city size, total employment in the non-tradable sector can be a constant fraction of total employment in a given city. Specifically, we assume:

$$\text{share}^{Diff.} = \frac{l_{m,Demand}^{Diff.}}{h_m} \times \frac{1}{L_m} \tag{1.40}$$

[32]Rewrite $\frac{\partial L_m}{\partial A_m} = -\frac{\widetilde{V_m^i}}{A_m \frac{\partial \widetilde{V_m^i}}{\partial L_m}}$ and integrate both sides:

$$\frac{1}{\widetilde{V_m^i}} \frac{\partial \widetilde{V_m^i}}{\partial L_m} dL_m = -\frac{1}{A_m} dA_m \Leftrightarrow \ln \widetilde{V_m^i}(L_m) = -\ln A_m + C$$

For a given constant C^* pinned down by the population constraint:

$$\widetilde{V_m^i}(L_m^*) = -A_m \exp(C^*) \Leftrightarrow L_m^* = \widetilde{V_m^i}^{-1}(-A_m \exp(C^*))$$

Given $\frac{\widetilde{V_m^i}}{\partial L_m} < 0$, if $\frac{\partial^2 \widetilde{V_m^i}}{\partial L_m^2} > 0$, then L_m^* must be an increasing and convex function of amenities $L_m^* \propto A_m^\rho$ ($\rho > 1$), if $\frac{\partial^2 \widetilde{V_m^i}}{\partial L_m^2} < 0$, then L_m^* must be an increasing and concave function of amenities ($\rho < 1$). By log-linearization of the expression for amenities given in Assumption 2, $\ln(m) \propto c_0 + \frac{1}{\rho\psi} L_m$. The Zipf's Law for the city size distribution follows if $\zeta = \frac{1}{\psi\rho} = -1$ and it is thus determined by the interaction of exogenous agglomeration (ψ) and endogenous agglomeration (ρ, which in spite of not admitting closed-form solution depends on the relative strength of agglomeration/congestion parameters featured in $\widetilde{U_m^i}$).

Dividing total labor demand in the differentiated sector by individual endowment of efficiency units returns total labor demand in per capita terms, or city employment in the non-tradable sector. While evaluating the model at each parameter combination, we set h_m such that eq.1.40 always holds: therefore, a higher gradient $\frac{\partial h_m}{\partial L_m}$ corresponds to parameter combinations that imply disproportionately higher total labor demand, $l^{Diff.}_{m,Demand}$, in large cities vs. small cities. For instance, higher values for α, the parameter governing substitution towards the differentiated sector would trivially need to be compensated by higher on average h_m. If the gradient of firm specific productivity in city size is too steep, the gradient $\frac{\partial h_m}{\partial L_m}$ will reflect this, and also be steep.[33]

Assumption 3. *The distribution of the firm/industry-specific component of marginal cost is inverse Pareto with average and dispersion given by:*

$$\mu_k = (1/b_k)\left(\frac{\theta_k}{\theta_k+1}\right) \quad \vee \quad \sigma_k^2 = (1/b_k)^2\left(\frac{\theta_k}{(\theta_k+1)^2(\theta_k+2)}\right)$$

The location-specific component of marginal cost is given by:

$$\begin{gathered}\widetilde{\mu_m} = \phi\left(1 - 2\times\left(\frac{\kappa l_m}{1+\kappa l_m}\right)\right)\\ 0 \leq \kappa \leq 1 \quad \vee \quad l_m = \min\left(\frac{L_m}{L_{max}}, 1\right) \quad \vee \quad \phi > 0\end{gathered} \tag{1.41}$$

The city-specific component of average marginal cost, $\widetilde{\mu_m} \in (0, \phi)$, where $\widetilde{\mu_m} = \phi$ for all cities if $\kappa = 0$ and $\widetilde{\mu_m} = \phi$ for a city of infinitesimal size and $\widetilde{\mu_m} = 0$ for a city of maximal size if $\kappa = 1$. Intuitively, ϕ identifies the relative magnitude of the location-specific vs. industry-specific component of the average marginal cost. As κ increases the average marginal cost decreases for all producers, and more so for those located in larger markets: therefore, κ identifies the strength of the comparative advantage of large cities in firm productivity.

The minimum and shape parameter of the industry-specific distribution of idiosyncratic cost is calibrated following Luttmer (2007). The calibration procedure is described in Appendix 1.5.

The task of finding a solution within a parameter region where endogenous selection does not arise demands a restriction on the set of values that the sunk cost can take. Additionally, we provide flexibility to the model by allowing the sunk cost to increase with market size, due to, for instance, advertising costs as in Arkolakis (2010).

[33] See Appendix 1.5 for details on the calibration of the efficiency units endowment.

Assumption 4.

$$\begin{aligned} f_m = \tfrac{L_m}{4\gamma}\left\{f_L + (f_H - f_L)\times\left(1 - 2\times\left(\tfrac{\chi l_m}{1+\chi l_m}\right)\right)\right\} \\ 0 \le \chi \le 1 \quad \vee \quad l_m = \min\left(\tfrac{L_m}{L_{max}}, 1\right) \end{aligned} \tag{1.42}$$

Per capita sunk cost is bounded between f_L and f_H, which are themselves function of α and ϕ.[34]

As for firm productivity, the parametric choice described in Assumption 4 features supermodularity between the city-specific component, l_m, and the parameter governing savings in (per capita) sunk cost, χ. Per capita sunk cost is, thus, lower in large markets, and more so the higher is χ.

We estimate $\delta = \eta/2\gamma$, rather than γ and η separately, due to lack of identification of γ (see also Ottaviano et al. (2017)). This parameter captures the negative impact of competition on firm profits: it depends positively on η, which captures the extent to which individual firm demand declines as aggregate production rises. The negative impact of an increase in η is magnified for low values of γ, since expected profits are higher when love for variety is high (low γ).

Due to possible lack of separate identification for ϕ and κ, we structurally estimate the parameters separately for each $\phi = \{.01, .025, .05, .1\}$, and select the model solution, and therefore value of ϕ, such that the elasticity of the city-specific component of firm productivity to city size is the closest to the average available empirical estimate of 5% (Rosenthal and Strange (2004)).

The final set of parameters is $\Theta = \{\kappa, \alpha, \delta, \chi, \beta, \psi\}$.

1.4.2 Choice of Moments and Identification

The parameters of the model are chosen so as to minimize the weighted sum of squared deviations between the moments observed the data and the ones predicted by the model:

$$\min_{\Theta} = (m - \hat{m}(\Theta))' W (m - \hat{m}(\Theta)) \tag{1.43}$$

where m is the set of moments observed in the data, $\hat{m}(\Theta)$ is the set of predicted moments for given Θ.

We have six parameters to estimate, four demand related parameters, $\{\kappa, \alpha, \delta, \chi\}$, and two supply related parameters $\{\beta, \psi\}$. We target the following two moments in order to identify the supply related parameters:

1. $p(75)/p(25)$ of the city size distribution.

[34] For more details on the implementation of the parameter restrictions and the selection of ϕ, please refer to Appendix 1.5.

2. $p(mean)/p(50)$ of the city size distribution.

In order to identify the demand related parameters, we target the following four moments:

1. median number of establishments in top-risky industries ($\widehat{\sigma^2}_k \geq \left\{\widehat{\sigma^2}_k\right\}_{median}$) rescaled by $\overline{N} = 250$ on average in the largest 36 cities, where largest cities are those located above the median of the city size distribution.

2. median number of establishments in bottom-risky industries ($\widehat{\sigma^2}_k < \left\{\widehat{\sigma^2}_k\right\}_{median}$) rescaled by $\overline{N} = 250$ on average in the largest 36 cities, where largest cities are those located above the median of the city size distribution.

3. median number of establishments in top-risky industries ($\widehat{\sigma^2}_k \geq \left\{\widehat{\sigma^2}_k\right\}_{median}$) rescaled by $\overline{N} = 250$ on average in the smallest 36 cities, where largest cities are those located above the median of the city size distribution.

4. average of log establishment size in the city/industry distribution.

The first demand related moment identifies most closely δ, which rescales up or down the entire firm count distribution. The first and second demand related moments jointly provide information on the within-city variation across industries. This moment is shown in the upper left plot of Fig.1.5 to be the most sensitive to changes in κ: intuitively, the larger is κ, the lower is average marginal cost in the largest cities for all industries, thus reducing the responsiveness of entry patterns to variation in industry-specific attributes.

The first and third demand related moments jointly provide information on within-industry variation across cities. This moment is shown in the upper right plot of Fig.1.5 to be especially sensitive to changes in the sunk cost: when the sunk cost decreases, it decreases more than proportionally for large cities, thus amplifying the responsiveness of entry patterns to variation in city-specific attributes.

Finally, the last moment, the average of log of establishment size responds primarily to changes in α, as shown in the bottom plot of Fig.1.5. When α increases, entry picks up in all industries and cities: this causes total employment in efficiency unit to increase and more so in larger markets, above the predicted 28% fraction of total employment. It must therefore be that worker productivities, or the per capita endowment of efficiency units, is sufficiently high, which implies a reduction in firm size expressed in per capita units.

We estimate the model by the method of simulated moments (McFadden (1989)).[35]

[35]We first solve the model on a sparse grid on the parameter space. We then select the 5 best

1.4.3 Model Fit

The model solution in terms of city size distribution consists of a regular geography, in spite of no reference to this property of the spatial equilibrium among the targeted moments. The model replicates 70% of the geographical concentration observed in the data as measured by the Herfindahl Index. The actual city size distribution is compared to the model solution in Fig.1.6. The model underpredicts actual size only in the top-2 most populated Metropolitan Areas in Germany.

The explanatory power with respect to the distribution of rescaled establishments counts across cities and industries is summarized by the R-squared of a OLS regression:

$$\text{relative establishment count}_{mk}^{data} = \alpha + \beta \text{relative establishment count}_{mk}^{model} + \epsilon_{mk} \tag{1.44}$$

The R-squared stands at $.44$: the unexplained variation in the data is mostly driven by city/industries featuring high establishment count. The degree of skewness observed in the data could be matched by means of introducing a stronger complementarity between city size and industry profitability (e.g., variation in sunk costs across industries).

To the best of our knowledge, there exists no prior paper whose focus was to provide a quantitative exercise aiming at describing variation in variety across goods and cities. Pre-existing quantitative papers in this field have devoted their attention to analyzing variation in the intensive margin, i.e., average firm size Gaubert (2018). Since the primary objective of this paper is to investigate how idiosyncratic uncertainty affects variety and competition (and indirectly firm size), we have chosen to target establishment counts. Simultaneous description of variation in establishment counts and establishment size across cities and industries is beyond the scope of this paper. However, Fernandes et al. (2017) recent analysis in the context of international trade highlights the importance of joint consideration of both intensive and extensive margin when assessing the properties of different distributional assumptions. Adapting this approach to the economic geography literature looks promising in terms of further refining the shape of agglomeration economies along the firm productivity distribution.

parameter combinations and run the bounded simulated annealing algorithm (Goffe, Ferrier and Rogers (1994)) starting from each initial point. A known shortcoming of simulated annealing is that it sacrifices precision to accuracy. We therefore further implement a Nelder-Mead simplex search method setting as initial value the solution returned by each run of simulated annealing.

1.4.4 Interpretation of Parameter Estimates

The estimated parameters are reported in the upper panel of Table 1.18. The estimate for the parameter capturing the strength of comparative advantage of cities, κ, implies agglomeration externalities in the order of magnitude of 7%. The elasticity of productivity to city size implied by the model corresponds to the slope coefficient of the OLS regression:

$$\ln(1/\mu_{mk}) = \gamma_k + \beta \ln L_m + \epsilon_{mk} \tag{1.45}$$

The value $\hat{\beta} = .07$ lies at the lower boundary of the range of existing empirical estimates (between 3% and 8%) provided in Rosenthal and Strange (2004).

To assess the plausibility of point estimates for the preference parameters, α and δ, we calculate the elasticity of the price index across cities and industries to city size. We obtain an elasticity of 2%, slightly higher than 1.1% in Handbury and Weinstein (2015): the mild overshoot of the model relative to existing evidence is most likely due to the absence of real income effects on firm markups, which would attenuate the pro-competitive role of market size as explained in Appendix 1.5.

1.4.5 Counterfactual Analysis and Policy

We undertake the following counterfactual. We solve the model at the estimated parameters in an environment characterized by reduced uncertainty, and therefore a more compressed productivity distribution in all industries. The impact on both consumer surplus and spatial equilibrium is summarized by Table 1.19. Following a decline in uncertainty by 2.75%, consumer surplus declines by 13% in the smallest city and by 5% in the largest city in the sample. These numbers are net of the spatial equilibrium adjustment: since the size of the smallest city increases by .56%, while the one for the largest city shrinks by .22%, the partial equilibrium response of consumer surplus would have been more (less) negative for the smallest (largest) city.

The losses to consumers accrue in the form of less varieties and higher prices. About 75% of consumer surplus decline in the smallest city is explained by less variety, while a reduction in the set of varieties accounts for 50% of consumer surplus decline in the largest city. The partition of consumer surplus losses into those due to less variety vs. those due to less competition is, thus, about 1/2-1/2 (3/4-1/4) for the largest (smallest) city in the sample. These two boundary estimates contain the 2/3-1/3 partition estimated in Feenstra and Weinstein (2017) in the context of gains following a reduction in trade barriers: however, the distance between them underscores the importance of accounting for asymmetries that hinge on product market size.

The small estimated macroeconomic impact of a drop in uncertainty is not surprising since it hinges on the response of a sector that accounts for a mere .28 share of aggregate employment. Growing evidence hints at the possibility that the advantages granted by urban density in consumption (i.e., lower prices and more variety in the non-tradable sector in large cities) may affect worker location decisions just as much as traditional advantages in production do (i.e., higher wages in large cities). However, there existed to date no test of the quantitative performance of agglomeration economies in consumption standing alone. Since the structural model estimated in this paper incorporates both kinds of advantages, the counterfactual exercise just described provides precisely this test.

The experiment conducted consists of lowering the degree of uncertainty and letting consumers relocate based on the relatively less advantageous terms for consumption in large cities, while keeping the spatial distribution of income constant: the limited impact on the city size distribution reveals that the advantages granted by urban density in consumption are second order at explaining individual location choices. More specifically, the extent of spatial reallocation triggered by the change in uncertainty is small given that estimated worker productivities - which are kept fixed during the counterfactual - represent the bulk of individual utility.[36]

We consider the problem of central government who observes the reallocation of economic activity caused by the exogenous decline in idiosyncratic uncertainty, and who is interested in minimizing aggregate losses in terms of consumer surplus. Since consumer surplus is not observed, we assume the government attempts to revert to the status quo by implementing a proportional subsidy on the sunk cost optimally chosen so as to minimize changes in the city size distribution:

$$\min_{\text{subsidy}} Loss = \sqrt{\sum_m \frac{L_m}{\overline{L}} \left[L_m \left(\sigma_k^{baseline}, f_m \right) - L_m \left(\sigma_k^{reduced}, (1 - \text{subsidy}) \times f_m \right) \right]^2}$$

subject to the competitive equilibrium.

[36]The solidity of this finding hinges on the assumptions of perfect elasticity of labor supply and absence of a direct response to declining uncertainty in the organization of production in the tradable sector. Both assumptions have been made with the intention of characterizing the responsiveness of the spatial equilibrium to one of the factors magnifying the advantages granted by urban density in consumption, i.e., idiosyncratic uncertainty. Removing the assumption of perfect competition in the tradable sector would magnify the negative impact of a reduction in uncertainty on spatial concentration for high enough trade barriers across cities. In this case, the disproportional drop in competition experienced by large cities would however be mitigated by a downward adjustment in the cost of labor. In the absence of better evidence on trade patterns across cities based on industry idiosyncratic uncertainty, a conclusion on the effect of a drop in idiosyncratic uncertainty on the city size distribution in the presence of these additional channels would be unwarranted.

We calculate this subsidy to be 6%. The percent change in size and consumer surplus for both the largest and smallest cities in the sample under the subsidy and relative to the baseline economic environment of uncertainty (equal to the level in 2000) is reported in Table 1.20.

It is not surprising that size is barely changed relative to size under baseline idiosyncratic uncertainty, given that this was the target of the central government. It is, however, interesting to notice that consumer surplus improves relative to the baseline scenario by 26% and 7% in the smallest and largest city, respectively. The explanation for this striking result hinges on the more geographically insulated impact of pro-competitive policies in the estimated model relative to changes in the productivity distribution.

1.5 Conclusion

This paper analyzes both empirically and theoretically the implications of declining business dynamism on consumer surplus and the spatial distribution of economic activity. First, it highlights how the decline in business dynamism came along with a substantial drop in idiosyncratic uncertainty, defined as the dispersion of fluctuations in establishment size that are not explained by observable establishment characteristics and not driven by aggregate risk factors. Secondly, it provides evidence consistent with the view that time-invariant idiosyncratic uncertainty is an asset from the viewpoint of potential entrepreneurs, as it unlocks growth options that raise the value of starting a firm: since the individual benefit outweighs the cost in terms of increased likelihood of facing productive competitors, competition must decline as idiosyncratic uncertainty falls, resulting in a consumer surplus loss. Additionally, it shows that the loss in terms of competitiveness is more sizeable in large markets, thus unveiling a further negative implication of declining uncertainty, i.e., a reduction in spatial concentration. Finally, it proposes a structural model to convey analytical intuition on the interpretation given to the reduced form empirical analysis. Counterfactual analysis run on the estimated model indicates that a reduction in microeconomic uncertainty of the magnitude seen in the data led to a 10% on average consumer surplus loss and a modest compression in the city size distribution.

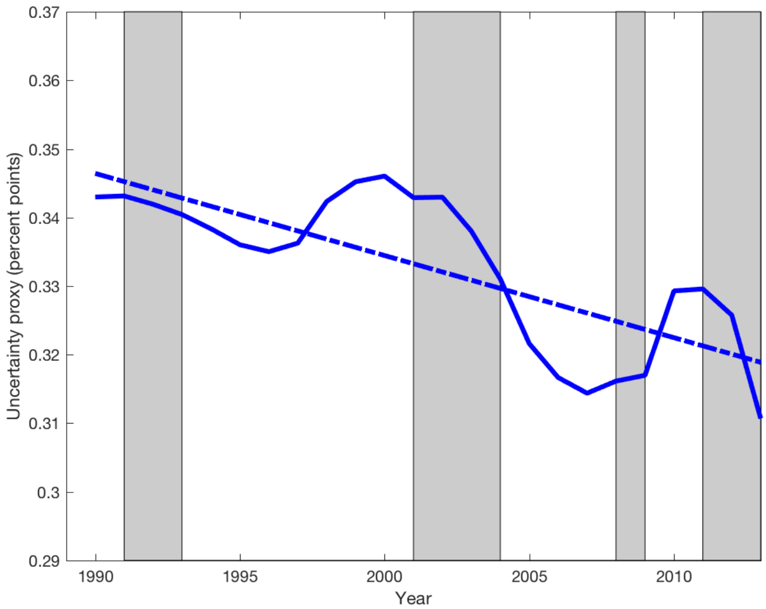
What could be behind the decline in idiosyncratic uncertainty? One possibility is that the type of technology employed by firms nowadays while being more productive on average, it is also less risky. We can think of the technology employed by a given firm as dependent on the characteristics of its labor force, and of an innovation to such technology as shaped by the luck with which any of the workers succeeds at improving the output of the tasks she is assigned with. Exogenous shifts in the composition of the labor force towards a worker type whose productivity is more predictable, i.e., high skilled workers, might have been responsible for a decline in microeconomic uncertainty faced by firms over their life-cycle. Alternatively, and in opposition with what assumed in this paper, the decline in microeconomic uncertainty could be mainly demand-driven. For example, customers are nowadays more loyal to the brands they consume and less prone to dramatic changes in preferences, which could be seen as the outcome on an increase in the effort put by companies into marketing and brand management strategies.

The findings presented in this paper highlight the positive association between entrepreneurial risk-taking and competition. The loss in competitiveness characterizing the decline in business dynamism can certainly be explained both in terms of a decline in actual uncertainty, as suggested throughout this analysis, as well as in terms of a decline in perceived uncertainty, if barriers preventing entrepreneurs from taking on risk and starting a business have increased over the past decades. A case in point is given by a tightening in the conditions on profitability to gain access to funding of new projects. A more detailed analysis distinguishing actual from perceived risk factors and their impact on competition is left for future research.

Appendices

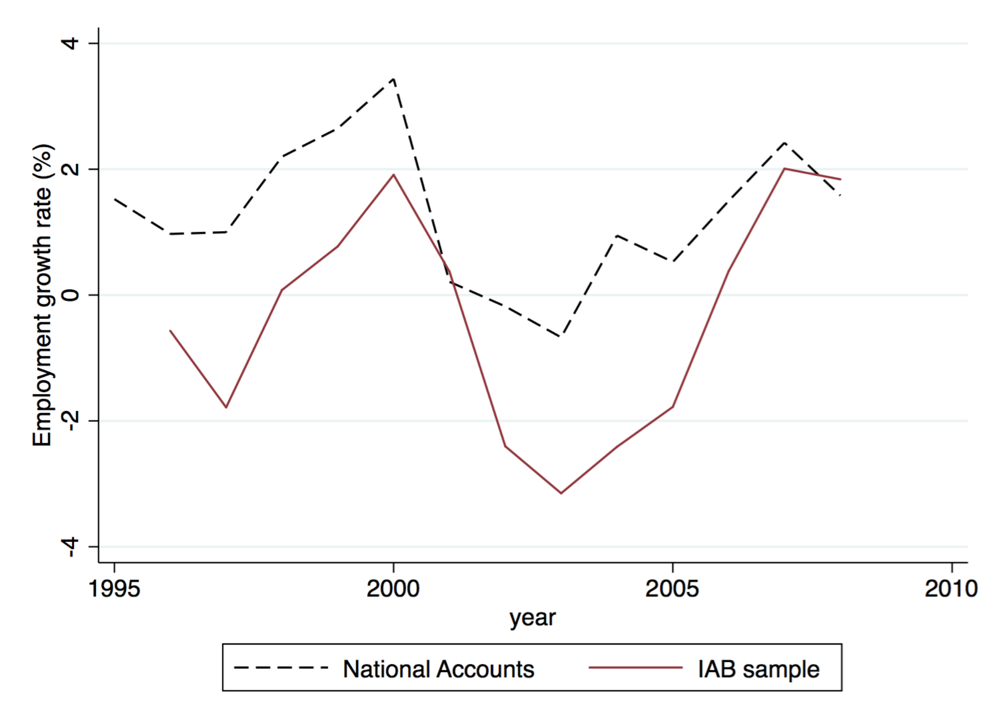
Figures

Figure 1.1: Declining idiosyncratic uncertainty over 1990-2014 in former West Germany



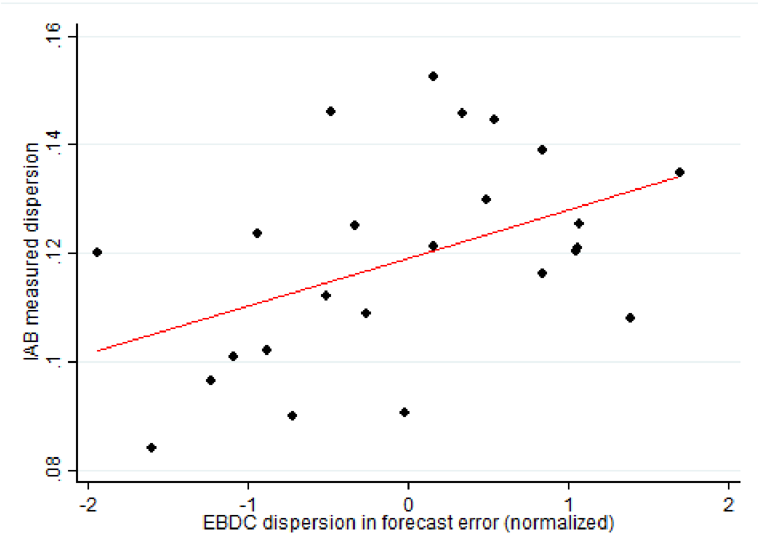
3-year moving average of idiosyncratic uncertainty as proxied by residual dispersion of establishment employment growth rates, linear trend coefficient .0011***. Population of establishments located in former West Germany. Shaded areas denote national recessions (OECD). Estimation of uncertainty proxy described in detail in Section 1.2. Source: AWFP.

Figure 1.2: Employment growth rate in AWFP and National Accounts



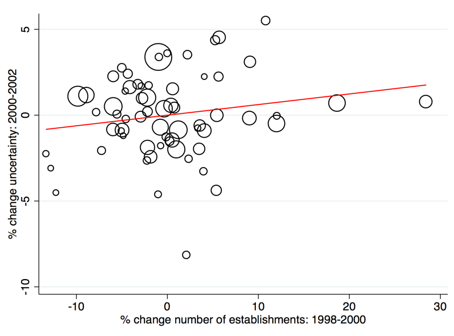
Yearly private sector full-time employment growth. Source: AWFP and DESTATIS.

Figure 1.3: Comparison between measures of idiosyncratic uncertainty obtained with AWFP vs. EBDC data on firms' expectations

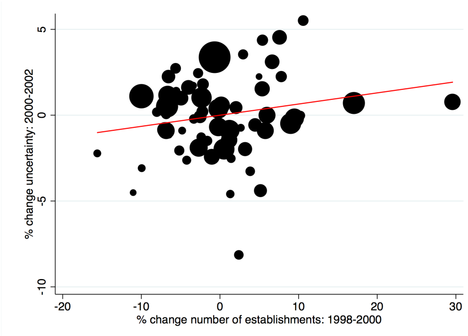


Expectations-based proxy for idiosyncratic uncertainty consists of the business-specific 3-month forecast error of business volume. The approach follows the methodology described in Bachmann et al. (2013). Source: AWFP and Economics and Business Data Centre (CESifo) Expectations Business Panel (2016).

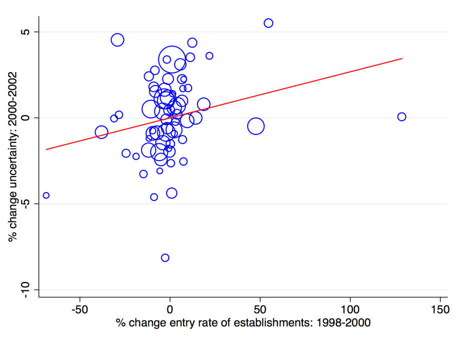
Figure 1.4: Change in uncertainty vs. change in the number of establishments



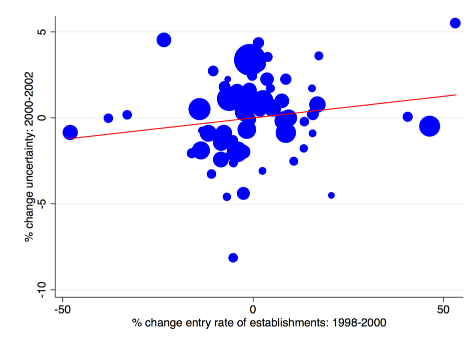
(a) Δ: all cities



(b) Δ: large cities



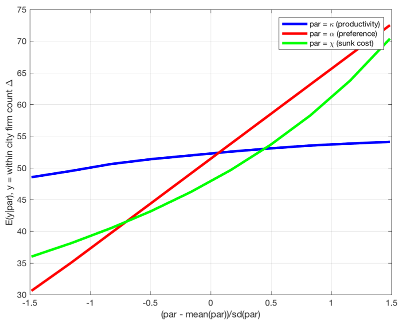
(c) Δ entrants: all cities



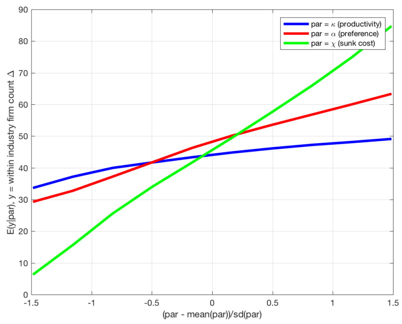
(d) Δ entrants: large cities

Growth rate in total establishment counts (entrant counts) by 5-digit non-tradable industry (y-axis) against lagged growth rate in uncertainty. In the bottom two panels only establishments located in large Metropolitan Areas (cities that cumulatively account for 50% of total employment after having ranked cities by size). Growth rates both in establishment counts and uncertainty are are demeaned by the sectoral average. Source: AWFP.

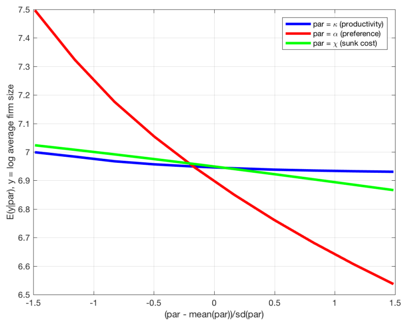
Figure 1.5: Simulated ensitivity of moments to model parameters



(a) Moment #2



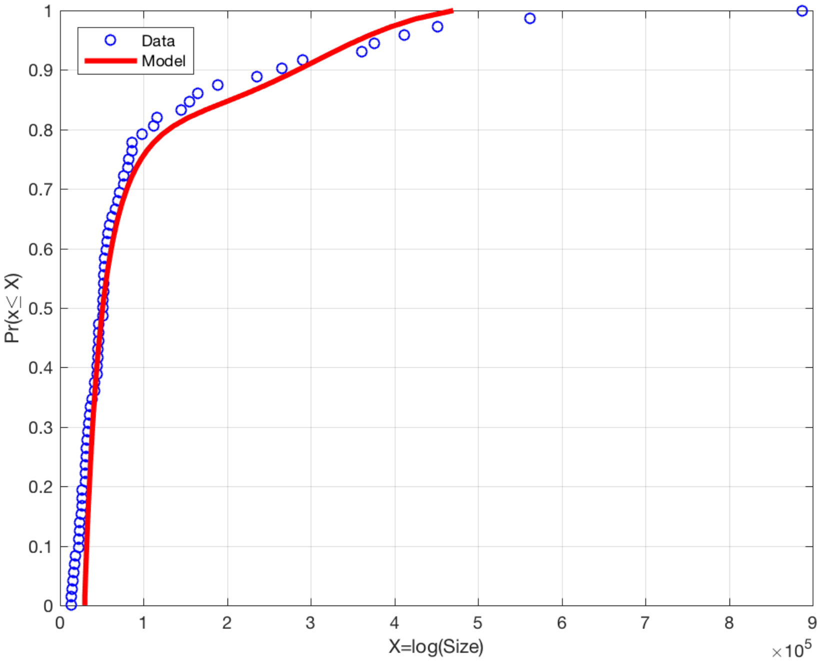
(b) Moment #3



(c) Moment #4

Moment # 2: within-city difference in establishment count (large city); Moment # 3: between-city difference in establishment count (high uncertainty industry); Moment # 4: average log establishment size. Dependence on κ (strength of agglomeration externalities in productivity), α (preference for differentiated good), χ (sunk cost per capita). $E(y|par)$ corresponds to median $y|par$ across $T = 1000$ combinations randomly drawn for excluded 3 parameters. The support for each parameter is standardized for comparability purposes.

Figure 1.6: Cumulative city size distribution in the model vs. data



Geographical concentration (Herfindahl index) model 70% of data.

Tables

Table 1.1: Mean establishment growth rate by age/size category

Age categories					
Group	$\widehat{\beta}_{max}$	$\widehat{\beta}_{min}$	$\widehat{\beta}_{mean}$	Sector: $\widehat{\beta} = \widehat{\beta}_{max}$	Sector: $\widehat{\beta} = \widehat{\beta}_{min}$
age: 1-4	.124	.022	.059	BS	H
age: 5-9	.041	-.009	.010	BS	C
age: 10-15	.020	-.024	-.001	BS	C
age: 16+	.010	-.039	-.014	RE	C
Size categories					
Group	$\widehat{\beta}_{max}$	$\widehat{\beta}_{min}$	$\widehat{\beta}_{mean}$	Sector: $\widehat{\beta} = \widehat{\beta}_{max}$	Sector: $\widehat{\beta} = \widehat{\beta}_{min}$
Size: 1-7	.0009	-.0360	-.0136	C	BS
Size: 8-25	.0043	-.0163	-.0055	C	RE
Size: 25-99	.0085	-.0063	.0001	M	O
Size: 100+	.0148	-.0148	-.0003	F	C

$\widehat{\beta}_s$ and $\widehat{\beta}_a$ in eq.1.7. Sector labels: Manufacturing ("M"), Construction ("C"), Retail and Wholesale ("RW"), Hospitality ("H"), Finance ("F"), Real Estate ("RE"), Business Services ("BS"), Other Services ("O"). Source: AWFP.

Table 1.2: R^2 of eq.1.7 estimated at the sectoral level

Sector	**Measure#1**	**Measure#2**	**Measure#3**	**N**
Manufacturing	1.05%	2.70%	0.82%	265,758
Construction	1.18%	2.00%	0.94%	401,283
Trade	0.63%	2.50%	0.60%	647,622
Hospitality	0.27%	1.64%	0.14%	233,077
Finance	1.02%	2.45%	0.74%	61,305
Real Estate	0.38%	1.71%	0.38%	108,383
Business Services	1.68%	3.60%	1.33%	548,819
Other Services	0.44%	1.71%	0.24%	282,835

Eq.1.7 is estimated at the sectoral level without city/5-digit industry/year dummies (uncertainty Measure#1), without age/size dummies (uncertainty Measure #2); with age/size dummies after having demeaned both dependent and independent variables by the city/5-digit industry/year mean (uncertainty Measure#3). The sample for each sector excludes low-density ($N < 30$) city/5-digit industry/year groups. Source: AWFP.

Table 1.3: Industry descriptive statistics: balance-sheet information

Statistics	W	M	**Tangible** I	**Intangible** I
		Other industries		
Mean	.242	.442	.210	.015
Std.Dev.	.192	.132	.476	.060
N	65			
Y share	.77			
E share	.65			
		Selected industries		
Mean	.282	.381	.132	.010
Std.Dev.	.156	.129	.266	.041
N	46			
Y share	.23			
E share	.35			

Descriptive statistics at the firm-level by industry (4-digit) for year = 2013: wage bill (W) divided by firm turnover (column 1), material cost (M) divided by firm turnover (column 2), ratio of firm total investment in fixed tangible assets (Tangible I) divided by turnover (column 3), ratio of firm total investment in fixed intangible assets (Intangible I) divided by turnover (column 4). Average of the industry-specific mean and standard deviation among firms (only industries with more than 50 firms are reported) within each selected group of industries: industries excluded fro the analysis (upper panel), industries included in the analysis (lower panel). Aggregate output and employment share for both industries subsamples is reported. Source: AMADEUS.

Table 1.4: Descriptive statistics of estimated industry-specific idiosyncratic uncertainty by establishment subgroups

Statistic	**All**	**Young/Small**	**Young/Large**	**Old/Small**	**Old/Large**
Mean	.33	.38	.34	.30	.22
Min	.18	.17	.24	.19	.17
Median	.32	.38	.33	.32	.22
Max	.43	.51	.48	.41	.31
Std.Dev.	.04	.05	.05	.03	.03

An establishment is classified as small if it has less than 5 employees and young if it is below 10 years of age. Industry-specific idiosyncratic uncertainty is measured by the sample standard deviation of estimated residuals in eq.1.7. Source: AWFP.

Table 1.5: Testing city size specificity of uncertainty measure

	Idiosyncratic uncertainty
Small/Middle	0.004
	(1.44)
Middle/Large	0.014^{***}
	(5.75)
Large	0.002
	(0.65)
Constant	0.321^{***}
	(181.71)
Observations	280

t statistics in parentheses. Standard errors are clustered at the AMR-level.
* $p < .1$, ** $p < .05$, *** $p < .01$

Cities are grouped in 4 city size classes as follows: a) they are ranked by total establishment count, b) they are classified as “Small”, “Small-Middle”, “Middle-Large”, “Large” if they correspond respectively to less than 25%, between 25% and 50%, between 50% and 75%, more than 75% of cumulative firm count share at the aggregate level. Next idiosyncratic uncertainty is calculated for each 5-digit industry and group of cities. The regression includes 5-digit industry fixed effects and the reference category is “Small” cities. Source: AWFP.

Table 1.6: Number of establishments and establishment size by industry uncertainty and market size

	All non-tradable industries	Retail and restaurants
	Rescaled establishment count	
Uncertainty$_k$	.002***	.022***
	(.001)	(.001)
Uncertainty$_k$ $\times$ Large$_m$	.012**	.017***
	(.001)	(.004)
	Log average establishment size	
Uncertainty$_k$	-.016***	-.048***
	(.003)	(.007)
Uncertainty$_k$ $\times$ Large$_m$	.012	-.016
	(.009)	(.020)
Observations	5040	1728

std.errors clustered at the AMR-level in parentheses.

* $p < .1$, ** $p < .05$, *** $p < .01$

Rescaled establishment count corresponds to the ratio between city/industry establishment count and the normalization factor ($N = 250 = 97.5^{th}$ percentile of the aggregate establishment count distribution). Industry-specific establishment-level uncertainty is measured by the sample standard deviation ($\times 100$) of estimated residuals according to eq.1.7 after exclusion of low-density cells from sectoral regressions. Retail and restaurants correspond to 2-digit group “52” and 3-digit groups “553” and “554”. Source: AWFP.

Table 1.7: Number of establishments by group-specific Industry uncertainty and market size

	Y/S	Y/L	O/S	O/L
	Rescaled establishment count			
Uncertainty$_k$	0.00248**	-0.00354***	-0.00146	0.00713*
	(2.28)	(-3.17)	(-0.74)	(1.94)
Uncertainty$_k$ $\times$ Large$_m$	0.0240	0.0299	0.00661	0.142**
	(1.19)	(1.35)	(0.18)	(2.17)
Observations	5040	5040	5040	5040

t statistics in parentheses. Standard errors are clustered at the AMR-level.

* $p < .1$, ** $p < .05$, *** $p < .01$

Rescaled establishment count corresponds to the ratio between city/industry establishment count and the normalization factor ($N = 250 = 97.5^{th}$ percentile of the aggregate establishment count distribution). Industry/group-specific Uncertainty is measured by sample standard deviation ($\times 100$) of estimated residuals after elimination of low-density cells from sectoral regressions. An establishment is considered young (Y) if it is below 10 years of age and old (O) otherwise, small (S) if it has less than 5 employees and large (L) otherwise. Source: AWFP.

Table 1.8: Number of establishments by idiosyncratic uncertainty and market size: shift/share analysis

All non-tradable industries			
	$\hat{\sigma}_k$	$\hat{\sigma}_k\|_{FixedUncertainty}$	$\hat{\sigma}_k\|_{FixedShare}$
	Rescaled establishment count		
Uncertainty$_k$	0.0021***	0.0077***	0.0008
	(3.87)	(6.78)	(1.51)
Uncertainty$_k$ $\times$ Large$_m$	0.012***	0.0205***	0.0089**
	(3.46)	(2.77)	(2.21)
Observations	5040	5040	5040
Retail and restaurants			
	$\hat{\sigma}_k$	$\hat{\sigma}_k\|_{FixedUncertainty}$	$\hat{\sigma}_k\|_{FixedShare}$
	Rescaled establishment count		
Uncertainty$_k$	0.0223***	0.0297***	0.0231***
	(9.08)	(8.03)	(7.55)
Uncertainty$_k$ $\times$ Large$_m$	0.0175*	0.0440***	0.0097*
	(1.95)	(2.77)	(1.79)
Observations	1728	1728	1728

t statistics in parentheses. Standard errors are clustered at the AMR-level.

* $p < .1$, ** $p < .05$, *** $p < .01$

Rescaled establishment count corresponds to the ratio between city/industry establishment count and the normalization factor ($N = 250 = 97.5^{th}$ percentile of the aggregate establishment count distribution). Industry-specific establishment-level uncertainty is measured by: baseline sample standard deviation (column 1), sample standard deviation constructed by means of weighting constant within-group uncertainty by industry/group-specific establishment count shares (column 2); sample standard deviation constructed by means of weighting industry-specific within-group uncertainty by constant group-specific establishment count shares (column 3). Industry/group-specific uncertainty is measured by sample standard deviation ($\times 100$) of estimated residuals after elimination of low-density cells from sectoral regressions. An establishment is considered young if it is below 10 years of age, small if it has less than 5 employees. Source: AWFP.

Table 1.9: Number of establishments by idiosyncratic uncertainty and market size: different subsamples

	$\hat{\sigma}_k$		N_{mk}	
	1995-2000	2000-2005	1998	2002
	Rescaled establishment count			
Uncertainty$_k$	0.00165***	0.00182***	0.00113**	0.00260***
	(3.24)	(3.48)	(2.16)	(4.72)
Uncertainty$_k$ $\times$ Large$_m$	0.0108***	0.0111***	0.00978***	0.0136***
	(3.00)	(3.37)	(2.78)	(3.80)
Observations	5040	5040	5032	5031

t statistics in parentheses. Standard errors are clustered at the AMR-level.

* $p < .1$, ** $p < .05$, *** $p < .01$

Rescaled establishment count corresponds to the ratio between city/industry establishment count and the normalization factor ($N = 250 = 97.5^{th}$ percentile of the aggregate establishment count distribution). Industry-specific establishment-level uncertainty is measured by sample standard deviation ($\times 100$) of estimated residuals after exclusion of low-density cells from sectoral regressions based on observations for years 1995-2000 (column 1), 2000-2005 (column 2), 1995-2005 (column 3-4). Establishment counts are for year 2000 (column 1-2), 1998 (column 3), 2002 (column 4). Source: AWFP.

Table 1.10: Number of establishments by idiosyncratic uncertainty and market size: volatility-based measure

	All	Surviving (1995-2005)
	Rescaled establishment count	
Uncertainty_k	0.00350^{***}	0.00904^{***}
	(4.62)	(7.00)
$\text{Uncertainty}_k \times \text{Large}_m$	0.0162^{***}	0.00952
	(3.19)	(0.66)
Observations	5040	5040

t statistics in parentheses. Standard errors are clustered at the AMR-level.

* $p < .1$, ** $p < .05$, *** $p < .01$

Rescaled establishment count corresponds to the ratio between city/industry establishment count and the normalization factor ($N = 250 = 97.5^{th}$ percentile of the aggregate establishment count distribution). Establishment volatility is calculated by 1) weighting each establishment-year observation by the inverse of the ratio between size in that year and average size for the establishment, 2) applying a degree of freedom correction (Davis et al., 2007). All establishments (column 1); only establishments surviving throughout 1995-2005 (column 2). Source: AWFP.

Table 1.11: Degree of freedom correction vs. elimination of "low density" groups

	All	Cells: $N > 30$	d.f. correction
	Rescaled establishment count		
Uncertainty_k	0.0168^{***}	0.00576^{***}	0.00868^{***}
	(7.43)	(2.87)	(4.14)
$\text{Uncertainty}_k \times \text{Large}_m$	0.172^{***}	0.0945^{***}	0.115^{***}
	(4.42)	(2.60)	(3.09)
Observations	5040	5040	5040

t statistics in parentheses. Standard errors are clustered at the AMR-level.

* $p < .1$, ** $p < .05$, *** $p < .01$

Rescaled establishment count corresponds to the ratio between city/industry establishment count and the normalization factor ($N = 250 = 97.5^{th}$ percentile of the aggregate establishment count distribution). Industry-specific establishment-level uncertainty is measured by: sample standard deviation ($\times 100$) of employment growth rate demeaned by city/industry/year average (column 1), sample standard deviation ($\times 100$) of employment growth rate demeaned by city/industry/year average after elimination of low density, $N < 30$, groups (column 2), sample standard deviation ($\times 100$) of employment growth rate demeaned by city/industry/year average with degrees of freedom correction according to unbiased pooled variance estimator (column 3). Source: AWFP.

Table 1.12: Number of establishments by idiosyncratic uncertainty and market size: industry controls.

	Rescaled establishment count			
$\hat{\sigma}_k$	0.0213***	0.0309***	0.000277	0.00819***
	(6.30)	(5.19)	(0.29)	(7.57)
$\hat{\sigma}_k \times \text{Large}_m$	0.00258	0.0214	0.0229***	0.0292***
	(0.20)	(0.86)	(3.96)	(4.13)
SunkCost_k	-0.0234***	0.100***		
	(-9.01)	(2.73)		
$\text{Large}_m \times \text{SunkCost}_k$	-0.0149	0.227		
	(-1.04)	(1.11)		
$\hat{\sigma}_k \times \text{SunkCost}_k$		-0.00362***		
		(-3.26)		
$\hat{\sigma}_k \times \text{Large}_m \times \text{SunkCost}_k$		-0.00709		
		(-1.18)		
Public_k			0.0618***	0.503***
			(9.15)	(8.45)
$\text{Large}_m \times \text{Public}_k$			0.176***	0.528
			(3.76)	(1.42)
$\hat{\sigma}_k \times \text{Public}_k$				-0.0132***
				(-7.70)
$\hat{\sigma}_k \times \text{Large}_m \times \text{Public}_k$				-0.0105
				(-1.00)
Observations	1152	1152	2016	2016

t statistics in parentheses. Standard errors are clustered at the AMR-level.

* $p < .1$, ** $p < .05$, *** $p < .01$

Rescaled establishment count corresponds to the ratio between city/industry establishment count and the normalization factor ($N = 250 = 97.5^{th}$ percentile of the aggregate establishment count distribution). Baseline idiosyncratic uncertainty measure is used. Industry controls: sunk cost measured by the log of the ratio between investment in fixed tangible asstes and aggregate industry gross output multiplied by the employment share of a median-sized establishment, normalized by the minimum value (column 1-2); predominance of public ownership measured at the 3-digit industry-level corresponding to a dummy variable that takes value 1 if the share of establishments with public ownership structure is above the cross-industry median fraction, 0 otherwise (column 3-4). Source: AFWP, DESTATIS and EBDC.

Table 1.13: Heterogenous and endogenous loadings on aggregate shocks

	Rescaled establishment count
Uncertainty$_k$	.0009**
	(.0004)
Uncertainty$_k$ $\times$ Large$_m$	.0104***
	(.0012)
	Log average establishment size
Uncertainty$_k$	-.007**
	(.003)
Uncertainty$_k$ $\times$ Large$_m$	.006
	(.008)
Observations	4032

std.errors clustered at the AMR-level in parentheses.

* $p < .1$, ** $p < .05$, *** $p < .01$

Industry-specific establishment-level uncertainty is measured by sample standard deviation of estimated residuals according to eq.1.12 after exclusion of cells with less than 500 establishments. All non-tradable industries. Source: AWFP.

Table 1.14: Baseline results: bootstrapped standard errors

	All non-tradable industries	Retail and restaurants
	Rescaled establishment count	
Uncertainty$_k$	.002***	.022***
	(.001)	(.001)
Uncertainty$_k$ $\times$ Large$_m$	.012**	.017***
	(.002)	(.006)
	Log average establishment size	
Uncertainty$_k$	-.016***	-.048***
	(.004)	(.009)
Uncertainty$_k$ $\times$ Large$_m$	.012	-.016
	(.013)	(.030)
Observations	5040	1728

bootstrapped std. errors in parentheses.

* $p < .1$, ** $p < .05$, *** $p < .01$

Bootstrap procedure: 1) run establishment-level (stage 1) and city/industry-level (stage 2) regression and store residuals; 2) for each bootstrap sample of residuals from establishment-level regression, construct dependent variable sample, re-estimate it, generate uncertainty proxy, bootstrap sample of residuals from city/industry-level regression, construct dependent variable sample, re-estimate it using uncertainty proxy. Repeat 500 times. Source: AWFP.

Table 1.15: Value of estimated idiosyncratic uncertainty by industry: bottom uncertainty

Counter	WZ1993 Code	Description	Uncertainty
1	15130	Production of meat and poultrymeat products	.289
2	15812	Manufacture of fresh pastry, cakes, pies, tarts, etc.	.288
3	22220	Printing n.e.c.	.282
4	28111	Manufacture of metal structures and parts of structures, except for underground operations	.348
5	28521	Boring, milling, eroding, planing, lapping, broaching, levelling, sawing, welding, splicing, etc.	.322
6	33104	Manufacture of equipment for dental laboratories; manufacture of artificial teeth and other	.317
7	45211	Building of complete constructions or parts thereof; civil engineering without specialization	.377
8	45212	Building of complete constructions or parts thereof (excl. prefabricated constructions)	.381
9	45221	Erection of roofs, roof covering	.343
10	45223	Erection of frames and constructional timber works	.335
11	45310	Installation of electrical wiring and fittings	.33
12	45331	Installation of plumbing and sanitary equipment, gas fittings	.304
13	45332	Installation of heating, ventilation, refrigeration or air-conditioning equipment and ducts	.317
14	45420	Joinery installation	.341
15	45436	Equipping of rooms without specialization	.317
16	45441	Painting and lacquering	.334
17	45442	Glazing	.287
18	50103	Retail sale of motor vehicles	.308
19	50201	Maintenance and repair of motor vehicles (excluding electrical repairs)	.287
20	50203	Spraying and painting of motor vehicles	.319
21	50303	Retail sale of motor vehicle parts and accessories	.298
22	50501	Retail sale of automotive fuel in the name of a chain (filling stations acting as agencies)	.339
23	51190	Agents involved in the sale of a variety of goods	.306

Table 1.16: Value of estimated idiosyncratic uncertainty by industry: mid uncertainty

Counter	WZ1993 Code	Description	Uncertainty
24	51431	Non-specialized wholesale of electrical goods and accessories	.336
25	52111	Retail sale in non-specialized stores of food, beverages or tobacco	.345
26	52310	Dispensing chemists	.308
27	52320	Retail sale of medical and orthopaedic goods	.31
28	52332	Retail sale in drugstores	.331
29	52421	Retail sale of clothing in non-specialized stores	.333
30	52423	Retail sale of ladies' wear and clothing accessories	.331
31	52431	Retail sale of footwear	.3
32	52441	Retail sale of furniture	.315
33	52445	Retail sale of household furnishing articles of textile materials and carpets	.306
34	52451	Retail sale of electrical household appliances and electrical products n.e.c.	.325
35	52452	Retail sale of radio and television goods, sound recording or reproducing apparatus and	.335
36	52463	Retail sale of do-it-yourself material and equipment	.317
37	52471	Retail sale of stationery, school and office supplies	.315
38	52472	Retail sale of books and technical journals	.304
39	52483	Retail sale of flowers, plants, live animals, animal food etc. and seeds	.326
40	52484	Retail sale of precision, photographic and optical equipment, computers and software	.337
41	52485	Retail sale of watches, clocks, precious metal products and jewellery	.291
42	52487	Retail sale of bicycles, their parts and accessories, sports and camping goods (excluding	.326
43	55111	Hotels	.31
44	55301	Restaurants with service	.379
45	55303	Cafès	.368
46	55304	Ice-cream parlours	.368
47	55305	Snack bars	.367
48	55510	Canteens	.386

Table 1.17: Value of estimated idiosyncratic uncertainty by industry: top uncertainty

Counter	WZ1993 Code	Description	Uncertainty
49	65121	Activities of commercial banks including branches of foreign banks	.301
50	65125	Activities of credit cooperatives	.251
51	67201	Activities of insurance agents	.286
52	67202	Activities of insurance brokers	.303
53	70113	Activities of real estate developing companies dealing with residential buildings	.351
54	70310	Real estate agencies	.342
55	70320	Management of real estate on a fee or contract basis	.282
56	71100	Renting of automobiles	.372
57	72202	Software development and supply	.416
58	74112	Activities of law offices without notaries public	.305
59	74123	Activities of tax consultants, tax consulting companies	.288
60	74141	Business and management consultancy activities	.403
61	74201	Consulting architectural activities in building construction and interior design	.369
62	74204	Engineering activities for projects involving civil, hydraulic and traffic engineering	.342
63	74205	Engineering activities for projects in specific technical fields	.346
64	74301	Engineering control and analysis	.31
65	74401	Activities of advertising consultants, window-dressing	.386
66	74701	Cleaning of buildings, rooms and equipment	.431
67	74811	Activities of the photographic industry	.315
68	91311	Activities of churches and church-related religious organizations	.276
69	91337	Activities of other membership organizations n.e.c.	.323
70	92621	Activities of sports associations and clubs	.313
71	92711	Operation of gambling dens and slot machines	.378
72	93013	Dry cleaning and dyeing of clothes	.309
73	93021	Activities of barbers and hairdressers	.32

Table 1.18: Parameter estimates

Description	**Parameter**	**Value**
Relative city/industry-specific productivity	ϕ	.10
City-size specific firm productivity	κ	.25
Preference for differentiated good	α	3.13
Impact of competition	δ	.004
City-size specific sunk cost	χ	.57
Utility scaling factor for housing cost	β	1.36
Utility elasticity to amenities	$\lvert\psi\rvert$	.016

Bootstrapped standard errors under construction.

Table 1.19: Simulated impact of counterfactual 2.7 percent point decline in idiosyncratic uncertainty for all industries

	Levels		Change (%)	
	Smallest	Largest	Smallest	Largest
Employment	29597	470282	0.56%	-0.22%
	(29764)	(469254)		
Consumer surplus	1460	5503	-13.65%	-5.37%
	(1274)	(5216)		
Decomposition of consumer surplus change				
	Δ		$\Delta(\%)$	
	Smallest	Largest	Smallest	Largest
Prices fixed	1312	5349	-10.75%	-2.85%
Varieties fixed	1416	5347	-3.10%	-2.88%

In the upper left part of the table employment and consumer surplus are reported for the largest and smallest city when uncertainty is at the baseline level (2000) vs. 2.75 percent point lower $\forall k$ (in parenthesis). In the upper right part of the table the percent change between the two scenarios is reported for each variable and city. Consumer surplus measured in terms of the price for the homogenous good. In the bottom left part of the table consumer surplus is reported for the largest and smallest city in the sample following the 2.75 percent point decline in uncertainty holding prices/varieties fixed. In the bottom right part of the table the percent change in consumer surplus due to the decline in uncertainty is decomposed into the part due to varieties and prices respectively.

Table 1.20: Simulated impact of 6% proportional subsidy on sunk cost

	Policy intervention		Change (%)	
Outcome	Smallest	Largest	Smallest	Largest
Employment	29560	470271	-0.12%	0.00%
Consumer Surplus	1901	5905	26.37%	7.04%

The values reported in column 1-2 are relative to a scenario characterized by $\left\{\text{Uncertainty}_k^{Reduced}, (1 - \text{subsidy}) \times f_m\right\}$. The percent change in column 3-4 is with respect to the model solution obtained under baseline uncertainty. Consumer surplus measured in terms of the price for the homogenous good.

Empirical Implementation

Details on Sample Selection

Consistency of Estimated Common Shocks To account for small sample bias discussed in Section 1.2.1 we consider the pooled variance unbiased estimator given by eq.1.46:

$$x\hat{\sigma}_k^2 = \frac{1}{N_k - P_k} \sum_{p=1}^{P_k} \sum_{n=1}^{N_{k,p}} \left(g_{nk,p} - \overline{g}_{k,p}\right)^2 \tag{1.46}$$

where P_k is the total number of city/year cells for industry k and we compare it to the biased one:

$$\hat{\sigma}_k^2 = \frac{1}{N_k - 1} \sum_{p=1}^{P_k} \sum_{n=1}^{N_{k,p}} \left(g_{nk,p} - \overline{g}_{k,p}\right)^2 \tag{1.47}$$

In the absence of age/size-specific dummies in eq.1.7, the estimator in eq.1.47 is equivalent to the one in eq.1.8. The pooled variance estimator in eq.1.46 on the other hand applies a degrees of freedom correction to account for the noise associated with the estimation of P_k within-group averages. The degree of freedom correction eschews the concern of establishment-level uncertainty being underestimated in industries characterized by low establishment count on average across markets that introduces a source of positive bias in the coefficient of eq.1.1.

Since the specification in eq.1.7 does not only entail demeaning the growth rate of each establishment but also subtracting the predictable component attributable to size or age, we retain eq.1.8 as baseline measure of industry-specific idiosyncratic risk and restrict the attention to city/industry/year cells that satisfy a minimum density requirement, i.e. $N_{mkt} > 30$.

We show in Table 1.11 that this solution if anything seems to over-correct the bias associated with eq.1.47 relative to the correction available through eq.1.46. Specifically, we estimate eq.1.1 after having constructed three measures of idiosyncratic risk based on the sample variance of *employment growth rates*. The first one is constructed according to the biased estimator in eq.1.47 (column 2); the second one is constructed based on the unbiased estimator in eq.1.46 (column 3); the third and last one is constructed according to the biased estimator in eq.1.47 after having discarded all city/industry/year cells featuring less than 30 observations (column 1). The non-interacted coefficient is even smaller compared the one obtained using the unbiased estimator, thus suggesting that, if anything, the crude method of eliminating low-density cells seems to over-correct the mechanical bias induced by inconsistent estimates of city/industry/year averages in eq.1.7 when estimated over the entire sample.

Sample Selection and Industry Imputation We provide here details on the establishment sample selection. Given unavailability of harmonized 5-digit industry codes over the period considered we impute the industry in which an establishment operated for 1994-1997 and 2004-2007 as follows:

1. we first drop all establishments changing 5-digit industry during 1998-2003 (only 1.5% drop of establishments during the selected timeframe).

2. we next impute the product category for establishments that exist during the 1998-2003 timeframe conditional on 1) having reported at least 4 times the same product category and 2) being the first/last year in which they report information on the industry they operate into distant no more than 4 years from the year of imputation, respectively for years $y = 1994/1997$ and years $y = 2004/2007$.[37]

Before imputing the industry there are are approximately 20.872.000 observations in the timeframe considered. Of these, around 12 millions correspond to years for which the WZ 1993 5-digit classification was not in use. 85% of establishment-year observations correspond to continuing establishments. After applying the imputation procedure just described, the sample consists of approximately 14.208.000 observations. Of these about 87% are continuing and the fraction of continuing establishments is higher at the beginning and at the end of the sample by construction. We drop establishments operating in Agriculture and Mining, Utilities and Transportation, Public Administration. We are left with approximately 13.158.000 observations: the median stay in the sample is about 8 years and out of a total of 1.686.000 establishments, only 26% survive (in the sense of reporting positive employment) throughout the timeframe considered. Of 13.588.000 observations, 6.043.000 are located in urban areas (the Metropolitan Areas selected).

Details on Model Solution

In the version of Luttmer (2007) without endogenous imitation, the productivity of entrants grows at rate θ_E, while incumbents productivity grows deterministically with age at a industry-specific rate $\theta_{I,k}$. The drift and variance of the stochastic

[37]This means that the only chance for an establishment in 1994 to be in the final sample is for this establishment to have been active at least four years during 1998-2003 and reporting positive employment in 1998. Similarly, the only chance for an establishment in 2007 to be in the final sample is for this establishment to have been active at least four years during 1998-2003 and reporting positive employment in 2003. An establishment entering in 2000 and reporting positive employment every year through 2010 is not in the sample since it has reported information on the product category only for 3 years.

process for firm specific employment depend on 1) the difference between the growth rate of entrants productivity and the growth rate of incumbents productivity, 2) the variance of the stochastic process for firm specific productivity, 3) the elasticity of employment to productivity, equal to the elasticity of substitution minus 1. Set this elasticity constant across industries and equal to $1/(1-\beta)$. Then the drift and variance for employment are given by:

$$\begin{bmatrix} \mu_k \\ \sigma_k \end{bmatrix} = \tfrac{\beta}{1-\beta} \begin{bmatrix} \theta_{I,k} - \theta_E \\ \sigma_{Z,k} \end{bmatrix} \tag{1.48}$$

Luttmer (2007) characterizes the equilibrium firm size density conditional on a given age for the firm by solving the Kolmogorov forward equation associated with the continuous-time stochastic process for firm-specific employment conditional on two boundary conditions: 1) as age goes to zero, the size distribution must converge to the size distribution for entrants, 2) the mass of newly born firms with size lower than the size cutoff is zero. The solution yields a clear expression for the shape of firm size distribution *conditional* on 1) all firms entering with the same productivity level, and 2) this being equal to the size cutoff:

$$\begin{aligned} \theta_k = -\tfrac{\mu_k}{\sigma_k^2} + \sqrt{\left(\tfrac{\mu_k}{\sigma_k^2}\right)^2 + \tfrac{\eta}{\sigma_k^2/2}} \quad if \quad \eta \geq 0 \\ \theta_k = -\tfrac{\mu_k}{\sigma_k^2} \quad if \quad \eta = 0 \end{aligned} \tag{1.49}$$

where η is the rate at which labor supply grows in all industries. If $\eta \geq 0$ and $\eta > \mu_k + \sigma_k^2/2$, then $\theta_k > 1$ and the firm size distribution has well-defined mean.

Re-define $\varphi_k = \frac{\mu_k}{\sigma_k^2}$. The size cutoff is endogenously determined and given by:

$$e_{m,k} = \left(\frac{\varphi_k + \sqrt{\varphi_k^2 + \frac{r-\kappa}{\sigma_k^2/2}}}{1 + \varphi_k + \sqrt{\varphi_k^2 + \frac{r-\kappa}{\sigma_k^2/2}}} \right) \left(1 - \frac{\mu_k + \sigma_k^2/2}{r-\kappa} \right) \tag{1.50}$$

where r stands for the interest rate and κ the rate of growth of per capita consumption along the balanced growth path. These are both endogenous objects given respectively by: $r = \rho + \gamma\kappa$ and $\kappa = \theta_E + \left(\frac{1-\beta}{\beta}\right)\eta$. The interest rate depends negatively on the intertemporal elasticity of substitution $1/\gamma$ and positively on the discount rate ρ. The rate of growth of per capita consumption depends on the common growth rate of entrants productivity and on population growth. It is necessary that 1) $r - \kappa > 0$, so that utility is finite, and that 2) $r - \kappa > \mu_k + \sigma_k^2/2$ for all k, for the size cutoff to be positive and well-characterized.

The equilibrium firm size distribution in industry k follows a Pareto distribution:

$$\text{Prob}(e_k > q) = \left(\frac{e_{m,k}}{q}\right)^{\theta_k} \tag{1.51}$$

In order to derive the associated firm productivity distribution, we follow the same steps as di Giovanni and Levchenko (2012). Firm productivity is related to firm size by: $z_k = e_k^{(1-\beta)/\beta}$. Then:

$$\begin{aligned}
\text{Prob}(z_k > q) &= \text{Prob}(e_k^{(1-\beta)/\beta} > q) \\
\Leftrightarrow \text{Prob}(z_k > q) &= \text{Prob}(e_k > q^{\beta/(1-\beta)}) \\
\Leftrightarrow \text{Prob}(z_k > q) &= \left(\frac{e_{m,k}}{q^{\beta/(1-\beta)}}\right)^{\theta_k} \\
\Leftrightarrow \text{Prob}(z_k > q) &= \left(\frac{e_{m,k}^{(1-\beta)/\beta}}{q}\right)^{\theta_k \beta/(1-\beta)}
\end{aligned} \tag{1.52}$$

The location parameter is $z_{m,k} = e_{m,k}^{(1-\beta)/\beta}$ and the Pareto tail is provided by $\theta_k^z = \theta_k \beta/(1-\beta)$.

We calibrate externally $\{\beta, \eta, \theta_E, \gamma, \rho\}$ and estimate $\theta_{I,k}$ and $\sigma_{Z,k}$. The external parameters are chosen analogously to Luttmer (2007):

Table 1.21: External parameters in productivity calibration

Parameter	**Description**	**Value**
$1/(1-\beta)$	Elasticity of substitution of varieties (5-digit industries)	6
η	Population growth	.01
θ_E	To match growth of consumption in BGP $\kappa = .02$	.018
γ	Inverse of intertemporal elasticity of substitution	2
ρ	Discount factor	.99

Next, we back-out μ_k and σ_k using eq.1.48 given $\widehat{\theta}_{I,k}$ and $\widehat{\sigma}_{Z,k}$. In particular, we re-estimate eq.1.7 separately for each 3-digit industry.[38] We weight the estimated average growth rate for establishments in a given industry at different stages of the life-cycle, $\widehat{\beta}_a$ for $a \in K_a$ (the number of age dummies), according to weights, ω_a, that are proportional to how many years does each age dummy

[38]The trade-off is between more precise estimates and a higher level of disaggregation. Re-estimating eq.1.7 to back-out 5-digit industry-specific estimates for $\theta_{I,k}$ should not alter the outcome of the calibration.

span.[39] Table 1.22 reports the summary statistics for $\widehat{\theta}_{I,k} = \left(\frac{1-\beta}{\beta}\right) \sum_{a=1}^{K_a} \omega_a \widehat{\beta}_a$ in each sector. Finally, we set $\widehat{\sigma}_{Z,k} = \left(\frac{1-\beta}{\beta}\right) \widehat{\sigma}_k$.

Table 1.22: Estimated 3-digit industry specific incumbents employment growth $\widehat{\theta}_{I,k}$

Sector	**Min**	**Mean**	**Max**
Business Services	0.02	0.04	0.05
Construction	-0.03	-0.00	0.01
Finance	-0.06	-0.02	0.02
Hospitality	-0.01	0.01	0.04
Manufacturing	-0.05	0.01	0.04
Other Services	0.00	0.01	0.02
Real Estate	0.02	0.02	0.02
Retail and Wholesale	-0.03	0.03	0.07
Total	-0.06	0.02	0.07

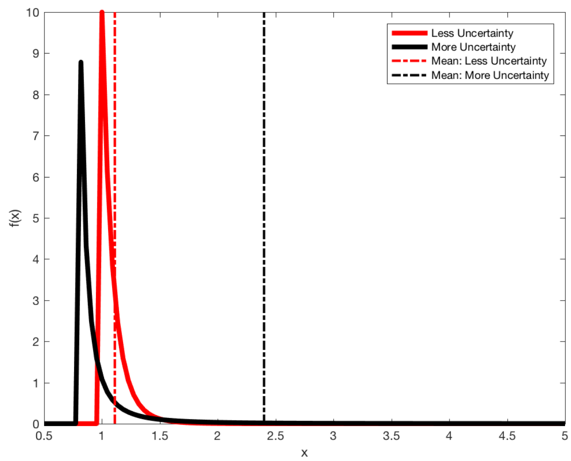
Estimated employment growth by age for 3-digit industries. The estimates are obtained by weighting $\widehat{\beta}_a$ in eq.1.7 run at the 3-digit level according to weights that are proportional to how many years does each age dummy span.

Fig.1.7 reports the stationary productivity distribution implied by lower uncertainty (red solid line), and higher uncertainty (black solid line). Two are the consequences of a decline in uncertainty. The first one is a reduction in dilation: the right tail shifts in, as fewer firms in the cross-section are likely to have received a sequence of favorable and sizeable productivity shocks. The second one is an increase in selection: absent the opportunity to grow in the future, the value of a firm characterized by low productivity is now sufficiently low to force this firm out of the market.

We report in Table 1.23 and Table 1.25 the estimates for $z_{m,k}$ and θ_k^z in each industry after having sorted industries in ascending order for θ_k^z. The median value for θ_k^z, 5.37, is very close to 5.3 in di Giovanni and Levchenko (2012). The median is 5.77 and the standard deviation 2.63.

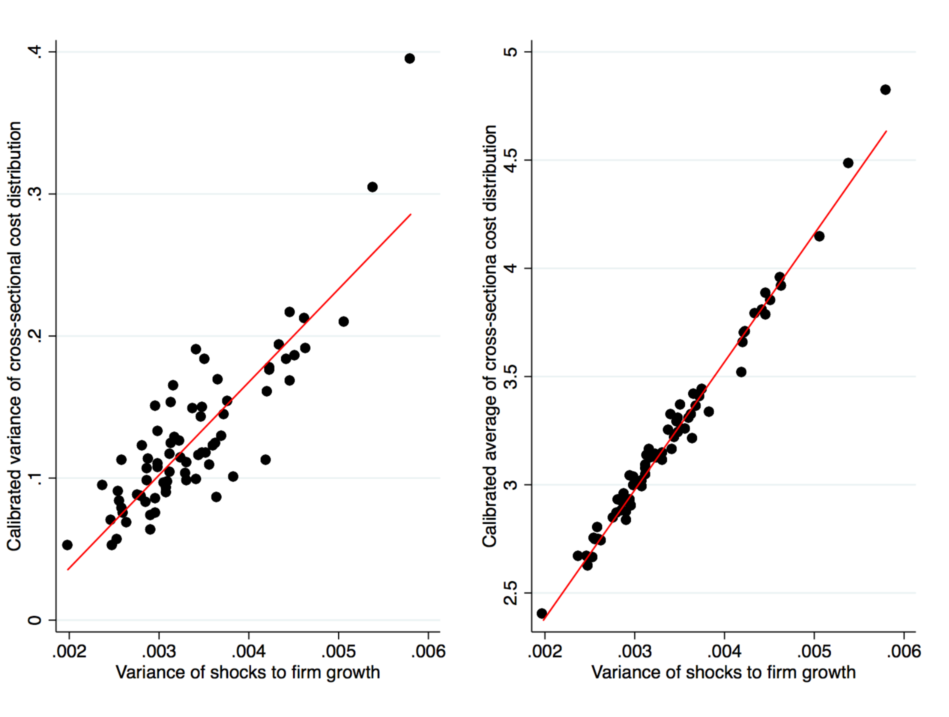
[39]Specifically, remember that $a \in K_a$ is for age $\in \{0, 1-4, 5-9, 10-15, 16+\}$: the weights are $\omega_1 = 4$ years, ω_2 = 5 years, ω_3 = 6 years and ω_4 = 15 years.

Figure 1.7: Stationary productivity distribution implied by different degrees of uncertainty



High = black; low = red. Dashed line indicates the expected value.

Figure 1.8: First and second moment of unconditional marginal cost distribution



Industries are labeled by sector.

Table 1.23: Estimated minimum and shape of productivity distribution

$z_{m,k}$	θ_k^z	Industry Description
.94	2.17	Retail sale in drugstores
.92	2.34	Business and management consultancy activities
.94	2.36	Retail sale of medical and orthopaedic goods
.94	2.37	Dispensing chemists
.92	2.9	Software development and supply
.93	2.95	Canteens
.94	3.13	Retail sale of automotive fuel in the name of a chain (filling stations acting as agencies)
.92	3.28	Cleaning of buildings, rooms and equipment
.93	3.51	Consulting architectural activities in building construction and interior design
.95	3.59	Activities of law offices without notaries public
.93	3.69	Activities of advertising consultants, window-dressing
.93	3.7	Operation of gambling dens and slot machines
.94	3.91	Engineering activities for projects in specific technical fields
.95	3.92	Activities of tax consultants, tax consulting companies
.94	3.99	Engineering activities for projects involving civil, hydraulic and traffic engineering
.94	4.31	Boring, milling, eroding, planing, lapping, broaching, levelling, sawing, welding, splicing, etc.
.94	4.33	Real estate agencies
.95	4.58	Activities of the photographic industry
.95	4.78	Spraying and painting of motor vehicles
.94	4.86	Manufacture of metal structures and parts of structures, except for underground operations
.94	4.89	Restaurants with service
.94	4.99	Non-specialized wholesale of electrical goods and accessories
.94	5.01	Retail sale of precision, photographic and optical equipment, computers and software

Ascending order of industries according to the shape parameter. First 33% section of industries is reported.

Table 1.24: Estimated minimum and shape of productivity distribution

$z_{m,k}$	θ^z_k	Industry Description
.94	5.07	Retail sale of radio and television goods, sound recording or reproducing apparatus and
.95	5.11	Manufacture of fresh pastry, cakes, pies, tarts, etc.
.94	5.12	Retail sale of clothing in non-specialized stores
.94	5.13	Ice-cream parlours
.94	5.15	Retail sale of ladies' wear and clothing accessories
.94	5.15	Cafès
.94	5.17	Snack bars
.95	5.29	Agents involved in the sale of a variety of goods
.95	5.3	Retail sale of bicycles, their parts and accessories, sports and camping goods
.95	5.31	Retail sale of flowers, plants, live animals, animal food etc. and seeds
.95	5.32	Retail sale of electrical household appliances and electrical products n.e.c.
.95	5.37	Activities of sports associations and clubs
.95	5.55	Activities of insurance brokers
.95	5.58	Retail sale of do-it-yourself material and equipment
.94	5.59	Joinery installation
.95	5.66	Retail sale of furniture
.95	5.66	Retail sale of stationery, school and office supplies
.95	5.75	Maintenance and repair of motor vehicles (excluding electrical repairs)
.95	5.78	Painting and lacquering
.95	5.82	Retail sale of motor vehicles
.95	5.89	Retail sale of motor vehicle parts and accessories
.95	5.96	Retail sale of household furnishing articles of textile materials and carpets
.95	6.03	Retail sale of books and technical journals
.95	6.09	Manufacture of equipment for dental laboratories; manufacture of artificial teeth and other

Ascending order of industries according to the shape parameter. Mid 33% section of industries is reported.

Table 1.25: Estimated minimum and shape of productivity distribution

$z_{m,k}$	θ_k^z	Industry Description
.95	6.11	Management of real estate on a fee or contract basis
.95	6.16	Retail sale of footwear
.95	6.19	Activities of insurance agents
.95	6.22	Activities of barbers and hairdressers
.95	6.32	Installation of electrical wiring and fittings
.95	6.38	Equipping of rooms without specialization
.95	6.39	Activities of other membership organizations n.e.c.
.95	6.52	Retail sale of watches, clocks, precious metal products and jewellery
.95	6.64	Dry cleaning and dyeing of clothes
.94	6.66	Building of complete constructions or parts thereof (excl. prefabricated constructions)
.95	6.72	Production of meat and poultrymeat products
.95	6.78	Installation of heating, ventilation, refrigeration or air-conditioning equipment and ducts
.94	6.79	Building of complete constructions or parts thereof; civil engineering without specialization
.95	7.36	Installation of plumbing and sanitary equipment, gas fittings
.95	7.7	Glazing
.95	7.76	Hotels
.95	7.89	Retail sale in non-specialized stores of food, beverages or tobacco
.95	8.15	Erection of roofs, roof covering
.95	8.53	Erection of frames and constructional timber works
.96	8.57	Activities of churches and church-related religious organizations
.96	13.67	Activities of commercial banks including branches of foreign banks
.96	14.29	Printing n.e.c.
.97	19.31	Activities of credit cooperatives

Ascending order of industries according to the shape parameter. Last 33% section of industries is reported.

Finally, we calculate the unconditional first and second moment of the marginal cost distribution given by $c_k = 1/z_k$:

$$E(c_k) = \frac{1}{z_{m,k}} \frac{\theta_k^z}{\theta_k^z + 1} \quad Var(c_k) = \left(\frac{1}{z_{m,k}}\right)^2 \frac{\theta_k^z}{(\theta_k^z+1)^2(\theta_k^z+2)} \tag{1.53}$$

and plot $E(c_k)$ and $Var(c_k)$ vs. σ_k in Fig.1.8. The Pareto tail is declining in residual dispersion: in principle, if the correlation between μ_k and σ_k was negative the Pareto tail could be increasing in residual dispersion, since a smaller drift implies a less dispersed firm size distribution. This is not the case since the correlation between μ_k and σ_k is .08 and statistically insignificant.

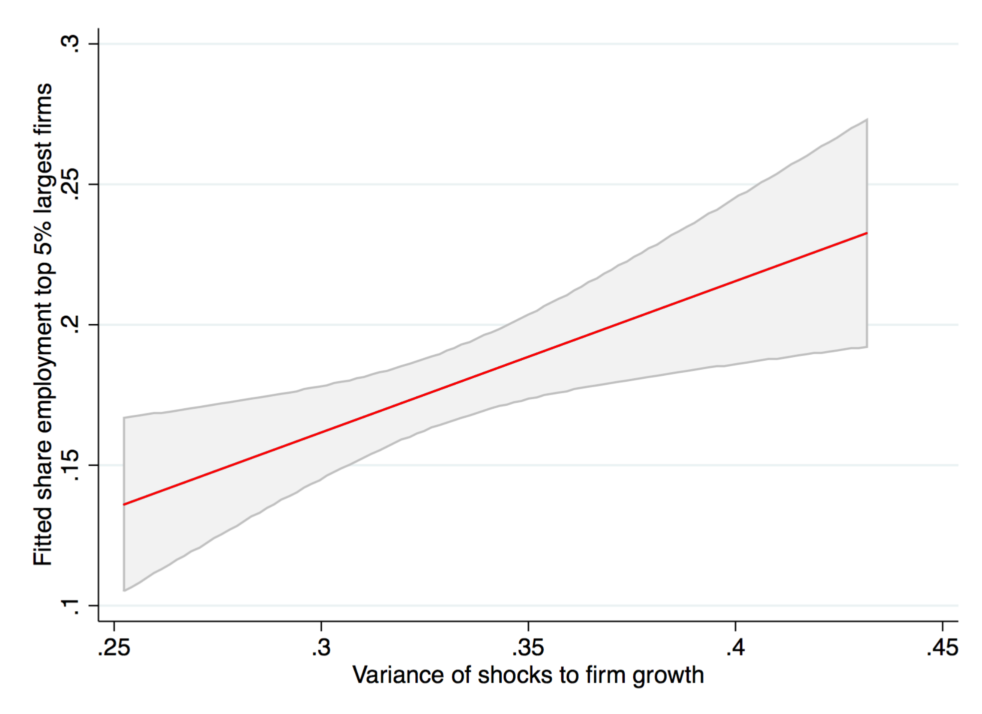
The maximum of the cost distribution is increasing in residual dispersion. However, the variation in the calibrated location parameter is very small, as one can see from the first column in Table 1.25, and insufficient to offset the positive (negative) correlation between average (dispersion of the) marginal cost and the Pareto tail. Thus, industries where the volatility of residuals is very high should be characterized by a steady state marginal cost distribution that has lower mean and higher variance.

While the framework just described provides an intuitive mapping between idiosyncratic uncertainty and the cross-sectional productivity distribution, it is worth considering whether the predictions of such framework are aligned with the actual degree of concentration registered across industries characterized by various degrees of uncertainty. Specifically, are industries corresponding to higher idiosyncratic uncertainty also more concentrated? The answer is positive according to Fig.1.9, where we plot the linear prediction of the employment share for the top 5% largest establishments based on idiosyncratic uncertainty according to the regression:

$$\text{share top 5\% largest establishments}_{k,t} = \gamma_{sector} + \gamma_t + \beta \text{uncertainty}_k + \varepsilon_{k,t} \tag{1.54}$$

Additionally, concentration at the aggregate level as measured in terms of the employment share in the hands of the top 5% largest establishments has gone down from .66 in 1995 to .63 in 2014.

Figure 1.9: Industry concentration as predicted by industry idiosyncratic uncertainty



On the y-axis is: fitted share top 5% largest establishments$_{industry,t}$ $= \widehat{\beta}$uncertainty$_{industry}$ + $\widehat{\varepsilon}_{industry,t}$. Observations are 5-digit/year non-tradable industries $(t = \{1998, 2000, 2002\})$.

Solving the model The estimation of the model is run over a specified region for the parameter space. The set of parameters to estimate is $\Theta = \{\kappa, \alpha, \delta, \chi, \beta, \psi\}$. The parameter space is not unconstrained. In particular the rescaled per capita sunk entry cost must be bounded to ensure that $N_{mk} \in (0, \widetilde{N})$, where $\widetilde{N}$ corresponds to the measure of firms that would turn demand negative given the parameters, $\widetilde{N} = \frac{1}{\delta}\frac{\alpha-\phi-\{b_k\}_{max}}{\{b_k-\mu_k\}_{max}}$. f_L and f_H defined in the main text thus correspond to:

$$f_L = \{\sigma_k^2\}_{max} + (\alpha - \{\mu_k\}_{min})^2 \left(1 + \delta\widetilde{N}\right)^{-2}$$
$$f_H = \{\sigma_k^2\}_{min} + (\alpha - \phi - \{\mu_k\}_{max})^2$$

We account for the non-linear restrictions on the parameters by allowing f_m to be a convex combination between the lower and upper boundary, which are themselves a function of α, ϕ, b_k and θ_k $\forall k$. This approach blurs the interpretation of α, but it has the advantage of achieving a wider search region.

We do not estimate ϕ, the parameter governing the relative magnitude between city and industry-specific component of average marginal cost. Instead, we solve the model on a grid for $\phi = \{.01, .025, .05, .1\}$, and select the model solution that yields the closest match between the elasticity of the city-specific component of firm productivity on city size in the model and the average available empirical estimate of 5%.

We set $\phi + \{b_k\}_{max} \leq \alpha \leq \overline{\alpha}$ and find ε such that:

- $(\phi + \{b_k\}_{max}) \times (1+\varepsilon) \leq \alpha \leq (1-\varepsilon) \times \overline{\alpha}$ (with $\overline{\alpha} = 5$),
- $f_H - f_L$ is maximized.

Finally, remember that congestion forces must dominate agglomeration forces for the equilibrium to be stable. We find numerically a lower bound for β, $g_\beta(\kappa, \alpha, \delta, \chi)$, and estimate $\widetilde{\beta}$, where $\beta = \exp(\widetilde{\beta})g_\beta(\kappa, \alpha, \delta, \chi)$.

We solve the model in the following steps:

1. For a given value of Θ we set up a 1×100.000 grid for population such that $0 < L_m < \overline{L}$ and calculate the solution to the labor demand block of the model and indirect utility for each value. We numerically compute the derivative of the indirect utility $\widetilde{V^i}(L_m)$ by means of a first-order Taylor expansion. We set:

$$g_\beta(\kappa, \alpha, \delta, \chi) = \max\left\{\frac{\frac{\partial(h(L_m)+CS(L_m))}{\partial L_m}}{\frac{\partial L_m^\xi}{\partial L_m}}\right\} \tag{1.55}$$

2. We next solve for the spatial equilibrium according to an iterative procedure.

   (a) We set $\Omega_j = M + 1 - j$ and start from a uniform initial guess for market size in each city $m \in \Omega_j$.

      i. Given L_m for $m = 1, 2, .., \Omega_j$ we solve for the labor demand side of the economy.

      ii. Given N_{mk} for $m = 1, 2, ..., \Omega_j$ and $k = 1, 2, ..., K$ we solve for the spatial equilibrium such that:

$$L_m^{sol} = \left(\frac{1}{\beta}\left(\tau^i + h_m^i + CS(L_m^i) - \frac{\overline{V}^i}{A_m}\right)\right)^{1/\xi} \tag{1.56}$$

      for $m = 1, 2, ..., \Omega_j$ jointly with the labor market clearing condition to pin down $\overline{V}^i$.

      iii. We update L_m^{i+1} according:

$$L_m^{i+1} = L_m^i + .5|L_m^{sol} - L_m^i| \times sign(L_m^{sol} - L_m^i) \tag{1.57}$$

   (b) We repeat step a) to c) until convergence or until $L_m^{i+1} < 0$ for some m.

   (c) If $L_m^{i+1} < 0$ for some m (non-negativity constraint), we go back to step (a) and drop the lowest ranked city.

Notice that throughout this procedure, it is necessary to check that $\tau = \int_\Omega l_m r_m dm$ is greater than τ_{min}, the minimum transfer for the largest city to feature positive consumption of the homogenous good.
By $0 < \xi < 1$, $\tau = \int_\Omega L_m^{1+\xi} dm$ is increasing in expected value and dispersion but decreasing in expected value and dispersion *jointly*.

$$\tau \approx \mu_l^{1+\xi} + \frac{1}{2}\xi\mu_l^{\xi-1}\sigma_l^2 \tag{1.58}$$

Both mean, μ_l, and dispersion in city size, σ_l^2, are increasing at each new iteration of the *inner* loop of the contraction procedure illustrated above. They are also increasing at each new iteration of the *outer* loop of the contraction procedure illustrated above: anytime that a location must be dropped for the non-negativity constraint to be satisfied, the new equilibrium size for the remaining locations must be at least as high/low as it was during the last iteration of the previous outer loop for top/bottom locations.
Hence, anytime that $\tau < \tau_{min}$ the procedure described above is interrupted, and a very large value assigned to the loss function corresponding to the combination

of parameters being evaluated.
In spite of being in violation with data on housing prices, assuming $\xi > 1$ would not force to rule out regions of the parameter space for which the equilibrium city size distribution features too high mean and dispersion, which we know to be the empirically relevant case.

Theory and Proofs

Proofs

Proposition 1: Consider the solution for the measure of varieties and the price level:

$$
\begin{aligned}
N_{mk} &= \frac{1}{\delta}\left(\frac{\alpha-\mu_{mk}(s)}{\sqrt{f_m-\sigma_k^2(s)}}-1\right)\\
p_{mk} &= \frac{1}{2}\sqrt{f_m-\sigma_k^2(s)}+\mu_{mk}(s)
\end{aligned}
\tag{1.59}
$$

Differentiating with respect to size, uncertainty and the interaction between the two:

$$
\begin{aligned}
\frac{\partial N_{mk}}{\partial s} &= \frac{1}{\delta}\left(\frac{1}{2}\frac{\alpha-\mu_{mk}(s)}{\left(f_m-\sigma_k^2(s)\right)^{3/2}}\left(\sigma_k^2(s)\right)'-\frac{1}{\left(f_m-\sigma_k^2(s)\right)^{1/2}}\mu_k'(s)\right)>0\\
\frac{\partial p_{mk}}{\partial s} &= -\frac{1}{2}\left(f_m-\sigma_k^2(s)\right)^{-1/2}\left(\sigma_k^2(s)\right)'+\mu_k'(s)<0\\
\frac{\partial N_{mk}}{\partial L} &= \frac{1}{\delta}\left(-\frac{1}{2}\frac{\alpha-\mu_{mk}(s)}{\left(f_m-\sigma_k^2(s)\right)^{3/2}}\left(f_m'\right)-\frac{1}{\left(f_m-\sigma_k^2(s)\right)^{1/2}}\left(\mu_m'\right)\right)>0\\
\frac{\partial p_{mk}}{\partial L} &= \frac{1}{2}\left(f_m-\sigma_k^2(s)\right)^{-1/2}\left(f_m'\right)+\left(\mu_m'\right)<0\\
\frac{\partial^2 N_{mk}}{\partial L\partial s} &= \frac{1}{\delta}\left(-\frac{3}{4}\frac{\alpha-\mu_{mk}(s)}{\left(f_m-\sigma_k^2(s)\right)^{5/2}}\left(f_m'\right)+\frac{1}{2}\frac{1}{\left(f_m-\sigma_k^2(s)\right)^{3/2}}\left(\mu_m'\right)\right)\left(\sigma_k^2(s)\right)'>0\\
\frac{\partial^2 p_{mk}}{\partial L\partial s} &= \frac{1}{4}\left(f_m-\sigma_k^2(s)\right)^{-3/2}\left(f_m'\right)\left(\sigma_k^2(s)\right)'<0
\end{aligned}
\tag{1.60}
$$

provided that $\mu_k' < 0$ and $(\sigma_k^2)' > 0$, i.e., the average productivity and the dispersion of productivities increases as uncertainty rises, and where we have represented by s the "state" of the world, i.e., the degree of uncertainty.
Hence, the measure of variety (price level) is increasing (decreasing) in uncertainty, market size, and the combined effect of both.

Next, consider the consumer surplus:

$$
CS_{mk} = \frac{1}{2}\left\{\left(\frac{N_{mk}}{\gamma+\eta N_{mk}}\right)\left(\alpha-p_{mk}\right)^2+\frac{1}{4\gamma}\sigma_k^2(s)\right\}
$$

Consumer surplus, hence, depends directly on uncertainty, through the dispersion in consumer prices $\sigma_k^2(p) = \frac{1}{4}\sigma_k^2$, and indirectly through the measure of varieties and the level of prices. Substituting in the solution for the measure of variety and the level of prices:

$$CS_{mk} = \frac{1}{2\eta}\left[(\alpha - \mu_{mk}(s)) - \lambda_{mk}\right]\left[(\alpha - \mu_{mk}(s)) - \frac{1}{2}\lambda_{mk}\right] + \frac{1}{8\gamma}\sigma_k^2(s) \quad (1.61)$$

with $\lambda_{mk} = \left(f_m - \sigma_k^2(s)\right)^{1/2}$.

Notice that consumer surplus depends on γ through the dispersion of consumer prices: intuitively, the lower is γ, the more substitutable varieties are, so that the consumer takes more advantage of highly dispersed prices. Also, notice that consumer surplus depends on η individually, through the imperfect substitutability between the extensive margin - N_{mk} - and the intensive margin - $\frac{1}{\gamma+\eta N_{mk}}$. Hence, in principle γ and η could be identified separately through the spatial equilibrium condition. Simulations, show, however, that the size distribution is very mildly responsible to variation in γ, which is why we choose to abstract from estimation of γ and η separately.

$$\begin{aligned}\frac{\partial CS_{mk}}{\partial s} &= \frac{1}{2\eta}\left[-\mu_k'(s) + \frac{1}{2}\left(f_m - \sigma_k^2(s)\right)^{-1/2}\left(\sigma_k^2(s)\right)'\right]\left[(\alpha - \mu_{mk}(s)) - \frac{1}{2}\lambda_{mk}\right] + \\ &+ \frac{1}{2\eta}\left[(\alpha - \mu_{mk}(s)) - \lambda_{mk}\right]\left[-\mu_k'(s) + \frac{1}{4}\left(f_m - \sigma_k^2(s)\right)^{-1/2}\left(\sigma_k^2(s)\right)'\right] + \\ &+ \frac{1}{8\gamma}\left(\sigma_k^2(s)\right)' > 0 \end{aligned} \quad (1.62)$$

$$\begin{aligned}\frac{\partial CS_{mk}}{\partial L} &= \frac{1}{2\eta}\left[-\mu_m' - \frac{1}{2}\left(f_m - \sigma_k^2(s)\right)^{-1/2} f_m'\right]\left[(\alpha - \mu_{mk}(s)) - \frac{1}{2}\lambda_{mk}\right] + \\ &+ \frac{1}{2\eta}\left[(\alpha - \mu_{mk}(s)) - \lambda_{mk}\right]\left[-\mu_m' - \frac{1}{4}\left(f_m - \sigma_k^2(s)\right)^{-1/2} f_m'\right] > 0 \end{aligned} \quad (1.63)$$

The joint effect of an increase in uncertainty and city size is also positive: the formal derivation is included in the working paper version of the chapter.

Hence, consumer surplus is increasing in uncertainty and city size and in the combined effect of an increase in both of them.

Proposition 2: Consider indirect utility associated with location m:

$$V_m^i = A_m\left(\tau + \omega_m - r_m + CS_m^i\left(L_m\right)\right) \quad (1.64)$$

In equilibrium it must be:

$$\overline{V} - A_m \left(\tau + \omega_m - r_m + CS_m^i\left(L_m\right)\right) = 0 \tag{1.65}$$

Applying the implicit function theorem to eq.1.65:

$$\frac{\partial L_m}{\partial A_m} = -\frac{\tau + \omega_m - r_m + CS_m^i\left(L_m\right)}{A_m\left(\frac{\partial}{\partial L_m}\left(CS_m^i\left(L_m\right) - r_m + \omega_m\right)\right)} \tag{1.66}$$

Since demand of the numeraire good must be positive:

$$q_{0,m}^i = \tau + \omega_m - r_m - \sum_{k \in K} \int_{N_{mk}} p_{mkn} q_{mkn}^i dn > 0$$

and:

$$CS_m^i\left(L_m\right) = \sum_{k \in K}\left[\alpha \int_{N_{mk}} q_{mkn}^i dn - \frac{\gamma}{2}\int_{N_{mk}}\left(q_{mkn}^i\right)^2 dn - \frac{\eta}{2}\left(\int_{N_{mk}} q_{mkn}^i dn\right)^2\right] + \\ - \int_{N_{mk}} p_{mkn} q_{mkn}^i dn$$

with $\alpha \int_{N_{mk}} q_{mkn}^i dn - \frac{\gamma}{2}\int_{N_{mk}}\left(q_{mkn}^i\right)^2 dn - \frac{\eta}{2}\left(\int_{N_{mk}} q_{mkn}^i dn\right)^2 > 0$ for $k = 1, 2, ..., K$ and $m = 1, 2, ..., M$, then $\tau + \omega_m - r_m + CS_m^i\left(L_m\right) > 0$ for $m = 1, 2, ..., M$.

If $A_m\left(\frac{\partial}{\partial L_m} CS_m^i\left(L_m\right) - r_m + \omega_m\right) < 0$ for $L_m \in \mathbf{L}$, then $\frac{\partial L_m}{\partial A_m} > 0$. It follows that the equilibrium is unique, i.e. there exists only one solution L_m given A_m and $\overline{V}$ such that eq.1.65 holds. Furthermore, the equilibrium is stable. To see why, consider moving a mass ε of consumers from away from their initial location m to market m'. Since indirect utility is strictly decreasing in city size, the utility attached to the new location must be strictly lower than the one experienced in the previous location. Thus, the only equilibrium is also stable.

The Homogenous Good

Two conditions need to be satisfied for internal consistency of the equilibrium: 1) production of the homogenous good must be positive in order for the wage in the differentiated goods sector to be pinned down; 2) demand for the homogenous good must also be positive in order for the demand function to be as in the text. The homogenous good is assumed to be perfectly tradable across locations.

The initial endowment consistent with market clearing for the homogenous good is:

$$\tau = \int_{\Omega} l_m r_m dm = \beta \int_{\Omega} L_m^{1+\xi} dm \tag{1.67}$$

This value must be consistent with the condition of positive demand in all locations:

$$\tau_{min} + h_m - r_m - \sum_{k\in K} \int_{N_{mk}} p_{mnk} q^i_{mnk} dn > 0 \tag{1.68}$$

By Proposition 3:

$$\frac{\partial}{\partial L_m} \left\{ h_m - r_m + \sum_{k\in K} \left[CS^i_{mk} - \int_{N_{mk}} p_{mnk} q^i_{mnk} dn \right] \right\} < 0 \tag{1.69}$$

Since $\frac{\partial CS^i_{mk}}{\partial L_m} > 0$, then:

$$\frac{\partial}{\partial L_m} \left\{ h_m - r_m - \sum_{k\in K} \int_{N_{mk}} p_{mnk} q^i_{mnk} dn \right\} < 0 \tag{1.70}$$

Thus a sufficient requirement for τ is:

$$\tau_{min} > -h_m \left(L^{max}\right) + r_m \left(L^{max}\right) + \sum_{k\in K} \int_{N_{mk}} p_{mnk} \left(L^{max}\right) q^i_{mnk} \left(L^{max}\right) dn) \tag{1.71}$$

or, after substituting $\sum_{k\in K} \int_{N_{mk}} p_{mnk} q^i_{mnk} dn = \sum_{k\in K} \int_{N_{mk}} \left(l_{mkn} + f\right) dn$:

$$\tau_{min} > -\left(1 - \text{share}^{Diff.}\right) h_m \left(L^{max}\right) + r_m \left(L^{max}\right) \tag{1.72}$$

since the chosen calibration for h_m implies that expenditure on either sector is fixed and independent of market size. The restriction $\tau > \tau_{min}$ is always checked ex post during the structural estimation, and whenever the condition does not hold, the parameter combination is discarded.

Positive demand for the homogenous good in all locations is not sufficient. It must also be the case that in equilibrium:

$$h_m L_m - \sum_{k\in K} \int_{N_{mk}} \left(l_{mkn} + f\right) dn > 0 \tag{1.73}$$

in terms of efficiency units of labor. Labor income is not endogenously modelled but it affects the spatial equilibrium. We therefore estimate the model for values of h_m that reproduce the following feature of the data, namely a labor share employed in the non-tradable (differentiated) sector that is invariant to city size (.28 in the data). This is labelled as $\text{share}_m^{Diff.} = \text{share}^{Diff.}$ and set h_m as follows:

$$\begin{aligned}
L_m \text{share}^{Diff.} &= \frac{1}{h_m} \sum_k N_{mk} \left\{ \frac{L_m}{2\gamma} \left[\frac{(\alpha - \mu_{mk})\, \mu_{mk}}{1 + \delta N_{mk}} - \sigma_k^2 \right] + f \right\} \\
h_m &= \frac{1}{\text{share}^{Diff.}} \sum_k N_{mk} \left\{ \frac{1}{2\gamma} \left[\frac{(\alpha - \mu_{mk})\, \mu_{mk}}{1 + \delta N_{mk}} - \sigma_k^2 \right] + \frac{f}{L_m} \right\}
\end{aligned} \tag{1.74}$$

where the first term into squared brackets corresponds to expected firm size in market m and industry k.

We next provide the conditions such that h_m, on which the utility of locating in city m as well as the characterization of the spatial equilibrium hinges, is increasing in market size. We analyze distinctly the two components of h_m, total labor demand for production (the first term in curly brackets) and total labor demand employed at the entry stage.

1. Total labor demand in differentiated goods sector:

$$\begin{aligned}
E(l_{mk}) = N_{mk} E(l_{mkn}) &= \frac{L_m}{\eta} \left((\alpha - \mu_{mk})\lambda_{mk} - 1\right) \left(\frac{\mu_{mk}}{\lambda_{mk}} - \sigma_k^2 \right) \\
\Leftrightarrow E(l_{mk})/L_m &= \frac{1}{\eta} \left(\left(\mu_{mk} - \lambda_{mk}\sigma_k^2\right) \left((\alpha - \mu_{mk}) - \frac{1}{\lambda_{mk}} \right) \right) \\
\Leftrightarrow \frac{\partial E(l_{mk})/L_m}{\partial L_m} &= \frac{1}{\eta} \left(C_{mk} + \frac{1}{2}\lambda_{mk}^3 F_m \sigma_k^2 \right) \left((\alpha - \mu_{mk}) - \frac{1}{\lambda_{mk}} \right) + \\
&\quad + \frac{1}{\eta} \left(\mu_{mk} - \lambda_{mk}\sigma_k^2\right) \left(-C_{mk} - \frac{1}{2}\lambda_{mk} F_m \right) \\
\Leftrightarrow \frac{\partial E(l_{mk})/L_m}{\partial L_m} &= \frac{1}{\eta} C_{mk} \left((\alpha - 2\mu_{mk}) - \frac{1}{\lambda_{mk}} \left(1 - \lambda_{mk}^2 \sigma_k^2\right) \right) + \\
&\quad + \frac{1}{2\eta} \lambda_{mk} F_m \left(\lambda_{mk}^2 \sigma_k^2 (\alpha - 2\mu_{mk}) - \mu_{mk} \left(1 - \lambda_{mk}^2 \sigma_k^2\right)\right) > 0
\end{aligned} \tag{1.75}$$

with $1 - \lambda_{mk}^2 \sigma_k^2 > 0 \Leftrightarrow 4\gamma f_m > 2\sigma_k^2$ and $\frac{\mu_{mk}}{\alpha - \mu_{mk}} - \lambda_{mk}^2 \sigma_k^2 > 0 \Leftrightarrow 2\mu_{mk} > \alpha$. The last condition ensures that, for given level of competition, firm demand per capita is increasing in city/industry efficiency. In equilibrium it has been shown that more efficiency city/industries are always competitive enough

for individual firms to receive relatively lower demand, so that firm size is always unambiguously smaller. However, conditional on given level of competition firms operating in city/industries characterized by higher efficiency level can be smaller or larger, depending on whether they receive sufficiently larger demand to compensate for the fact that, being more productive, they need less labor to produce given output. The restriction $2\mu_{mk} > \alpha$ ensures that the first channel dominates.
Finally the fixed cost has to be sufficiently high $4\gamma f_m > 2\sigma_k^2$ for firms in large cities to be better shielded from overwhelming competition that would otherwise follow. Notice that this restriction may be binding as well as it may not be, provided that $4\gamma f_m > \sigma_k^2 + \left(\frac{\sigma_k^2}{\mu_{mk}}\right)^2$.

2. Total labor demand employed in the entry stage:

$$\frac{\partial F_{mk}/L_m}{\partial L_m} = \frac{1}{2\eta}\left\{-C_{mk}\lambda_{mk}\left(4\gamma f_m\right) + \left[(\alpha-\mu_{mk})\lambda_{mk}\left(1-\frac{\lambda_{mk}^2}{2}\left(4\gamma f_m\right)\right)-1\right]\right\} > 0$$

with F_{mk} denoting total entry cost in industry k and city m, and:

$$(\alpha-\mu_{mk})\lambda_{mk}\left(1-\frac{\lambda_{mk}^2}{2}\left(4\gamma f_m\right)\right)-1 < 0 \Leftrightarrow -2\gamma f_m\lambda_{mk}^3+\lambda_{mk}-(\alpha-\mu_{mk})^{-1} < 0$$

Hence, $\frac{\partial h_m}{\partial L_m} > 0$ if $\alpha < 2\mu_{mk}$.

Chapter 2

THE MICRO-ORIGINS OF BUSINESS CYCLES: A SPATIAL APPROACH

joint with Heiko Stüber

2.1 Introduction

When Amazon.com announced the opening of a second headquarter, 238 U.S. cities signed up for becoming second home to the second most valuable U.S.-listed company. The presence of a large company greatly benefits local employment and productivity, both directly and indirectly through spillovers on neighboring firms. Nevertheless, attracting a large firm such as Amazon is not without drawbacks. For instance, a textbook disadvantage of a granular economy is the strengthened market power of the largest firms and therefore higher inefficiency of the economy as a whole.[1]

Higher market power is not, however, the only disadvantage. A strand of literature (Gabaix (2011), di Giovanni and Levchenko (2012) and Carvalho and Grassi (2017)) has also shown theoretically that the concentration of economic activity in the hands of few firms causes greater volatility. This aspect of granular economies is also of interest to macroeconomists arguing how higher uncertainty is detrimental for aggregate investment (Bloom (2009)). In this paper, we assess empirically to what extent granularity is responsible for macroeconomic volatility. In particular, we test three empirical implications of Carvalho and Grassi's

[1]The combination of granularity and oligopolistic competition has been for instance explored in Gaubert and Itskhoki (2018) and Grassi (2018).

(2017) modeling framework: 1) the more fat-tailed the steady state firm size distribution, the more persistent macroeconomic aggregates conditional on firm-level shocks persistence; 2) the more fat-tailed the steady state firm size distribution, the higher aggregate unconditional volatility; 3) an increase (decrease) in concentration Granger-causes an increase (decrease) in aggregate conditional volatility. Furthermore, we provide narrative evidence supporting the theory on the micro-origin of business cycles.

The intuition for these theoretical results is as follows. When a large firm receives a negative (positive) and persistent shock, even if other firms meanwhile benefit (suffer) from more (less) favorable idiosyncratic conditions, their impact on aggregate employment is not sufficient to counteract the effect of the shock to the large firm: as a result, a single firm is capable of plunging the economy into a recession (triggering an expansion). Hence, higher concentration on average leads to higher unconditional volatility: furthermore, since concentration varies over time, an increase in concentration leads a rise in aggregate conditional volatility. Moreover, since shocks to individual firms are assumed to be persistent, the wage and aggregate employment stay depressed (experience a boom) for longer.

To test the empirical implications in Carvalho and Grassi (2017) we exploit spatial variation in local employment business cycles across Metropolitan Areas in Germany and time variation in concentration within each Metropolitan Area. Metropolitan Areas represent local labor markets and therefore lend themselves well as units of analysis, so long as risk-sharing at the national level is not perfect. The notion of Metropolitan Areas is not introduced just for methodological purposes: we expect the impact of granularity on aggregate volatility to be especially visible at the local level, where the law of large numbers is more likely to fail. While the idea that the destiny of cities and regions is often determined by the economic success of the industries to which they are tied has received great attention both among academics and policy-makers.[2] A more rigorous empirical assessment of the extent to which individual firms are actually responsible for local slumps and expansions is, however, missing, which further motivates our analysis.

To test all three predictions in Carvalho and Grassi (2017) we require information at a business cycle frequency of macroeconomic aggregates and the moments of the firm size distribution. The second set of variables is especially difficult to gather since they build on data on the population of firms (or establishments). The measure of concentration derived from a sample of firms - as typically available in many datasets - can, in fact, turn out to be a very poor approximation of actual

[2]Famous examples are the demise of the UK textile industry or the US car industry and the ensuing surge in unemployment and negative growth experienced in the North of England or the Midwest in the US. In a more similar flavor to the one in this paper, Simon (1988) discusses the impact of industrial diversity on local labor market volatility, unemployment and wages.

concentration: by the definition of fat-tailed distribution, an increasingly larger sample is needed to be able to observe the largest firms. Furthermore, the information contained in most firm-level datasets is at the yearly frequency, which is a too low frequency to study the effects of concentration on the occurrence and shape of local business cycles. The database we rely on - the Administrative Wage and Labor Market Flow panel (AWFP) from the Institute of Employment Research in Germany (IAB) - allows us to circumvent both issues related to data availability since it contains information on employment for each German establishment at the quarterly frequency.

Local concentration is measured in terms of the fraction of local employment accounted by establishments situated above the 99^{th} of the local establishment size distribution. We extract the cyclical component of local employment and proxy its persistence with the estimated coefficient on the first lag of an autoregressive process. We find that an increase in local concentration by one standard deviation is associated with an increase in persistence by .23. We measure unconditional volatility of the cyclical component of local employment, and find that an increase in local concentration by one standard deviation is associated with an increase in volatility by 52 basis points. We follow McConnell and Perez-Quiros (2000) in constructing an estimate of instantaneous residual volatility and find a positive and statistically significant correlation with one-period-ahead concentration. Interestingly, we find evidence of a statistically significant association with previous lags. Finally, we also report a negative and statistically significant association with contemporaneous concentration. These findings are compatible with the micro-origin of recessions *only*. When concentration is high, the likelihood of a further increase (positive shock to a large firm) is lower than a decline (negative shock to a large firm): if it was identical, a period characterized by higher residual volatility and preceded by higher-than-usual concentration should be no more likely to feature low than high concentration.

Next, we cast these results in the terminology of the business cycle literature: specifically, we investigate whether a) concentration is associated with steeper recessions/recoveries, b) it predicts the occurrence of turning points. We calculate turning points for each Metropolitan Area and find that cities characterized by higher average concentration tend to feature recessions that are steeper on average. Furthermore, local recessions characterized by higher-than-usual concentration up to the peak (start of recession) tend to be steeper, while recoveries accompanied by higher-than-usual concentration tend to be less steep. The timing of the uncovered statistical correlation provides support in favor of the causality link between concentration and steepness of recessions. Finally, we find that a small positive deviation of concentration from steady state at peak (trough) increases (reduces) the probability of a peak (trough) occurring by 1.5 times as much as the change in concentration. We take this evidence that concentration Granger-causes recessions but not expansions, thus rejecting the symmetry of the theory in Carvalho and Grassi (2017). We suspect that the empirical rejection of Gibrat's Law - that their results build upon - stands as a potential explanation for the disagreement between the theory and the data, but we leave a more accurate investigation to future research.

Next, we provide narrative evidence in favor of the micro-origins of local business cycles. We combine AFWP data with stock price information on 15 industrial goods companies that are part of the DAX index and describe a list of episodes in which idiosyncratic firm conditions have been the likely driver of local employment business cycles, ranging from the takeover battle between Procter & Gamble and Beiersdorf AG and the protracted slump in Hamburg aggregate employment in 2002/2003, to the long series of strikes organized by IG Metall during the same period and the difficulties experienced by the BMW supply chain and aggregate employment in Landshut and Schwandorf, where two major BMW plants are located.

Finally, we discuss the role of economic policy. In many countries large firms are the subject of multiple regulations: this special attention is a telling signal of the important role played by large corporations for economic development. While it is extremely likely that the efficiency cost of such regulations for the overall economy (Garicano et al. (2016)), the analysis presented in this paper suggests that a strictly positive amount of asymmetric regulation might be socially optimal. The economy described in this paper is characterized by firms that expand or shrink out of luck without internalizing the impact of their firing/hiring decisions on local labor market conditions: it follows that when a large firm receives a negative shock, it fires too many workers, while when it receives a positive shock, it hires too many. A social planner should therefore tax large firms receiving a positive shock, and subsidize large ones being hit by a negative shock. Furthermore, a more than proportional intervention in the former case would cater to the macro-prudential objective of "leaning" against the formation of too large corporations, thus reducing the scope for "cleaning" ex post. Given the vast literature on the *costs* (e.g., economic, bureaucratic, etc.) imposed on large firms, we provide a digression on size-contingent policies in place across countries aimed at providing relief to companies experiencing difficulties.[3]

The rest of the paper is organized as follows: Section 2 describes the conceptual framework; Section 3 presents the empirical evidence; Section 4 provides the narrative evidence; Section 5 contains a survey of policies in place in a group of major European economies that address - often partially - the negative externality imposed by large firms on the rest of the economy in terms of endogenous

[3]These transfers are ofter financed by extra-social security payments to which large businesses are subject.

volatility; Section 6 concludes.

2.2 Conceptual Framework

In this section, we briefly sketch the setup in Carvalho and Grassi (2017) (henceforth CG 2017) and their main theoretical findings that are related to the analytical characterization that the authors provide of the dynamics of aggregate productivity and output in the presence of granularity.[4]

For simplicity, we consider a partial equilibrium in which the size of a given city is exogenously determined and employment fluctuates according to the quantity of leisure that households choose to consume. Consider a representative city. The aggregate state at time t for the local economy corresponds to the productivity distribution $\mu_t = (\mu_{1,t}, \mu_{2,t}, ..., \mu_{s,t}, ..., \mu_{S,t})'$, where $\mu_{s,t}$ denotes the number of firms at time t characterized by productivity φ^s. The productivity space is a S-tuple $\Phi = \left\{\varphi^1, ..., \varphi^S\right\}$, where $\varphi > 1$ so that $\varphi^1 < ... < \varphi^S$. There is a total of N firms with $N \in \mathbb{N}$: each firm's productivity is assumed to follow a Markov chain with transition matrix P given by:

$$P = \begin{bmatrix} a+b & c & 0 & ... & ... & 0 & 0 \\ a & b & c & ... & ... & 0 & 0 \\ ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & ... & ... & a & b & c \\ 0 & 0 & ... & ... & 0 & a & b+c \end{bmatrix} \tag{2.1}$$

Hence, the probability of a firm's productivity to improve is c, to decline is a, to stay the same is $b = 1 - a - c$. Cordoba (2008) proves that the Markovian process described in eq.2.1 leads to a Pareto distribution, thus generalizing the continuous state-space result. For a given firm i characterized by productivity level $\varphi^{s_{i,t}}$, as $t \to \infty$ the probability of having productivity level φ^s is:

$$\lim_{t\to\infty} P(\varphi^{s_{i,t}} = \varphi^s) = K(\varphi^s)^{-\delta} \tag{2.2}$$

where K is a normalization constant and $\delta = \frac{log(a/c)}{log\varphi}$ is the tail index of the Pareto distribution to which the Markovian process described in eq.2.1 converges. As the probability of receiving a favorable shock c rises relative to the probability of receiving an unfavorable shock a the tail index declines and the firm size distribution becomes more fat-tailed.

[4] Since in the empirical analysis section we refrain from distinguishing between the intensive (growth of continuing establishments) and extensive (entry/exit of establishments) margin of economic activity, we focus on the set of results in CG 2017 that are based on the absence of entry/exit.

Aggregate productivity is obtained by summing across firms productivities and it depends on the productivity distribution through:

$$A_t = \sum_{s=1}^{S} (\varphi^s)^{\frac{1}{1-\alpha}} \mu_{t,s} \tag{2.3}$$

One of the contributions of CG 2017 is to show that aggregate productivity follows the law of motion:

$$A_{t+1} = \rho A_t + O_t^A + \sigma_t \varepsilon_{t+1} \tag{2.4}$$

with:

$$\rho = a\varphi^{-\frac{1}{1-\alpha}} + b + c\varphi^{\frac{1}{1-\alpha}} \qquad \sigma_t^2 = \varrho D_t + O_t^\sigma \tag{2.5}$$

where $D_t = \sum_{s=1}^{S} \left((\varphi^s)^{\frac{1}{1-\alpha}} \right)^2 \mu_{s,t}$ is proportional to the second moment of the firm size distribution at time t. Hence, volatility of the aggregate productivity shock $\tilde{\varepsilon}_{t+1} = \sigma_t \varepsilon_{t+1}$ is time-varying and proportional to the dispersion in the firm size distribution, D_t.[5]

In order to derive a law of motion for aggregate employment as a function of the law of motion for aggregate productivity, the economic environment must be specified. CG 2017 assume that firms produce a homogenous good with price normalized to 1 using only labor and according to a decreasing returns to scale technology as in Hopenhayn (1992). This gives rise to labor demand $L_t^d = \left(\frac{\alpha}{w_t}\right)^{\frac{1}{1-\alpha}} A_t$. The model is closed by a ad-hoc supply schedule $L_t^s = M w_t^\gamma$, with M being a scaling factor.

The equilibrium wage and employment are, respectively: $w_t = \left(\alpha^{\frac{1}{1-\alpha}} A_t\right)^{\frac{1-\alpha}{\gamma(1-\alpha)+1}} M^{\frac{\alpha-1}{\gamma(1-\alpha)+1}}$ and $L_t = \left(\alpha^{\frac{1}{1-\alpha}} A_t\right)^{\frac{\gamma(1-\alpha)}{\gamma(1-\alpha)+1}} M^{\frac{1}{\gamma(1-\alpha)+1}}$.

Defining with $\widehat{X}_t$ the percentage deviation of variable X from its steady state, the law of motion for aggregate employment is:

[5] O_t^A and O_t^σ are two correction terms that vanish as the bounds of the productivity space increase, $E(\varepsilon_{t+1}) = 0$ and $Var(\varepsilon_{t+1}) = 1$. Omitting the terms related to the productivity distribution of the entrants, the two correction terms in CG 2017 (Online Appendix) are:

$$O_t^A = \left(\varphi^{\frac{1}{1-\alpha}}\right)^S (\rho_S - \rho)\mu_{S,t} \qquad O_t^\sigma = \left(\varphi^{\frac{1}{1-\alpha}}\right)^{2S} (\sqrt{\varrho_S} - \sqrt{\varrho})^2 \mu_{S,t}$$

where:

$$\rho_S = \begin{bmatrix} \varphi^{-\frac{1}{1-\alpha}} & 1 \end{bmatrix} \begin{bmatrix} a \\ b+c \end{bmatrix} \qquad \varrho_S = \begin{bmatrix} \varphi^{-\frac{1}{1-\alpha}} & 1 \end{bmatrix} \begin{bmatrix} a(1-a) & -a(1-a) \\ -a(1-a) & a(1-a) \end{bmatrix} \begin{bmatrix} \varphi^{-\frac{1}{1-\alpha}} \\ 1 \end{bmatrix}$$

$$\widehat{L}_{t+1} = \rho\widehat{L}_t + \kappa\widehat{O}_t^A + \psi\frac{\sigma_t}{A}\varepsilon_{t+1} \tag{2.6}$$

where $\widehat{L}_t = \psi\widehat{A}_t$, with $\psi = 1 - \frac{1}{\gamma(1-\alpha)+1}$, $\kappa = \psi\frac{O^A}{A}$ and A equal to the steady state productivity level.

Proposition 3. ***Carvalho and Grassi (2017).*** *If $\delta(1-\alpha) \geq 1$, then the persistence of the aggregate employment, ρ, satisfies the following properties:*

a. Holding δ constant, aggregate persistence is increasing in firm-level persistence b: $\frac{\partial\rho}{\partial b} \geq 0$,

b. Holding b constant, aggregate persistence is decreasing in the tail index of the stationary productivity distribution: $\frac{\partial\rho}{\partial\delta} \leq 0$,

c. If the productivity distribution is Zipf, aggregate productivity dynamics contain a unit root: if $\delta = 1/(1-\alpha)$, $\rho = 1$.

Proposition 1 (Proposition 3, CG 2017) states that the dynamics of aggregate employment in an economy characterized by a very skewed firm size distribution should feature a higher degree of persistence, holding firm-level persistence constant (b). Through the lenses of the theory, an increase in the degree of skewness holding persistence constant is achieved through an increase in the probability for an individual firm to grow matched by a tantamount decrease in the probability of shrinking. Additionally, it states that if firm-level persistence increases, higher persistence of aggregate employment follows (a).[6]

Furthermore, *conditional* (on the previous period) volatility of aggregate employment is:

$$Var(\widehat{L}_{t+1}) = \psi\frac{\sigma_t^2}{A^2} = \psi\left(\varrho\frac{D}{A^2}\frac{D_t}{D} + \frac{O^\sigma}{A^2}\frac{O_t^\sigma}{O^\sigma}\right) \tag{2.7}$$

Proposition 4. ***Carvalho and Grassi (2017).*** *Then:*

a If $1 < \delta(1-\alpha) < 2$, the unconditional expectation of the variance of aggregate employment is satisfies:

$$\mathbb{E}\left(\frac{\sigma_t^2}{A^2}\right) \sim_{N\to\infty} \frac{\varrho G}{N^{2\left(1-\frac{1}{\delta(1-\alpha)}\right)}}$$

where G is a function of other parameters and independent of N,

[6] Carvalho and Grassi (2017) derived all their results in terms of aggregate output. Due to output isoelasticity in employment, their findings extend straightforwardly to aggregate employment.

b The dynamics of conditional variance of aggregate employment depend on the dispersion of firm size, $\frac{\partial Var(\widehat{L}_{t+1})}{\partial D_t} > 0$.

Proposition 2 (Proposition 4, CG 2017) states that unconditional variance is an increasing function of the degree of skewness of the firm size distribution (a) and that the conditional variance is an increasing function of the degree of dispersion (b). In the rest of the paper, we test Proposition 1b, Proposition 2a and 2b.

2.3 Empirical Analysis

2.3.1 Data

The analysis is based on administrative records from the Administrative Wage and Labor Market Flow Panel (AWFP), a dataset recording information on major labor market outcomes (employment and wages) for the universe of German establishments (*Betriebe*). A detailed description of the data can be found in Seth and Stüber (2017).

We draw from the AWFP a sample corresponding to the universe of establishments reporting at least one full-time worker receiving social security contributions over the period from 1990q1 to 2014q4.[7] All sectors of economic activity are included except from Agriculture and Mining and Construction, and retain exclusively establishments located in *Kreise* mapping into local labor markets of former West Germany.[8] This leaves us with a total of 72 Metropolitan Areas (or “Arbeitsmartkregion”, sometimes abbreviated as AMR from here onwards), as shown in Fig.2.1. Rural areas are thus excluded from the analysis.

We aggregate establishment-level employment at the 3-digit industry and *Kreis* level. *Kreise* are the German equivalent of US counties: to reflect differences in seasonality induced by the type of economic activity or legislation enforced in different *Kreise*, we seasonally adjust employment data at the 3-digit industry and *Kreis*-level.[9] Next, we compute AMR private sector full-time employment based on Eurostat geographical delineations.

[7] As standard when working with establishment-level data distributed by the Institute for Labor Market Research in Germany (IAB), we miss a few categories of workers, such as self-employed and civil servants. Additionally, we exclude part-time workers.

[8] We exclude the Berlin Metropolitan Area from the sample.

[9] We follow Chodorow-Reich and Wieland (2018) and take advantage of the seasonal adjustment code kindly made available by the authors at https://scholar.harvard.edu/chodorow-reich/data-programs.

2.3.2 Methodology

The goal of the empirical section is to test three important theoretical results of CG 2017:

1. higher steady state concentration is associated with higher persistence of employment (Prediction 1b);

2. higher steady state concentration is associated with higher unconditional volatility of employment (Prediction 2a);

3. an increase (decrease) in concentration leads an increase (decrease) in conditional volatility (Prediction 2b).

We exploit variation in employment at the local labor market level to test these predictions.

After having analyzed the relationship between the moments of the local employment time series and the shape of the city-specific establishment size distribution, we establish a bridge between our results and the business cycle literature. Specifically, we ask ourselves: what are the implications of higher persistence and unconditional volatility for the intensity of recessions and recoveries? The intensity of a recession (recovery) is defined as the per quarter average loss (gain) in employment (or output, investment, etc.). The intensity is therefore a meaningful proxy for the severity of recessions/extent of recoveries.

Then, while on the one hand higher persistence should translate into *longer* business cycle phases, both recessions and recoveries alike, higher volatility would on the other seem to counteract this effect, by inducing both *shorter* and/or *deeper* business cycles (i.e., that have higher amplitude). It follows that local labor markets characterized by higher concentration should unambiguously experience *steeper* (more intense) recessions/recoveries, although not necessarily longer ones.

Furthermore, one advantage of running the analysis at the local business cycle level is that we are able to tell apart endogenous volatility associated with aggregate vs. recessions that are local in nature, which represent the proper testing ground of the theory.

Before moving to the details of the empirical strategy, we describe the tools borrowed from the business cycle literature upon which we rely extensively.

Business cycle dating The bulk of the literature on business cycle dating draws inspiration from the seminal work by Burns and Mitchell (1946). Their empirical definition of business cycle goes as follows: *"business cycles are a type of fluctuation found in the aggregate economic activity of nations that organize their work mainly in business enterprises: a cycle consists of expansions occurring at about* *the same time in many economic activities, followed by similarly general recessions, contractions, and revivals which merge into the expansion phase of the next cycle.”*

Hence, there are two peculiar features of business cycles: 1) the comovement of many individual economic series, and 2) the different behavior of the economy during expansions and contractions (Diebold and Rudebusch (1996)). In this paper, we focus on the second feature of business cycle fluctuations. We identify local and aggregate business cycles by means of the widely exploited Harding and Pagan (2002) business cycle dating algorithm. The algorithm performs three tasks:

1. Determination of a potential set of turning points i.e. the peaks and troughs in a series.

2. A procedure for ensuring that peaks and troughs alternate.

3. A set of rules that re-combine the turning points established after steps one and two in order to satisfy pre-determined criteria concerning the duration and amplitudes of phases and complete cycles, the so-called “censoring rules”.

We apply the same censoring rules as in the original paper by Harding and Pagan (2002): we impose that a cycle (i.e, from peak-to-peak) lasts at least 5 quarters, and that a phase (either from peak-to-trough or from trough-to-peak) lasts at least 2 quarters.

In the existing literature no business cycle dating exercise has been so far conducted at the local level that can be referred to as benchmarks: instead, we assess the goodness of the algorithm based on the turning points that the procedure identifies at the national level and the comparison with the turning points identified by the OECD.[10]

The first dating exercise is based on (log) aggregate private sector full-time employment. The series is plotted in Fig.2.2 together with the start and end of recession dates identified according to the two procedures (OECD - left, Harding and Pagan (2002) - right). There are a number of potential explanations behind the discrepancy observed in terms of the two sets of turning points. A first possible reason could be that the analysis in this paper is based on former West Germany

[10]Data on OECD-based turning points can be found at https://fred.stlouisfed.org/series/DEURECM. Turning points are identified by the OECD at the monthly frequency: this raises a conflict with the dating procedure employed in this paper, which exploits data at the quarterly frequency. To reconcile the different frequencies, we let a recession start in the same quarter where the month identified by the OECD dating procedure falls (and similarly for the trough).

only, while the OECD considers Germany as a whole. In support of this candidate explanation, one can see that the OECD dating procedure identifies a few more national cycles at the beginning of the sample, during which time it is conceivable that the integration between former West and East Germany gave rise to higher aggregate volatility.

A second potential explanation has to do with the economic indicator used for business cycle dating: the OECD relies on industrial production, which provides timely information on economic activity at the monthly frequency and for a large set of countries. while this analysis is based on private sector employment.

We repeat the dating exercise and this time use the cyclical component of (log) aggregate employment, which is identified according to either linear detrending (and break adjustment) or HP filtering.[11] The cyclical component is plotted in Fig.2.3 together with the start and end of recession dates identified by the business cycle dating algorithm on either of the two series: the series obtained through HP filtering reveals a higher number of turning points.

While we are only marginally interested in low frequency fluctuations in employment, one can see from Fig.2.2 that since the reunification full-time aggregate employment has declined substantially in Germany. The mirror image of this picture is the surge in part-time employment: part-time employment has increased by 108% from January 2000 to January 2014: while it represented 11.9% of the stock of employees subject to social security contributions in 2000, this figure stands at 32.6% at the end of our sample in 2014.[12]

A business cycle is defined as the period ranging from peak to next peak: within a business cycle we distinguish two phases, a recession (lasting from peak to next trough) and a recovery (lasting from trough to next peak). We can characterize the shape of business cycles through a set of properties. Index by i the set of all local recessions identified by the dating procedure, and by m the Metropolitan Area. Furthermore, define $C^i_{m,t} \in \{1, ..., I\}$ where I is the set of all local recessions: $C^i_{m,t}$ therefore denotes the recession the m^{th} economy is going through at time t. Similarly, index by j the set of all local recoveries identified by the dating procedure, and by m the Metropolitan Area. Furthermore, define $C^j_{m,t} \in \{1, ..., J\}$ where J is the set of all local recoveries: $C^j_{m,t}$ therefore denotes the recovery the m^{th} economy is going through at time t.

The business cycle properties employed in the analysis are listed below:

- Depth recession$_{m,i} = 100 \times |y_{m,t_2} - y_{m,t_1}|$, where $t_2 = \max\left\{t | C^i_{m,t} = i\right\} + 1$

[11]We linearly detrend aggregate employment by running $y_t = \alpha + \beta t + \gamma I_{t|t \geq 2005q1} + \delta t \times I_{t|t \geq 2005q1} + \epsilon_t$ and take $\hat{y}_t = \hat{\epsilon}_t$, thus allowing for a break in correspondence with the Hartz reforms.

[12]Data on part-time employment are available at the monthly frequency on the Bundesagentur für Arbeit website at https://statistik.arbeitsagentur.de/.

and $t_1 = \min\left\{t|C^i_{m,t} = i\right\}$, with t_2 and t_1 being, respectively, the period immediately after the end (trough) and the start (peak) of the current recession.

- Depth recovery$_{m,j} = 100 \times (y_{m,t_2} - y_{m,t_1})$, where $t_2 = \max\left\{t|C^j_{m,t} = j\right\} + 1$ and $t_1 = \min\left\{t|C^j_{m,t} = j\right\}$, with t_2 and t_1 being, respectively, the period immediately after the end (peak) and the start (trough) of the current recovery.
- Length recession$_{m,i} = t_2 - t_1$, where $t_2 = \max\left\{t|C^i_{m,t} = i\right\} + 1$ and $t_1 = \min\left\{t|C^i_{m,t} = i\right\}$.
- Length recovery$_{m,j} = t_2 - t_1$, where $t_2 = \max\left\{t|C^j_{m,t} = j\right\} + 1$ and $t_1 = \min\left\{t|C^j_{m,t} = j\right\}$.
- Intensity recession$_{m,i}$ = Depth recession$_{m,i}$/Length recession$_{m,i}$.
- Intensity recovery$_{m,j}$ = Depth recovery$_{m,j}$/Length recovery$_{m,j}$.

The *depth* of a business cycle phase captures the cumulated increase or decrease in employment from peak to trough (recession) or from trough to peak (recovery), while the *intensity* of a business cycle phase measures the speed at which employment plunges or recovers.

Finally, we describe the measure of local concentration adopted for this analysis. There are many candidate measurements of concentration in local economic activity. The main difference is the extent to which they account for the entire shape of the establishment size distribution. The Herfindahl index, for instance, considers dispersion in the size distribution as a whole, while the fraction of local employment accounted by establishments situated above the x^{th} percentile of the local establishment size distribution - with x being a large percentile - places more weight on the size of the largest establishments. A third alternative to measure concentration consists of the shape parameter estimated by fitting a Pareto distribution on the empirical density of establishment sizes: this option is the least appealing since it imposes a too strong parametrization that is effectively rejected in the data.[13]

Since the focus in this analysis is on large establishments, we set as benchmark measure of concentration the fraction of local employment employed by establishments situated above the 99^{th} percentile, and label it $x_{m,t}$.[14] Existing literature has documented how establishment size is complementary with the size of

[13]The firm size distribution is Pareto but only in the right tail. See Fernandes et al. (2017) for a discussion in the context of international trade.

[14]The findings do not depend on this choice and they carry through also with other top-percentiles, such as 90 or 95.

the local labor market (Combes et al. (2012) and Gaubert (2018)). We document this empirical regularity also for German cities in Table 2.1: a 1% increase in city size translates into a 3.5 percentage point increase in concentration as defined above. The magnitude of the coefficient increases moving from the 90^{th} to the 99^{th} percentile.

We seasonally adjust local concentration at the Metropolitan Area level: based on the seasonally adjusted series, we find evidence of a pronounced downward trend in establishment size dispersion across all Metropolitan Areas in the sample: in Fig.2.5 we plot the median, the 10^{th} and 90^{th} percentile of the cross-sectional distribution in the seasonally adjusted index of local concentration at any point in time during 1990q1-2014q4. This finding is in line with Moral-Benito and Queirós (2018) who also find that concentration has gone down in the population of Spanish firms. Hence, we extract the cyclical component of local concentration according to the same procedure (either linear detrending or HP filtering) employed to calculate cyclical employment and local turning points.

Empirical Strategy Estimation of the persistence and unconditional volatility of employment in Germany must account for the fact that the German labor market underwent a sequence of structural reforms in the early 2000s that reached full implementation in January 2005. We therefore adjust the local concentration time series for the 72 local labor markets in our sample by estimating eq.2.8 at the Metropolitan Area m level:

$$x_{m,t} = \alpha_{1,m} + \alpha_{2,m} I_{t|t\geq 2005q1} + \beta_{1,m} t + \beta_{2,m} t \times I_{t|t\geq 2005q1} + \nu_{m,t} \quad m \in \{1, 2, ..., M\} \tag{2.8}$$

where $x_{m,t}$ is seasonally adjusted local concentration. The city-specific coefficients on the break in the intercept and linear time trend are statistically significant for the majority of Metropolitan Areas. We define the cyclical component of local concentration, $x^C_{m,t} = \widehat{\nu}_{m,t}$.[15]

Since unconditional volatility and persistence are shaped by average concentration over the time span considered, we allow our measure of average concentration to reflect potential breaks in the intercept in 2005q1: specifically, we set average concentration equal to a weighted average of the estimated intercepts on the two subsamples, $\overline{x}_m = .6 \times \widehat{\alpha}_{1,m} + .4 \times (\widehat{\alpha}_{1,m} + \widehat{\alpha}_{2,m})$, where .6 and .4 are weights reflecting the length of each subsample over the total.

The set of hypotheses is tested through the following statistical model:[16]

[15]We also test the robustness of our results to the HP filter-based trend/cycle decomposition for local concentration.

[16]An alternative would be to take a Bayesian approach, jointly estimating the system in eq.2.9-eq.2.10 (e.g., Cogley and Sargent (2005)).

$$y_{m,t} = \gamma_{1,m} + \gamma_{2,m} I_{t|t\geq 2005q1} + \delta_{1,m} t + \delta_{2,m} t \times I_{t|t\geq 2005q1} + \rho_m y_{m,t-1} + \epsilon_{m,t} \quad m \in \{1, 2, ..., M\} \tag{2.9}$$

$$\sqrt{\frac{\pi}{2}} |\widehat{\epsilon}_{m,t}| = \kappa_m + \varphi x^C_{m,t-1} + \kappa_t + \eta_{m,t} \tag{2.10}$$

where $y_{m,t}$ is seasonally adjusted log employment, and $I_{t|t\geq 2005q1}$ is an indicator taking value 1 after 2005q1 and 0 otherwise. Hence, we allow for city-specific intercepts and trends on each subsample, which turn out to be statistically significant for the majority of Metropolitan Areas.[17]

Eq.2.10 is estimated on the pooled sample. The dependent variable, $\sqrt{\frac{\pi}{2}}|\widehat{\epsilon}_{m,t}|$, is an unbiased estimator for the instantaneous volatility of the innovation to an autoregressive process (Davidian and Carroll (1987)): this definition of instantaneous volatility has been used by McConnell and Perez-Quiros (2000), Stock and Watson (2003), Carvalho and Gabaix (2013) in different contexts. The addition of year/quarter fixed effects, κ_t, controls for changes in macroeconomic volatility, while city fixed effects, κ_m, allow for cross-city differences in average instantaneous volatility. As a robustness check to eq.2.9, we include contemporaneous concentration or the lagged value of the dependent variable:

$$\sqrt{\frac{\pi}{2}} |\widehat{\epsilon}_{m,t}| = \kappa_m + \psi \sqrt{\frac{\pi}{2}} |\widehat{\epsilon}_{m,t-1}| + \varphi_0 x^C_{m,t} + \varphi_1 x^C_{m,t-1} + \kappa_t + \eta_{m,t} \tag{2.11}$$

We do not incur the small sample bias highlighted in Nickell (1981) when we estimate eq.2.11 given the large longitudinal dimension of our panel. We test the presence of heterogenous effects in φ_1 in eq.2.11 through the following equation estimated at the Metropolitan Area m level:

$$\sqrt{\frac{\pi}{2}} |\widehat{\epsilon}_{m,t}| = \kappa_m + \psi_m \sqrt{\frac{\pi}{2}} |\widehat{\epsilon}_{m,t-1}| + \varphi_{0,m} x^C_{m,t} + \varphi_{1,m} x^C_{m,t-1} + \eta_{m,t} \quad m \in \{1, 2, ..., M\} \tag{2.12}$$

The coefficient on the autoregressive term in eq.2.9, $\widehat{\rho}_m$, measures the persistence in the local employment time series. Unconditional volatility, $\widehat{\sigma}_m$, is calculated based on predicted employment in the absence of breaks or time trends,

[17] It could be that local employment is non-stationary in one or more cities: in this case, eq.2.9 would have to be estimated in first differences. Since the objective of the empirical analysis is to investigate the extent of persistence in local employment, we prefer not to impose a priori a unit root in its data generating process. We acknowledge that the lack of a formal unit root test - that should account for the presence of stochastic and potentially non-stationary volatility - represents a weakness of the current version of the paper, and we are working towards including a more thorough analysis of data stationarity post-detrending.

$\widehat{y}_{m,t} = \widehat{\gamma}_{1,m} + \widehat{\rho}_m \widehat{y}_{m,t-1} + \widehat{\epsilon}_{m,t}$.[18] We test Prediction 1b and Prediction 2a with a simple linear regression of $\widehat{\rho}_m$, $\widehat{\sigma}_m$ on $\overline{x}_m$.[19] Finally, the sign and magnitude of $\widehat{\varphi}_1$ in eq.2.11 delivers a test of Prediction 2b.

Next, we consider the implications of our results for the intensity of recessions/recoveries. These are variables of interest to academics and policy-makers: for example, both contagion to different markets/economies and the probability of a regime switch depend on the severity of recessions.

A more persistent series should be characterized by longer cycles; a more volatile one by both shorter and deeper cycles. In other words, business cycle phases should be *steeper* in Metropolitan Areas characterized by higher steady state concentration of economic activity.

Define the following business cycle *properties*:

$$BC_{m,i} = \left\{ \text{Depth Reces.}_{m,i}, \text{Intens. Reces.}_{m,i}, \text{Length Reces.}_{m,i} \right\}$$

$$BC_{m,j} = \left\{ \text{Depth Recov.}_{m,j}, \text{Intens. Recov.}_{m,j}, \text{Length Recov.}_{m,j} \right\}$$

Local turning points marking the start of recessions/recoveries are calculated through the Harding and Pagan (2002) algorithm on the *cyclical* component of log employment, which is extracted via either linear detrending or HP-filtering. While HP-filtering automatically delivers a trend-cycle decomposition, under linear detrending we set the cyclical component equal to the residual $\widehat{\epsilon}_{m,t}$ in eq.2.9. Results for both detrending procedures are presented.

For each above defined business cycle property we estimate eq.2.13:[20]

$$\begin{aligned} BC_{m,i} &= \alpha + \beta \overline{x}_m + \epsilon_{m,i} \\ BC_{m,j} &= \alpha + \beta \overline{x}_m + \epsilon_{m,j} \end{aligned} \tag{2.13}$$

As robustness check to eq.2.13, we exclude local business cycles that are aggregate in nature, i.e., local business cycles taking place during generalized downturns.

The relationship between concentration and business cycle intensity in eq.2.13 can be further improved upon by exploiting time-variation in the cyclical component of local concentration. When recessions/recoveries have a micro-origin,

[18]Predicted employment is initialized at $\widehat{y}_{m,0} = y_{m,0}$.

[19]Since according to CG 2017 higher concentration should translate into higher aggregate persistence conditional on firm-level persistence, we also provide rudimentary evidence of a lack of association between the two at the Metropolitan Area level.

[20]We employ a Poisson type of regression when the dependent variable is the length of recessions/recoveries.

we expect a buildup in concentration *up to* the turning point (peak or trough) to be associated with steeper business cycle phases (recessions or recoveries). A positive and statistically significant relationship could emerge also *after* the local peak/trough: this evidence, however, would provide weak support in favor of *granularity-driven* recessions/recoveries.

Eq.2.14 tests the sign and magnitude of the relationship between concentration and business cycle steepness at leads and lags:

$$\begin{aligned} BC^{Recession}_{m,i} &= \alpha + \beta x^C_{m,i,k} + \epsilon_{m,i} \qquad \forall k = -4,-3,-2,-1,0,+1,+2,+3,+4 \\ BC^{Recovery}_{m,j} &= \alpha + \beta x^C_{m,j,k} + \epsilon_{m,j} \qquad \forall k = -4,-3,-2,-1,0,+1,+2,+3,+4 \end{aligned} \tag{2.14}$$

where $x^C_{m,i,k}$ corresponds to the cyclical component of local concentration at the start of the i-th local recession ($k = 0$), $|k|$ quarters before the start of the i-th local recession ($k < 0$), $|k|$ quarters after the start of the i-th local recession ($k > 0$). Similarly, $x^C_{m,j,k}$ corresponds to the cyclical component of local concentration at the start of the j-th local recovery ($k = 0$), $|k|$ quarters before the start of the j-th local recovery ($k < 0$), $|k|$ quarters after the start of j-th local recovery ($k > 0$).

As robustness check to eq.2.14, we exclude local business cycles that are aggregate in nature and include city fixed effects. We report results for local business cycles calculated through the Harding and Pagan (2002) algorithm on the cyclical component of log employment identified according to both linear detrending/break adjustment and HP-filtering.[21] The cyclical component of concentration is extracted likewise.

Finally, we test whether higher concentration translates into a higher probability of entering a local recession/recovery, therefore assessing empirically the relevance of concentration as systemic risk factor. We use a probability model to investigate the marginal effect of concentration at lags and leads on the probability of a peak (trough) occurring by means of eq.2.15:

$$\begin{aligned} P(t = \text{peak}_m) &= \frac{1}{1 + \exp(-\alpha - \beta x^C_{m,t+k} - \gamma_t)} \qquad \forall k = -4,-3,-2,-1,0,+1,+2,+3,+4 \\ P(t = \text{trough}_m) &= \frac{1}{1 + \exp(-\alpha - \beta x^C_{m,t+k} - \gamma_t)} \qquad \forall k = -4,-3,-2,-1,0,+1,+2,+3,+4 \end{aligned} \tag{2.15}$$

[21] The procedure for linear detrending/break-adjusting time series $y_{m,t}$ consists of estimating eq.2.9 *without* autoregressive term, such that $y^{trend}_{m,t} = \widehat{\gamma}_{1,m} + \widehat{\gamma}_{2,m} I_{t|t\geq 2005q1} + \widehat{\delta}_{1,m} t + \widehat{\delta}_{2,m} t \times I_{t|t\geq 2005q1}$ and $y^{cycle}_{m,t}$ is equal to the residual.

where $x^C_{m,t+k}$ denotes the cyclical component of local concentration in city m at time $t+k$, and δ_t is a set of time fixed effects.

2.3.3 Results

We start by presenting the empirical assessment of Prediction 1b and 2a. In Fig.2.6 we plot $\widehat{\rho}_m$ and $\widehat{\sigma}_m$ against $\overline{x}_m$: the data reveal a positive and statistically significant relationship between local employment persistence/unconditional volatility and local average concentration, thus validating both predictions. An increase by one standard deviation in local concentration - measured in terms of percentage of local employment accounted by establishments located above the 99^{th} percentile of the local establishment size distribution - is associated with a .23 increase in the AR(1) coefficient, and 52 basis point increase in local volatility. Both relationships are statistically significant at the 5% level.

Since Prediction 1b holds keeping firm-level persistence constant, we investigate whether firm-level persistence is positively associated with the degree of concentration. The finding that Gibrat's Law tends to be violated for small firms (e.g., Hall (1987)) would suggest that firm-level persistence should be on average lower in markets characterized by many large firms, thus hinting towards a negative relationship.

However, when it comes to testing this hypothesis, a few complications arise. On the one hand, the time-varying nature of firm-level persistence during the life-cycle prevents from estimating firm-specific levels of persistence. On the other, a cross-sectional assessment does not allow to control for time-invariant firm characteristics. Additionally, the theory needs to be reconciled with the reality of the data: in CG 2017 firms either grow to the next productivity level, or shrink to the one before, or stay the same, so that persistence is defined as the probability of productivity staying the same. We run a simplistic check and use a linear probability model to estimate based on yearly data the average probability of employment being unchanged at the establishment level in each year and city. In the estimation we include a full set of establishment fixed effects. The year-specific means are then averaged at the city level: the resulting city-specific average probability features a positive but statistically insignificant connection with the degree of concentration.

There are several caveats to such approach. The first one could be that employment is not changing because of the presence of real options or other forms of adjustment costs. Another one could be that we expect small firms to feature mechanically a higher probability of employment being unchanged due to indivisibility of labor. These are very valid reasons for concern: however, since the random growth assumption in CG 2017 - while handy from a practical perspective - is in practice empirically rejected, we refrain from further imposing counterfactual assumptions on the data and postpone a more accurate investigation to future research.

In Table 2.2 we report the coefficients from estimation of eq.2.10. Regression results highlight a consistently positive association between lagged local concentration and instantaneous volatility, thus validating Prediction 2b. Interestingly, instantaneous volatility is statistically significantly associated with lagged concentration only after controlling for contemporaneous concentration. Furthermore, while lagged concentration is positively associated with volatility, contemporaneous concentration features a negative correlation. This evidence strongly corroborates the theory on granularity-driven *recessions*: a buildup in local concentration leads to higher volatility and reverses the following period. This pattern is precisely what we would expect to observe if the recession was triggered by a negative shock to one of the large establishments present in the economy: the swing in aggregate employment the following period is large and concentration drops.

This finding anticipates some of the results coming later in the analysis, i.e., that granularity is not equally likely to Granger-cause recessions and recoveries alike through large establishments being hit by a negative/positive shock, respectively: if this was indeed the case, we would expect the probability of upward swings in concentration (i.e., large establishments to grow even larger) not to depend on the initial concentration level. Instead, the evidence presented in Table 2.2 provides empirical support to the mean-reverting nature of concentration, and to the asymmetric nature of systemic risk associated with a buildup in concentration.

We test whether the correlation between lagged concentration and contemporaneous volatility is heterogeneous at the Metropolitan Area level. We report in Fig.2.7 the estimated coefficients of eq.2.12 on lagged concentration grouped by statistical significance. Despite of considerable dispersion $\widehat{\varphi}_{1,m}$ is statistically significantly estimated only when it takes a positive value, thus lending support Prediction 2b. Overall, we estimate $\widehat{\varphi}_{1,m}$ to be statistically significant in 20% of cases.

Moving to the business cycle analysis, Table 2.3 provides summary statistics on business cycle properties according to both detrending procedures. According to linear detrending (and break-adjustment), both phases tend to last on average longer and consequently entail larger cumulated losses (gains). In terms of intensity, however, both detrending procedures yield very similar outcomes: recessions tend to be more intense than recoveries, in the sense of entailing a larger employment variation per unit of time (in absolute value terms). This evidence squares with the positive skewness characterizing the distribution of changes in the unemployment rate (Barnichon (2012)): positive changes in unemployment at the beginning of recessions are more sizeable in absolute value terms than negative changes when the recovery starts.

Business cycle dating based on linearly detrended/break-adjusted local employment identifies 337 local business cycles, against 447 identified according to HP-filtered data: Fig.2.4 reports the histogram for the number of local business cycles identified according to each detrending procedure at the Metropolitan Area level. The median number of local business cycles during the sample considered is equal to 6 according to the linear detrending procedure, while it is between 6 and 8 according to HP-filtering. The variability in the number of local business cycles is broadly similar across procedures.

Results from the estimation of eq.2.13 are presented in Table 2.4. The table is divided in two panels, each corresponding to a detrending procedure. Within each panel there are two blocks. The first block reports the estimated coefficients and t-statistics for each business cycle property using the entire set of local business cycles as a sample, while the second block excludes local business cycle that are aggregate in nature. A local business cycle is considered to be aggregate in nature if it occurs within 4 quarters from a national peak.[22] Thanks to this robustness, the empirical findings are not biased by differences in the shape of recessions/recoveries between local business cycles with a micro and macro-origin.

Both detrending procedures deliver a positive and statistically significant association between the intensity of local recessions and average local concentration when only purely local business cycles are considered: according to the most conservative estimate, 1 percentage point increase in concentration entails .01 higher percentage decline in employment per quarter of recession.

In Fig.2.8-2.9 we report the estimated coefficients of eq.2.14 applied to the intensity of local recessions/recoveries. The outcome is consistent across detrending procedures: while for recessions there is a positive and statistically significant correlation between local concentration *up* to the peak and the intensity of the ensuing recession, for recoveries the correlation is negative and emerges mostly *during* the recovery. While the first pattern provides support to the theory on the micro-origin of aggregate fluctuations, the second pattern does not, since the timing goes against the hypothesis of causation. Furthermore, even if the correlation was evident only up to the local economy entering an expansion, by the symmetry of the theory we would expect to find a positive relationship, while in fact we find a negative one.

Finally, we provide a more accurate test on the Granger-causality linking concentration of economic activity to local business cycles. The estimated marginal effects in eq.2.15 and corresponding 90% confidence intervals are reported in Fig.2.10 (linear detrending/break-adjustment) and Fig.2.11 (HP filtering). Within each figure the upper-left (right) plot reports the coefficients from the estima-

[22]We experimented also with an alternative definition of aggregate business cycles: we defined a peak as the start of a local business cycle that is aggregate when the cumulative number of cities also experiencing a peak within a year is above 75% of the total number of cities. The results are robust to this alternative definition.

tion of plain-vanilla eq.2.15 around the peak (trough) without time fixed effects; the bottom-left (right) plot reports the coefficients from the estimation of eq.2.15 around the peak (trough) with the inclusion of time fixed effects.

Quite strikingly, local concentration has consistently a statistically significant and positive (negative) association with the probability of the economy entering a recession (recovery) exclusively in the one-quarter neighborhood of the local peak (trough). Furthermore, the coefficient is maximized (in absolute terms) at the local peak (trough). The interpretation of the magnitude of the estimated marginal effects is unfortunately non-straightforward, since these are valid only locally: a small positive (negative) deviation from steady state in local concentration increases (reduces) the probability of a recession to occur (expansion to start) by 1.5 times as much as the change in concentration.

When concentration is high, the economy is vulnerable and susceptible to entering a downturn, so that the probability of a negative shock to a large establishment dragging the local economy into a recession is the highest. On the other hand, the data reject the hypothesis that the same holds true with respect to expansions. This evidence suggests therefore that 1) the presence of a few very large establishments is likely to plunge the economy into a steep recession (but not to start an expansion), 2) the protracted presence of a few very large establishments is in fact likely to delay the start of the recovery and, once it starts, it may be responsible for a sluggish one.

2.4 A Narrative Approach

Following Gabaix (2011), we provide a narrative of granularity-driven local business cycles. The characteristics of the German economy make it particularly well-suited to investigate the micro-origins of local business cycles, since it features a high number of large corporations active in the industrial goods/chemicals sector and with a dense network of establishments operating in Germany: there are indeed two sectors where labor adjustment costs are expected to be low, and, hence, a higher chance for idiosyncratic shocks to play out at the aggregate level.

Of the 30 companies that are included in the DAX index: 6 are active in the chemicals sector (BASF, Bayer, Beiersdorf, Henkel, Linde and Merck); 9 are active in the industrial goods sector (Adidas, BMW, Continental, Daimler, HeidelbergCement, Infineon Technologies, Siemens, ThyssenKrupp, Volkswagen Group); 2 in the energy sector (E.ON and RWE); 3 are government-owned companies (Deutsche Lufthansa, Deutsche Post, Deutsche Telekom); 2 in the medical sector (Fresenius and Fresenius Medical Care); 6 in the finance-insurance-real estate, or FIRE, sector (Allianz, Commerzbank, Deutsche Bank, Deutsche Börse, Munich Re, Vonovia); 1 in the media sector (ProSiebenSat.1 Media); 1 in the software industry (SAP).

We conduct a manual search of all German sites where the 15 companies belonging to either the chemicals or industrials goods sector are located. Since our employment definition does not include the construction industry, we further exclude HeidelbergCement from the sample. For each company we compile a list of *Kreise* where all German plants listed in the *careers* section of the company website are located. Where needed we supplement the search with information coming from Indeed.de, the popular vacancy advertising website, that allows to search for all vacancies posted by a specific company, grouping search results according to the location of advertised vacancies. The average number of German sites per company is 8, with substantial variation: while Bayer and BMW have 13 production facilities in Germany, Merck has only one, since most of the company's production takes place in the United States.

Next, we download for each company historical stock closing price data for regular shares from Yahoo Finance at the weekly frequency. We seasonally adjust the series, and calculate quarterly stock price growth rates.[23] Stock price data for a given company are matched to employment data for each *Kreis* where the company has a production site. Finally, we calculate the correlation for each of the 113 company/*Kreis* matched pairs.

The median correlation is positive and in several instances statistically significant. Fig.2.12 reports both the company stock price growth rate and the local employment growth rate for 5 company/location pairs featuring the highest and statistically significant correlation between the two series.

BMW, *Landshut* and *Wackersdorf* factories The first and second plot of Fig.2.12 show BMW stock price evolution together with employment in Landshut (Bayer) and Wackersdorf (Bayer). Nowadays, Landshut factory has 4100 employees, while the Wackersdorf facility employs 3000 people.[24] Landshut population stands at around 70000 inhabitants, while average private sector full-time employment in AWFP data is 18000. Wackersdorf is part of *Schwandorf* district (*Kreis*), which has population of 140000 and average employment in AWFP data of 30000.

In the first quarter of 2003 Landshut employment drops by 5%: this is the largest drop registered in Landshut over the entire sample. Schwandorf employment experiences a comparable drop in aggregate employment ($\approx 3\%$) anticipated by one year, hence during the first quarter of 2002. The period between the first quarter of 2002 until the second quarter of 2003 represents an unlucky spell for BMW stock as well: the closing price drops from 40.2 in 2002Q2 to 28.3 in 2003Q1.

[23]The price is set in a given quarter equal to the price during its first week.

[24]See https://www.bmwgroup-werke.com/.

What happened in 2002 that potentially drove down both BMW stock price and Landshut and Schwandorf employment? The first quarter of 2002 marks the start of a wave of strikes in Germany organized by IG Metall: the largest metalworkers union in Germany at that time demanded a shorter working week for factories located in the East, corresponding to a reduction from 38 to 35 weekly working hours, the latter being the number of working hours in place in West Germany factories. Disruptions in production were passed on to firms located further on along the supply chain, BMW being one of those. Newspapers articles report that several BMW production facilities in the first quarter of 2003 had to let go several thousands of workers for an unpredictable period of time.[25] In mid-2003 the strikes ceased, and both employment in Landshut and Schwandorf and BMW stock price rose again.

Bayer, *Kiel* facility Kiel (Schleswig-Holstein) has 240000 inhabitants and it is located in the north of Germany. Owing to the presence of Christian-Albrechts-Universität, founded in 1665 and largest university in the region, and its population of over 25000 students, Kiel was a natural choice as host of one of the facilities of Bayer HealthCare AG, subsidiary of Bayer AG. The site's name is KVP Pharma + Veterinary Products GmbH: it has recently experienced very fast growth since its inception in 1974. The site has nowadays over 800 employees and it produces 50% of all Bayer veterinary medicinal products sold worldwide.[26]

The Animal Health Business Group has featured steady positive growth in net sales over the past years for Bayer AG, and the company has committed to invest €92 millions in the expansion of the Kiel production site by 2021.[27] In spite of the fast growth experienced by this site, which has moved from 500 to more 800 employees in the last 7 years, it is still relatively small for a city with average employment equal to 63000 in AWFP data. Additionally, the Animal Health Business Group - for which Kiel represents one of the two main production sites worldwide - is solid but overall makes up a small fraction of total sales: in 2014 net sales for the animal health segment were €1320 million as opposed to €41340 million sales overall for the group.[28]

Nevertheless, Kiel is one of the most rapidly expanding German cities, with an average growth rate of 50 basis points that has been increasing by one basis point each year in AWFP data: while we are unable to test for the presence of any structural break prior to the opening of the Kiel production site due to unavailability of data, one can guess that the presence of Bayer has boosted local entrepreneurship

[25] See https://www.theguardian.com/world/2003/jun/21/germany.jeevanvasagar.

[26] See https://www.kiel.bayer.de/en/home/index.php.

[27] See https://biooekonomie.de/en/nachrichten/kiel-bayer-invests-millions-animal-health-site.

[28] See https://www.investor.bayer.de/en/reports/archive/.

and innovation-mix and paved the way for the strong trend in employment growth observed over the past 25 years.

Beiersdorf, *Hamburg* site Beiersdorf AG is a German company active in the chemicals sector. It was founded in 1882 and it is headquartered in Hamburg, where it employed 3100 workers in 2003.

"In 2003, a 2-year bidding war ended. Procter & Gamble, an American competitor, had sought to purchase Beiersdorf and proposed a take-over deal to Allianz insurance, which then held 19.6% of Beiersdorf's stock. Fearing that Procter & Gamble was interested only in Beiersdorf's brands and not in the company as a whole, many in Hamburg preferred to retain local ownership. The city of Hamburg and its state-owned holding company HGV created such a solution. The Herz family, owner of the German company Tchibo, who already had a stake in Beiersdorf, increased their holdings to 49.9%. Allianz still held 3.6%; Beiersdorf AG bought up 7.4% of its shares, of which 3% were given to the Beiersdorf pension fund. Another share holder, a private family, retained their share. This public-private alliance ensured that Beiersdorf's headquarters would remain in Hamburg and continue to provide hundreds of jobs, while paying taxes of approximately €200 million annually."[29]

The fear of investors during this time of uncertainty is fully captured by the evolution of the stock price: Beiersdorf stock lost more than 25% of its value during 2002, dropping from 44 in 2002Q1 to 33 in 2002Q3. At the same time, one can see from the fourth plot in Fig.2.12 that in 2003Q4 a spell of negative employment growth that started in 2001Q3 finally came to an end: this period of time coincides with the one over which the take-over battle by Procter & Gamble took place.

One can imagine that during this time of heightened uncertainty and decline in global sales the company paused the hiring process and sought a contraction in its labor costs. The timing of events and the importance of Beiersdorf for the local economic environment provide good proof of the causal link between the difficulties experienced by Beiersdorf in 2002/2003 and the long spell of subdued employment growth the city of Hamburg went through during the same period.

Daimler, *Rastatt* factory The Daimler AG plant in Rastatt nowadays employs 6500 employees and it is the lead plant in compact car production. Rastatt is a district (*Kreis*) located in Baden-Württemberg. It has a population of 227000 inhabitants, and average private sector full-time employment of 56000 in AWFP data.

[29] See https://www.abendblatt.de/wirtschaft/article106724487/Sieg-fuer-Hamburg-Tchibo-und-Stadt-kaufen-Beiersdorf.html.

During the Great Recession, and in the first quarter of 2009 specifically, the works council of the Rastatt plant agreed on switching to a short-time working regime, as a strategy to cope with the difficulties the company had been experiencing during the previous months.[30] Even if the short-time working arrangement minimized the impact of weak demand on layoffs, full-time employment dropped dramatically, as one can see from the fifth plot in Fig.2.12. Given the size of the production plant relative to total employment in Rastatt, it is likely that Daimler dismissals during the Great Recession have had sizeable repercussions on aggregate local employment. Nevertheless, it cannot argue that the slump in local employment experienced by the city of Rastatt during 2009 was caused by conditions specific to Daimler AG.

If it is difficult to point at episodes during which negative idiosyncratic experienced by Daimler AG affected employment in Rastatt, it is also because positive episodes in the history of this production plant by far outnumber negative episodes. In 1997, for example, employment in Rastatt grows by 3.5% over just two quarters. At the same time Daimler AG stock price rises by 42%, from 59 in 1997Q1 to 84 in 1997Q3. What happened during these two miraculous quarters? Large-scale production of the Mercedes-Benz A-Class kicked-off in June 1997: by the time production started, the Rastatt plant was employing just under 4000 people, almost twice as many as it had in October 1996.

Similarly, in 2004 Mercedes-Benz announced the start of B-Class production in the Rastatt plant: in 2004Q3 Rastatt aggregate employment grows by 1.4%, against an average growth rate of nearly zero over the 25 years considered. These two examples provide strong narrative evidence in support of granularity-driven local business cycles.

2.5 A Review of Restructuring Support Instruments

In many countries large firms are the subject of multiple regulations: this special attention is a telling signal of the important role played by large corporations for economic development. The analysis presented in this paper suggests that a strictly positive amount of asymmetric regulation might be socially optimal.

The economy described in this paper is characterized by firms that expand or shrink out of luck without internalizing the impact of their firing/hiring decisions on local labor market conditions: it follows that when a large firm receives a negative shock, it fires too many workers, while when it receives a positive shock, it hires too many. A social planner should therefore tax large firms receiving a positive shock, and subsidize large ones being hit by a negative shock. Furthermore, a

[30]See http://www.daimler.igm.de/news/meldung.html?id=28246.

more than proportional intervention in the former case would cater to the macroprudential objective of "leaning" against the formation of too large corporations, thus reducing the scope for "cleaning" ex post.

A few considerations undermine the external validity of this policy intervention, which should be seen as correct only in light of the assumptions governing the admittedly stylized behavior of large firms described in the theoretical section. First of all, and through the lenses of the results obtained in this paper, the impact of granularity on macroeconomic volatility appears to be asymmetric: large firms are not as likely to trigger an expansion as much as they are at starting a recession. A tax levied on large businesses experiencing "good times" would then serve primarily the above mentioned macroprudential goal, rather than the one of smoothing macroeconomic volatility. Secondly, the analysis presented in this paper does not view firms as actively working towards their success nor it allows for imperfect competition that would reinforce the competitive edge of large players. Both extensions represent a gain in realism and demand, at the same time, a more decisive intervention of the social planner in "good times": absent such disciplining device, firms would make excessive investments aimed at increasing their market share - and thus macroeconomic volatility - while remaining shielded by government subsidies should their business go sour. Lastly, the analysis presented in this paper completely overlooks positive externalities associated with the presence of few large firms in the economy, e.g., productivity spillovers: these externalities are likely to be relevant, especially for economic developments in the long-run.

Nevertheless, the relevance of these arguments does not call into question the appropriateness of a twofold public intervention aimed at 1) preventing the accumulation of systemic risk in the economy - through the imposition of more than proportional costs on larger companies and following reduction in their size, 2) cushioning the impact on the economy of shocks to larger companies. Given the vast literature on the first type of interventions (Guner, Ventura and Xu (2008) and Garicano, Lelarge and Van Reenen (2016)), in this section, we focus on the second set of policies.[31] In particular, we review the main schemes deployed by a subset of major European economies in support of large companies undergoing a phase of restructuring.

A restructuring event 1) is triggered by either a positive or negative shock to an individual company, and 2) it involves a substantial amount of job destruction or job creation.[32] Given their systemic importance, one of the missions of Eurofound, the European Foundation for the Improvement of Living and Working

[31]Guner, Ventura and Xu (2008) and Garicano, Lelarge and Van Reenen (2016) discuss both theoretically and empirically how size-contingent policies affect economy-wide distortions.

[32]While the interest of this paper lies in labor market fluctuations at the business cycle frequency, restructuring can also take place in response to permanent shocks, e.g., technological change.

Conditions, is precisely to monitor the incidence and magnitude of restructuring events involving EU large companies. The European Monitoring Centre on Change, one of the Eurofound Observatories, manages the European Restructuring Monitor, which offers support to one of the objectives of the agency, i.e., monitoring the pace of structural change in EU labor markets.

The European Restructuring Monitor (henceforth ERM) provides a database of all large-scale restructuring events that have taken place since 2002 in EU countries, and thus featuring 22000 restructuring events to date.[33] The ERM defines a large-scale restructuring event as a circumstance that entails the *"announced destruction or creation of at least 100 jobs, or at least 10% of the workforce at sites employing more than 250 people"*. Metadata are available for download by country, sector of economic activity and restructuring event type. Fig.2.13 and Fig.2.14 reports the frequency of internal restructuring (left) and business expansion (right) events during 2002-2018 across EU countries together with median job destruction and job creation, respectively, in terms of employee headcount.

The ERM also surveys public restructuring support schemes deployed by individual countries. There are several ways to group these policies. One possibility is to group them into a) policies that limit the impact on employment, and b) policies that limit the impact on total hours worked. An institutional background that allows firms to absorb shocks via variation in the intensive margin of labor is certainly a desirable feature for the economy: the firm is neither forced to hoard labor nor to lose trained workers; workers maintain their contracts opened and their bargaining power so that workers purchasing power does not decline much. An economy that features more flexibility along the intensive margin is more resilient to aggregate fluctuations. Another way of clustering these policies is into a) those that make the adjustment more costly for firms, and b) those that make adjustment by firms less costly for workers through a system of subsidies (e.g., short-time working allowances). Both types of policies are aimed at limiting the social cost of labor market fluctuations: the second set of policies is however in a better position to attain such goal since by not imposing additional costs onto firms already in financial distress, they can effectively minimize the rise in unemployment due to firm exit.

We focus primarily on the second type of policies because they entail lower efficiency losses. We only consider restructuring support schemes targeted to in-

[33]Microdata can be requested for research purposes by EU-based researchers: they are gathered through a collective effort entertained by a Network of Eurofound Correspondents across EU countries, who screens at the daily frequency a wide range of business press and online sources. The European Restructuring Monitor also conducts research on restructuring events in SMEs: given the smaller systemic importance of these episodes and the higher difficulty in detecting them, these are not included in the database. For more information see https://www.eurofound.europa.eu/observatories/emcc/european-restructuring-monitor.

dividual firms when they are still *going concerns*. In what follows, we provide an overview of the schemes currently in place in four major European economies: Germany, France, Italy and Spain.

France Size-dependent policies are present in an extraordinarily large set of labor market regulations in France (Garicano, Lelarge and Van Reenen (2016)). The subset pertaining to the public support offered to companies experiencing economic difficulties is however fairly narrow.

Companies employing more than 50 workers must appoint an expert (*Experts auprès des Comités d'entreprise*): he is in charge also of providing expert judgment and solutions on how to limit the social cost of large-scale dismissals.

National funds can finance allowances to employers resorting to short working hours (*activité partielle*). Companies can file for state support after having implemented a working time reduction if this is caused by 1) the economic situation, 2) supply difficulties, 3) a disaster or weather conditions of an exceptional nature, 4) the transformation, restructuring or modernisation of the company, 5) or any other exceptional circumstance such as the loss of a principal client. Employers must pay 70% of the gross hourly wage for each hour subtracted from the contracted working time per worker (and in any case not below the minimum wage): they receive a compensation equal to €7.74 (€7.23) per hour per worker for companies with fewer (more) than 250 workers.[34] The objective of this policy is to mitigate the labor market impact of crises by providing incentives for the adoption of flexible working time arrangements. Aside from the mild pecuniary difference in favor of smaller firms, the fact that firms with a works council (hence, firms with more than 50 employees) can only resort to this kind of instrument if it has been agreed upon with the works council puts large companies at a disadvantage when it comes to benefiting from public support in crisis management.

The resilience of larger firms to negative shocks is further challenged by the presence of asymmetric firing costs, since the *Code du Travail* establishes that firms with more than 50 employees must use a complex redundancy plan with oversight, approval, and monitoring from the Ministry of Labor in case of a collective redundancy for 9 or more employees (Garicano, Lelarge and Van Reenen (2016)).[35]

[34]See https://www.eurofound.europa.eu/it/observatories/emcc/erm/support-instrument/partial-activity.

[35]Lastly, firms with more than 50 employees must draw up a *plan de sauvegarde de l'emploi* (employment protection plan), whose goal is to minimize the number of redundancies in times of crisis and facilitate the redeployment of dismissed workers when redundancy is unavoidable. Different types of public funding is available depending on the circumstances: however, Eurofound highlights how this instrument is not very diffused, and that the conditions that determine the existence and magnitude of public re-

Germany The most famous form of employment support in Germany to firms undergoing difficulties with their business is the *Kurzarbeiterfund*, or short-time working allowances.

For a company to be eligible for short-time working allowances at least a third of the company workforce must be affected and total staff income loss must exceed 10%. Short-time working schemes can be of three types: 1) short-time working activated by temporary economic difficulties/shortfall of orders (*konjunkturelle Kurzarbeit*); 2) seasonal short-time working (*Saisonkurzarbeit*); 3) short-time working in the event of restructuring (*Transferkurzarbeit*).[36] Once a firm has obtained eligibility for short-time working allowances, the employer pays for the effective working time and the Federal Employment Agency contributes a short-time working allowance of 60% of the missing net wage.

While there is in principle no restriction on how large a firm must be before it can apply for short-time working so long as it has at least one registered job-holder, in practice it is large firms that have most intensely relied on these schemes during the Great Recession (Brenke, Rinne and Zimmermann (2011)).[37]

Short-time working schemes differ from other types of short-time working arrangements, such as the so-called "alliances for jobs" (*Betriebliche Bündnisse für Arbeit*): these are also negotiated between works council and the employer but the firm receives no subsidies from the authorities, and the cost of working time reduction is entirely born by the employees.

Italy The analogue to *Kurzarbeitergeld* for large firms in Italy is the *Cassa integrazione guadagni* or CIG (Wage Guarantee Fund). The CIG is a type of support provided by the state to companies experiencing difficulties: the *Ministero del Lavoro e delle Politiche Sociali* (Ministry of Labour and Social Policies) pays a wage guarantee in the form of 80% of the compensation that the employee would have received for the hours not worked and within the limit of 40 hours per week. The objective of the intervention is to avoid staff layoffs by reducing the labor costs weighing on the company's shoulders.

The discipline of the CIG has been reformed by the *decreto attuativo* Jobs Act

---

sources are very opaque. See https://www.eurofound.europa.eu/observatories/emcc/erm/support-instrument/employment-protection-plan-pse for additional information.

[36] See http://edz.bib.uni-mannheim.de/daten/edz-ma/esl/10/EF10632EN.pdf.

[37] At the peak of the crisis, one in five German companies with 500 and more employees were affected by short-time work: this pattern is easily explained by the fact that large manufacturing and export-oriented firms have been the most damaged by the negative global cycle during the Great Recession (Brenke, Rinne and Zimmermann (2011)). At the same time, short-time work has declined faster for large companies: of all short-time workers, one in three could be attributed to small companies at the beginning of 2011, while it was only one in ten in the second quarter of 2009.

n.148/2015, which divides it into *Cassa integrazione guadagni ordinaria* or CIGO (Ordinary Wage Guarantee Fund), available to all firms operating in certain industries and irrespective of their size, and *Cassa integrazione guadagni straordinaria* or CIGS (Extraordinary Wage Guarantee Fund), available to companies with more than 15 (or 50) employees depending on the industry.[38] The extraordinary wage guarantee can be activated in case of 1) company restructuring, 2) company crisis, 3) solidarity contracts. Payments cannot last longer than 24 months over a 5-year period in the first and third case, 12 months in the second case.[39]

The second type of support scheme for large firms are the *contratti di solidarietà* (solidarity contracts): these are agreements between labor unions and employers on the reduction of working time aimed at avoiding staff layoffs (*contratti difensivi*) or fostering job creation (*contratti espansivi*). Solidarity contracts are available to all firms falling within the perimeter of the CIGS legislation.[40]

Finally, it is worth mentioning that Italy features asymmetric firing costs based on firm size. Up until 2015 firms above 15 employees were subject to the obligation to reinstate (and compensate for the lost wage) the dismissed worker if a judge had ruled that there existed no just cause for the dismissal. Since 2015 firms above 15 employees are subject to the reinstatement obligation only if the dismissal took place on discriminatory grounds; when it instead takes place on economic grounds the firm must make a severance payment that increases in worker experience (*contratto a tutele crescenti*). No obligation to reinstatement/compensation after dismissal is applied to firms below 15 employees.

Labor flexibility is thus impaired especially for larger firms, and the special restructuring support schemes targeting large Italian firms just described help mitigate the asymmetry of treatment built into the Italian employment protection legislation.[41]

Spain Following Law 3/2012, Spanish companies can implement temporary layoff plans (*Expediente Temporal de Regulaciòn de Empleo*) irrespective of their size and the number of workers affected, and they do not need to obtain anticipated administrative authorization. Such plans can involve either temporary layoffs or

[38]The set of activities covered by the CIGS and the respective size thresholds can be found at https://www.guidafisco.it/cassa-integrazione-straordinaria-800.

[39]The CIG is funded by ordinary and extraordinary contributions that must be paid every period by employers and employees benefitting from the scheme. In the case of CIGS the ordinary contribution corresponds to .9% of a worker compensation, to be split 2/3-1/3 between employer and employee. More details can be found at: https://www.eurofound.europa.eu/it/observatories/emcc/erm/support-instrument/short-time-allowances-ordinary-wages-guarantee-fund-cigo-and-extraordinary-wages-guarantee-fund-cigs.

[40]See https://www.inps.it/nuovoportaleinps/default.aspx?itemdir=46131.

[41]Boeri and Garibaldi (2018) use a regression discontinuity design to study how the changed legislation of the Jobs Act affected firm growth for firms below the threshold of 15 employees.

a reduction in working time. In case of temporary layoffs, workers perceive unemployment benefits and the employer pays social security contributions, while in case of working time reduction workers are entitled to partial unemployment benefits. Upon termination of the restructuring process, workers go back to their job at the same contractual conditions as before the crisis.

During the Great Recession the Spanish government has worked towards making use of this instrument more appealing from the firms perspective: this action has mitigated the subsequent increases in unemployment after the dramatic spike in 2009. No special cushion is considered for difficulties experienced by large companies in Spain: conversely, collective dismissals - which are more likely to occur in large companies - traditionally have been very costly (Dolado, Garcia-Serrano and Jimeno (2002)).

2.6 Conclusion

Large firms are responsible for adding endogenous uncertainty to the economy: a negative shock to BMW in 2002/2003 was responsible for employment in the city of Landshut to drop by 5%. There exist many newspaper articles and policymaker speeches where specific plants are pinpointed as neuralgic centres of local economic activity.

Guided by narrative evidence and the theoretical framework in Carvalho and Grassi (2017), this paper tests three predictions in the data: 1) the more fat-tailed is the steady state firm size distribution, the more persistent are macroeconomic aggregates; 2) the more fat-tailed is the steady state firm size distribution, the higher is aggregate unconditional volatility; 3) an increase (decrease) in concentration Granger-causes an increase (decrease) in aggregate conditional volatility. With the support of quarterly data on the evolution of local employment and concentration of economic activity across 72 Metropolitan Areas in former West Germany over the past 25 years, we find evidence validating all of these predictions.

We consider the stylized fact of local concentration Granger-causing local busts as the most original finding of this paper. Additionally, we highlight how the data reject the symmetry of the theory in Carvalho and Grassi (2017): while high concentration is unlikely to trigger an expansion of the local economy, we show that it acts so as to reduce the speed of the recovery once this has started.

In light of the assumptions governing large firms behavior in this paper and the ensuing systemic risk externality imposed by them on the economy, a social planner should tax large firms receiving a positive shock, and subsidize large ones being hit by a negative shock. Furthermore, a more than proportional intervention in the former case would cater to the macroprudential objective of “leaning” against the formation of too large corporations, thus reducing the scope for “cleaning" ex post.

Of the described twofold policy intervention, we focus on public support schemes geared towards large firms experiencing economic difficulties. Among those, policies that incentives firms' adoption of flexible working times schemes appear to have been the most successful ones during the Great Recession: the firm is neither forced to hoard labor nor to lose trained workers; workers maintain their contracts opened and their bargaining power so that workers purchasing power does not decline much. We provide a review of the most relevant policies that aim at mitigating the impact on the labor market of shocks to large firms in four major European economies: we find very large variation in the adjustment margin incentivized by different policies, as well as to the extent of preferential treatment devoted to systemically important firms.

Appendices

Tables

Table 2.1: City size and concentration.

	E-prop. above 99th pct.	E-prop. above 95th pct.	E-prop. above 90th pct.
Log size	0.0342^{***}	0.0279^{***}	0.0212^{***}
	(3.94)	(4.59)	(4.48)
Observations	72	72	72

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Complementarity between large establishments and large cities: E-prop. above x^{th}percentile$_m$ $= \alpha + \beta \ln \text{Size}_m + e_m$, where m indexes cities. Source: AWFP.

Table 2.2: Instantaneous volatility and concentration of economic activity.

	Dep.var.: $\sqrt{\frac{\pi}{2}}\lvert\hat{\epsilon}_{m,t}\rvert$		
$x^C_{m,t}$		-0.153*	-0.147*
		(-2.14)	(-2.12)
$x^C_{m,t-1}$	-0.0558	0.0711**	0.0742**
	(-0.92)	(2.74)	(2.97)
$\sqrt{\frac{\pi}{2}}\lvert\hat{\epsilon}_{m,t-1}\rvert$			0.0944***
			(5.06)
Observations	7128	7128	7056
R^2	0.223	0.229	0.235
MA FE	yes	yes	yes
Time FE	yes	yes	yes

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Instantaneous volatility is given by $\sqrt{\frac{\pi}{2}}\lvert\hat{\epsilon}_{m,t}\rvert$ obtained by fitting at the city level an AR(1) process on log employment. The independent regressor $x^C_{m,t}$ is the cyclical component of local concentration. t-statistics are based on robust standard errors. Source: AWFP.

Table 2.3: Local business cycles properties descriptives.

	Cycle: linear detrending and break-adjustment			
	Mean	Std dev	Max	Min
Length Recession	8.20	5.77	31.00	2.00
Depth Recession	5.00	4.21	35.55	0.04
Intensity Recession	0.66	0.73	13.13	0.02
Length Recovery	7.98	4.93	20.00	2.00
Depth Recovery	4.91	3.68	29.67	0.11
Intensity Recovery	0.64	0.47	4.94	0.06
	Cycle: HP-filtering			
	Mean	Std dev	Max	Min
Length Recession	6.61	3.82	20.00	2.00
Depth Recession	3.54	2.39	26.30	0.09
Intensity Recession	0.61	0.68	13.15	0.02
Length Recovery	6.60	4.06	21.00	2.00
Depth Recovery	3.42	2.60	27.30	0.03
Intensity Recovery	0.55	0.40	4.55	0.01

Business cycles turning points are identified by the Harding and Pagan (2002) dating procedure applied to the cyclical component of the (log) of private sector full-time employment at the local level. The length of recessions/recoveries is measured in quarters; the depth in percentage points; the intensity is the ratio between the depth expressed in percentage points and the length expressed in quarters. Source: AWFP.

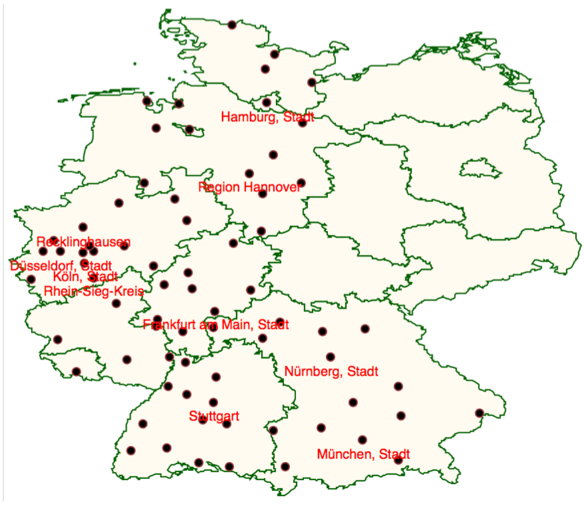
Table 2.4: Local business cycle properties and average local concentration.

Cycle: linear detrending and break-adjustment							
All local business cycles							
	Recession			*Recovery*			
	Depth	Intensity	Length	Depth	Intensity	Length	N
$\widehat{\beta}$	0.063	0.001	0.507	0.044	0.006	0.542	385
t-stat	2.431	0.344	1.750	1.633	1.964	1.801	.
Excluding aggregate cycles							
	Recession			*Recovery*			
	Depth	Intensity	Length	Depth	Intensity	Length	N
$\widehat{\beta}$	0.115	0.009	0.310	0.052	0.008	0.606	149
t-stat	2.579	1.741	0.563	1.179	1.685	1.466	.
Cycle: HP-filtering							
All local business cycles							
	Recession			*Recovery*			
	Depth	Intensity	Length	Depth	Intensity	Length	N
$\widehat{\beta}$	0.009	0.008	-0.252	0.014	0.004	-0.156	447
t-stat	0.524	1.633	-0.861	0.711	1.445	-0.521	.
Excluding aggregate cycles							
	Recession			*Recovery*			
	Depth	Intensity	Length	Depth	Intensity	Length	N
$\widehat{\beta}$	0.043	0.033	-1.034	0.047	0.008	0.021	115
t-stat	1.125	1.982	-1.569	1.067	1.238	0.041	.

Business cycles turning points are identified by the Harding and Pagan (2002) dating procedure applied to the cyclical component of the (log) of private sector full-time employment at the local level. The length of recessions/recoveries is measured in quarters; the depth in percentage points; the intensity is the ratio between the depth expressed in percentage points and the length expressed in quarters. Source: AWFP.

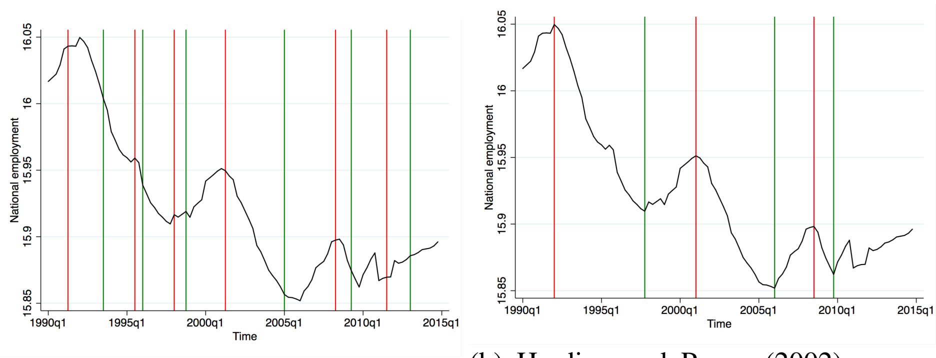
Figures

Figure 2.1: Metropolitan Areas (*Arbeitsmarktregionen*) in former West Germany.



Top-10 most populated (2013, DESTATIS) Metropolitan Areas have been labeled, the label corresponds to the most populated *Kreis*. Source: Eurostat.

Figure 2.2: Comparison between business cycle dating procedures.

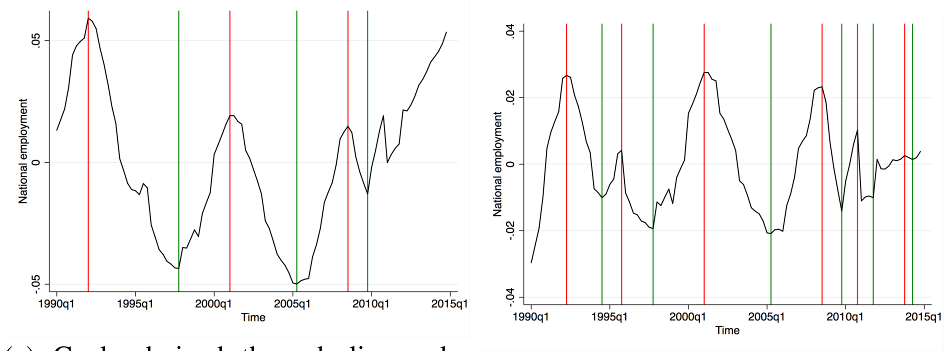


(a) OECD

(b) Harding and Pagan (2002) on national employment

OECD business cycle dates vs. Harding and Pagan (2002) dating methodology applied to national (log) employment. Red (green) vertical lines correspond to peaks (troughs). Source: AWFP.

Figure 2.3: Comparison between business cycle dates given different detrending methods.

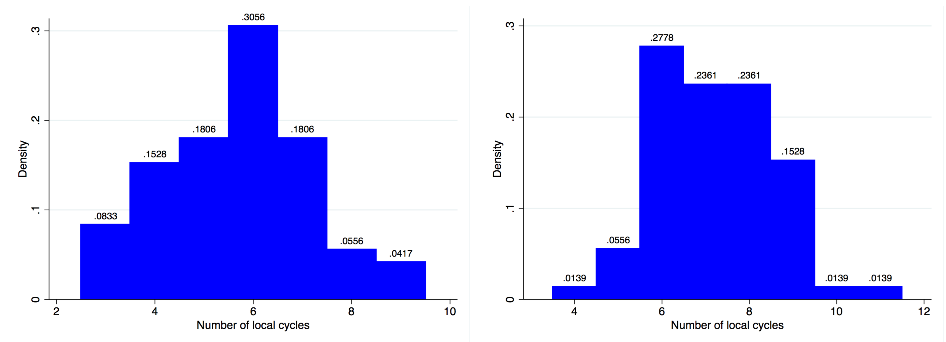


(a) Cycle derived through linear detrending

(b) Cycle derived through HP filter

Harding and Pagan (2002) dating methodology applied to the cyclical component of national (log) employment. he cyclical component has been isolated by means of a linear detrending procedure applied to national (log) employment (left), or by means of a Hodrick-Prescott filter (right). Red (green) vertical lines correspond to peaks (troughs). Source: AWFP.

Figure 2.4: Comparison between frequency of local business cycles given different detrending methods.

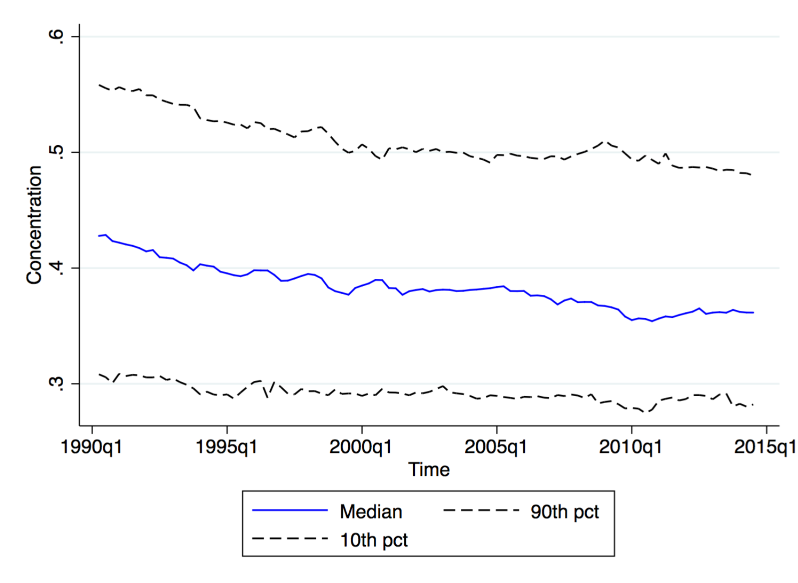


(a) Cycle: linear detrending

(b) Cycle: HP filtering

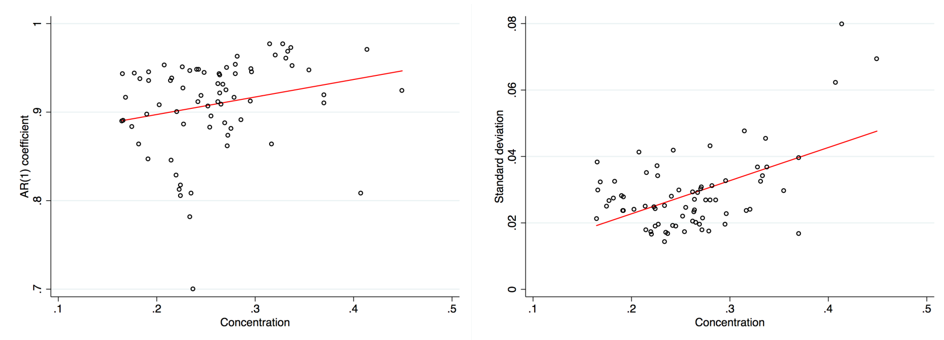
Histogram of the number of local business cycles identified at the Metropolitan Area level through Harding and Pagan (2002) methodology applied to the cyclical component of local (log) employment. The cyclical component has been calculated through a linear trend fitted on local (log) employment (left), a HP filter applied to local (log) employment (right). The frequency is reported on top of each bar. Source: AWFP.

Figure 2.5: Trend in local concentration.



Local concentration: median, 10^{th} and 90^{th} percentiles in the cross-sectional distribution at any point in time during 1990q1-2014q4. Concentration is measured by the fraction of local employment in establishments located above the local 99^{th} percentile of the establishment size distribution. Source: AWFP.

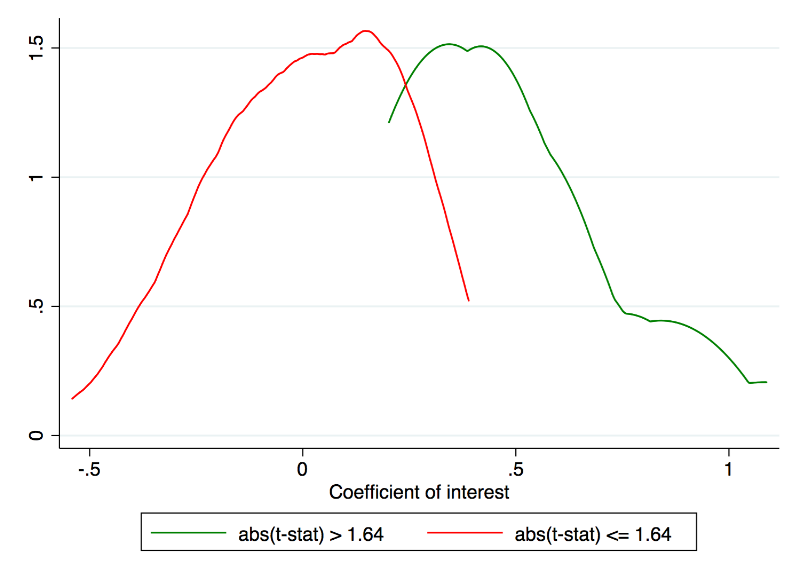
Figure 2.6: Employment persistence/volatility and local concentration.



(a) Persistence (b) Volatility

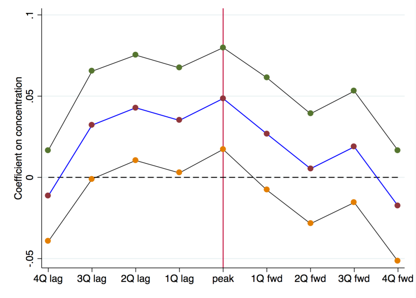
Estimated employment persistence and average volatility on average local concentration. All city-level series of log employment and concentration have been cleaned of linear trends and structural breaks both in the intercept and in the trend as a result of the labor market reforms in 2005. Source: AWFP.

Figure 2.7: Instantaneous volatility and local concentration: city-level regressions

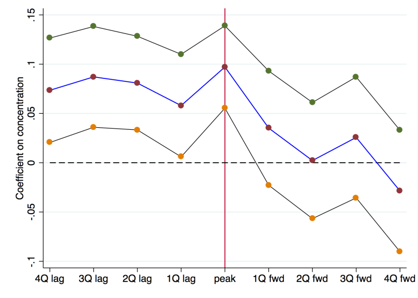


Coefficients on lagged concentration in eq.2.12 grouped by statistical significance. t-statistics are based on robust standard errors. Source: AWFP.

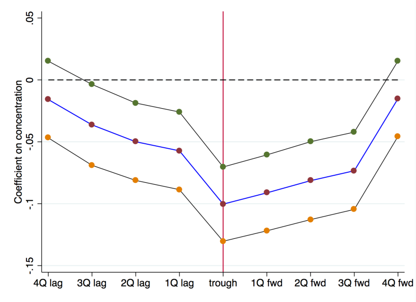
Figure 2.8: Local business cycles and concentration (linear detrending/break adjustment).



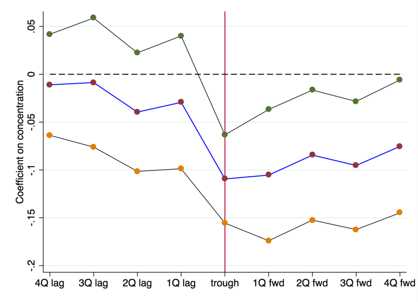
(a) Recession: all cycles



(b) Recession: only local



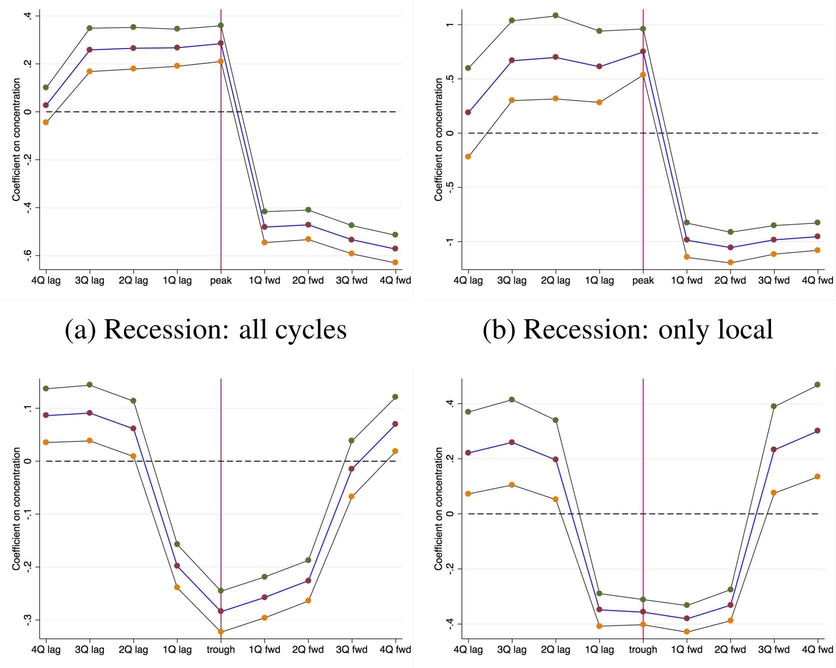
(c) Recovery: all cycles



(d) Recovery: only local

Local concentration at different lags and leads (-4/4 quarters) from the peak/trough and how it interacts with the intensity of recessions. On the vertical axis the estimated coefficient of eq.2.14 is reported together with confidence bands at $\alpha = .90$ significance level. Both local (log) employment and local concentration have been adjusted for the 2005 break and linearly detrended: the business cycle dating procedure by Harding and Pagan (2002) is applied to the cyclical component of local (log) employment, and business cycle properties are derived. The coefficients reported in left (right) column graphs are based on the unrestricted (restricted) sample. The restricted sample excludes local business cycles that are likely to be aggregate in nature. Source: AWFP.

Figure 2.9: Local business cycles and concentration (HP filtering).



(a) Recession: all cycles

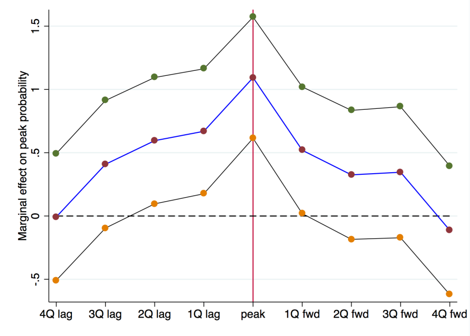
(b) Recession: only local

(c) Recovery: all cycles

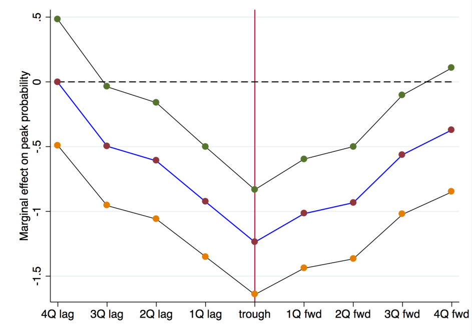
(d) Recovery: only local

Local concentration at different lags and leads (-4/4 quarters) from the peak/trough and how it interacts with the intensity of recessions. On the vertical axis the estimated coefficient of eq.2.14 is reported together with confidence bands at $\alpha = .90$ significance level. Both local (log) employment and local concentration have been HP filtered: the business cycle dating procedure by Harding and Pagan (2002) is applied to the cyclical component of local (log) employment, and business cycle properties are derived. The coefficients reported in left (right) column graphs are based on the unrestricted (restricted) sample. The restricted sample excludes local business cycles that are likely to be aggregate in nature. Source: AWFP.

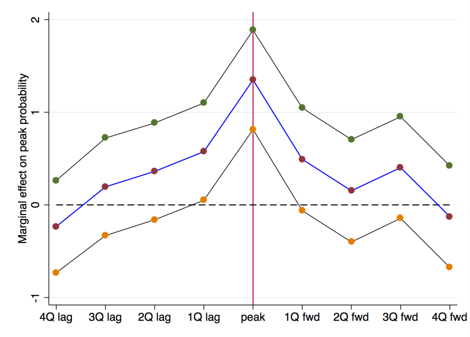
Figure 2.10: Probability of start of local recession and local concentration (linear detrending/break adjustment).



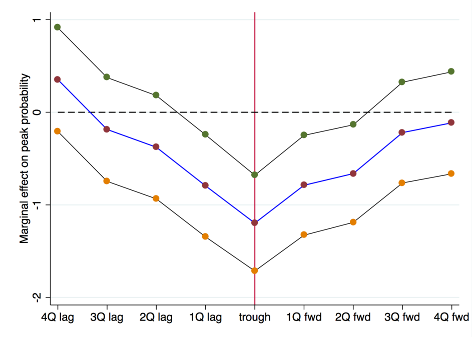
(a) PEAK: no controls



(b) TROUGH: no controls



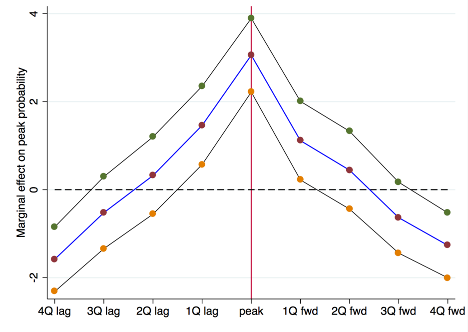
(c) PEAK: time fixed effects



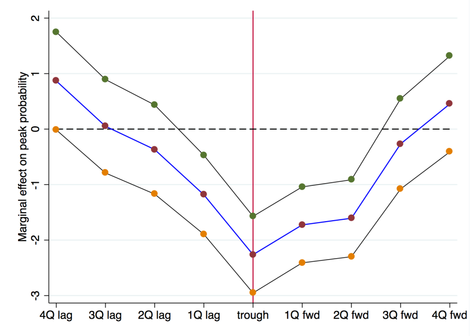
(d) TROUGH: time fixed effects

Marginal effects of the cyclical component of local concentration on the probability of start (end) of local recession, i.e., of a peak (trough) occurring, and confidence bands at $\alpha = .90$ significance level. Peaks and troughs are identified by means of the Harding and Pagan (2002) business cycle dating methodology applied to the cyclical component of local (log) employment post break adjustment/linear detrending. The cyclical component of local concentration is identified via a linear detrending procedure. The estimated model is: $P(p_{m,t} = 1) = \frac{1}{1+\exp(-\alpha-\beta x^C_{m,t+k})}$, where $p_{m,t} = 1$ if t is a local peak (trough), and $p_{m,t} = 0$ if t is not a local peak (trough), and $x^C_{m,t+k}$ is the cyclical component of local concentration at different lags ($k < 0$) and leads ($k > 0$) of t (a-b); $P(p_{m,t} = 1) = \frac{1}{1+\exp(-\alpha-\beta x^C_{m,t+k}-\delta_t)}$, where δ_t is a set of time FE for each quarter and year in the sample (c-d). The marginal effects are evaluated at the mean value of regressors. Source: AWFP.

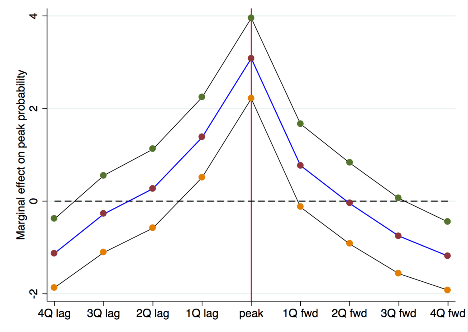
Figure 2.11: Probability of start of local recession and local concentration (HP filtering).



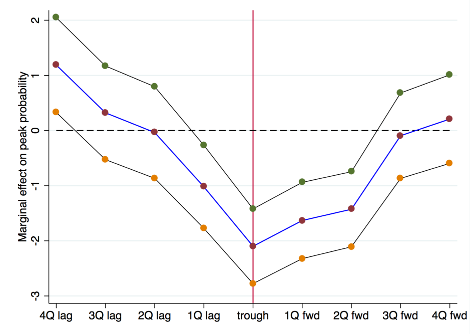
(a) PEAK: no controls



(b) TROUGH: no controls



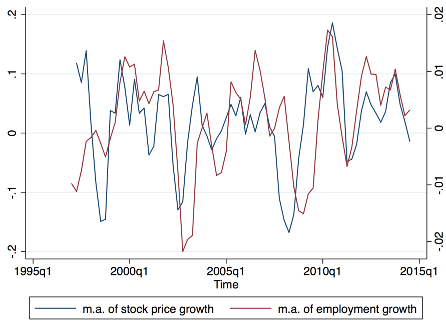
(c) PEAK: time fixed effects



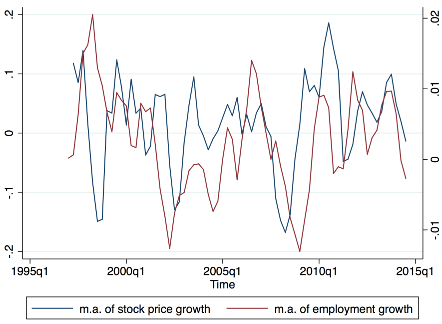
(d) TROUGH: time fixed effects

Marginal effects of the cyclical component of local concentration on the probability of start (end) of local recession, i.e., of a peak (trough) occurring, and confidence bands at $\alpha = .90$ significance level. Peaks and troughs are identified by means of the Harding and Pagan (2002) business cycle dating methodology applied to the cyclical component of local (log) employment identified via HP filter. The cyclical component of local concentration is identified via HP filter. The estimated model is: $P(p_{m,t} = 1) = \frac{1}{1+\exp(-\alpha-\beta x^C_{m,t+k})}$, where $p_{m,t} = 1$ if t is a local peak (trough), and $p_{m,t} = 0$ if t is not a local peak (trough), and $x^C_{m,t+k}$ is the cyclical component of local concentration at different lags ($k < 0$) and leads ($k > 0$) of t (a-b); $P(p_{m,t} = 1) = \frac{1}{1+\exp(-\alpha-\beta x^C_{m,t+k}-\delta_t)}$, where δ_t is a set of time FE for each quarter and year in the sample (c-d). The marginal effects are evaluated at the mean value of regressors. Source: AWFP.

Figure 2.12: Company stock price growth rate and local employment growth rate

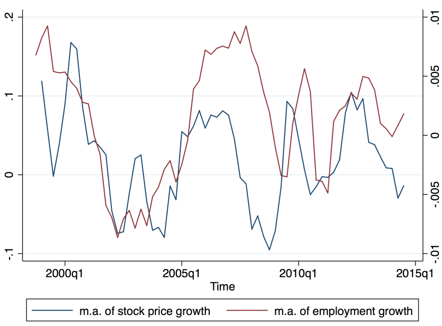


(a) BMW, Landshut: $\rho = .25$

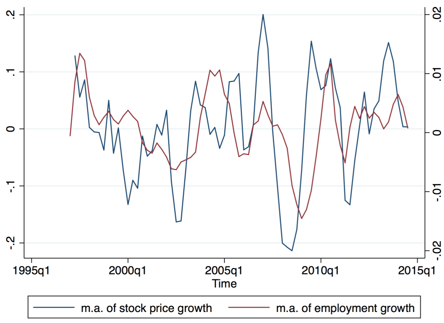


(b) BMW, Wackersdorf: $\rho = .21$

(c) Bayer, Kiel: $\rho = .21$



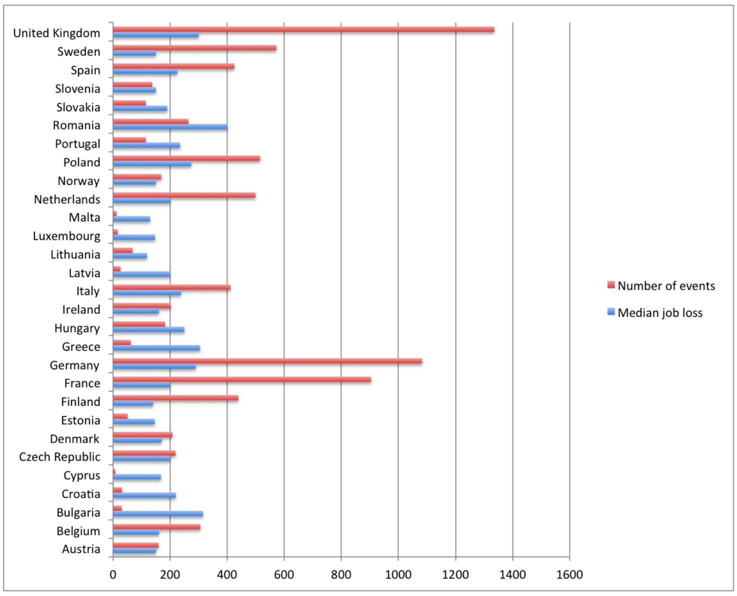
(d) Beiersdorf, Hamburg: $\rho = .24$



(e) Daimler, Rastatt: $\rho = .32$

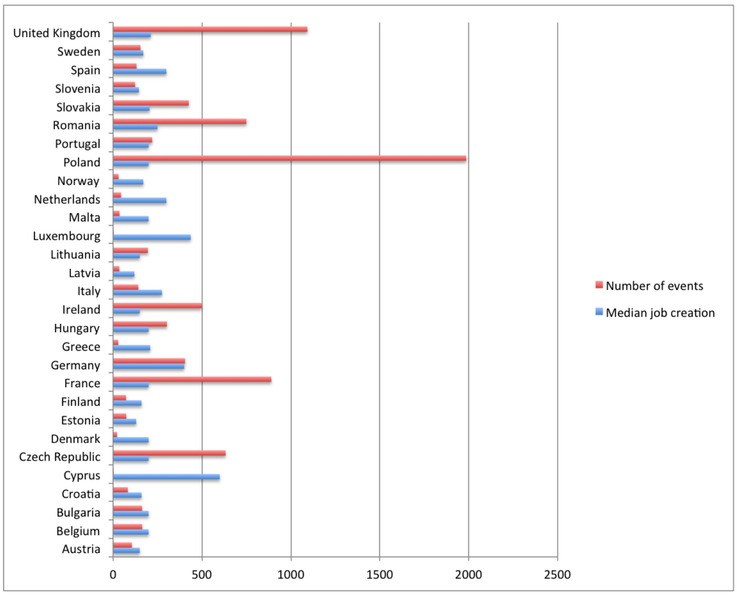
Company stock prices are downloaded from Yahoo Finance at the weekly frequency and closing price is used. Stock price growth rate corresponds to the log difference between the price during the first week of a quarter and the price during the first week of the previous quarter. Closing prices are seasonally adjusted and the correlation is calculated on the growth rate. For representation purposes, 3-quarters centred moving averages are reported. Source: AWFP and Yahoo Finance.

Figure 2.13: Internal restructuring events: number of cases and median job loss during 2002-2018 in EU countries.



Source: the European Restructuring Monitor. Data can be downloaded at: https://www.eurofound.europa.eu/observatories/emcc/erm/.

Figure 2.14: Business expansion events: number of cases and median job creation during 2002-2018 in EU countries.



Source: the European Restructuring Monitor. Data can be downloaded at: https://www.eurofound.europa.eu/observatories/emcc/erm/.

Chapter 3

SPECIALIZATION IN CITIES

3.1 Introduction

"There are some sorts of industry, even of the lowest kind, which can be carried on nowhere but in a great town. A porter, for example, can find employment and subsistence in no other place. A village is by much too narrow a sphere for him; even an ordinary market town is scarce large enough to afford him constant occupation. In the lone houses and very small villages which are scattered about in so desert a country as the Highlands of Scotland, every farmer must be butcher, baker and brewer for his own family."

A. Smith, An inquiry into the Nature and Causes of the Wealth of Nations (1776)

The notion of specialization as a source of individual productivity is an old one dating back to Adam Smith, as the above quote shows. A worker is more specialized when she is in charge of fewer tasks: being able to concentrate her time and effort on fewer assignments, according to the smithian argument she can perform each of these better than if she had been burdened with a larger number of tasks.

Learning by doing, i.e., the assumption that workers marginal productivity increases with the time spent on a task, is the key determinant of specialization-driven productivity premium. Learning by doing is often cited as the reason behind both wage growth (on-the-job or on-the-occupation) and firm productivity growth. The common aspect to these instances is that the productivity premium materializes over time: a worker can passively become more productive as she spends more time in a given job. Specialization-driven productivity, on the contrary, hinges on the active management of working time of an individual: a more efficient division of labor allows a worker to spend more on each task she is in charge of on any given day so that the productivity premium materializes faster.

The second observation highlighted in the above reported smithian excerpt is the fact that individual specialization is bounded by the extent of the market. Specialized occupations are especially valuable in larger cities because they allow for a better division of labor. A porter will not choose to locate in a small village because the compensation he would receive is too low: his too narrow focus would lower total production in the small village since it would be compensated by other workers taking up a larger subset of tasks, thus reducing the extent of individual learning by doing and aggregate productivity. These theoretical underpinnings make individual specialization a suitable candidate for the urban wage premium.

In this paper, I operationalize these insights and ask the following questions: Do specialized occupations earn a higher wage? And do they earn a higher wage especially in large cities? Are specialized occupations more than proportionately represented in large cities? Using data on US workers over three decades, I find a positive answer to all three questions, thus proving that the almost three-centuries-old smithian intuition bears strong ties with the organization of production in the US nowadays.

Furthermore, I provide an application of the specialization-driven wage premium by connecting with the literature that has explored the causes and consequences of the divergence in location choice for college and non-college workers between 1980 and 2000.[1] Specifically, I investigate to what extent differential specialization growth by city and worker type has accounted for the patterns observed in the data.

From a methodological standpoint, I construct a measure of occupational specialization by exploiting the information provided by the O*NET on the task content of occupational titles. Occupational specialization is defined as the sum of the tasks of which each occupation is comprised weighted by the frequency with which they occur. This metric is mapped into the time-consistent occupation scheme developed by Dorn (2009), thus allowing to investigate the evolution of specialization across Metropolitan Statistical Areas and years.

In terms of cross-sectional findings, I find that 1) average worker specialization is higher in larger cities, regardless of the Census year considered, 2) workers with a college degree are on average more specialized, regardless of the Census year and where they are located, 3) the rate at which specialization increases in city size is much higher for college workers relative to workers with less than a college degree. The results do not hinge on the local industry-mix.

Fig.3.4 provides a snapshot of the division of labor in US cities. A practical example of the increasing degree of specialization characterizing large cities is given by the geographical distribution of two occupations, the waiter and the catering attendant. Both occupations are typically employed in the restaurants' industry.

[1]See Moretti (2013), Diamond (2016), Giannone (2018).

While both waiter and catering attendant perform very similar duties, it is obvious that the waiter is a less specialized occupation: a waiter, on top of receiving guests and recording orders, must also perform tasks such as supply stockpiling, cleaning, or helping with the preparation of dishes. These three tasks are assigned to workers with a different occupation profile by the restaurant in Washington D.C., while they must be catered to by the waiter in the diner in Kalamazoo-Portage.

Next, I provide evidence on the specialization-driven wage premium. Exploiting Census data for multiple cross-sections, I show that specialized occupations earn a wage premium that is increasing in the size of the city where the worker is located, and that the magnitude of the interaction coefficient is almost double for college graduates. As a robustness check, I run a Mincerian type of regression on longitudinal worker-level data (Survey of Income and Program Participation), where the set of regressors is augmented to include worker specialization, local labor market size and their interaction: the specialization wage premium survives when controlling for individual-level unobservable characteristics.

Next, I investigate the role of specialization growth as a determinant of diverging location choices of college vs. non-college workers in 1980-2000. Specialization depends positively on total employment and negatively on the measure of tasks needed to produce. I take the latter as a primitive in the model so that a decline in the measure of tasks is interpreted as a shock to specialization-driven productivity.

The explanatory power of the "specialization" shock goes beyond aggregate productivity shocks resulting from the local industry-mix, as proxied in Bartik (1991). Local wage growth is positively associated with specialization growth, with cities experiencing a 1 standard deviation percent increase in workforce specialization receiving a .2 standard deviation percent wage boost. However, while over this period college workers moved to fast-growing specialization cities to reap the benefits of higher productivity, non-college workers migrated out of cities where specialization had been growing faster.

Next, I lay down a theoretical framework that is able to capture the patterns observed in the data and that can be used to quantify the relevance of specialization growth as a determinant of city and wage growth. The model is a two-skill version of the production framework outlined in Duranton and Puga (2003) enriched with a free-entry condition.

Cities differ ex-ante according to the city/skill-specific measure of symmetrically differentiated tasks that is employed in production, city/skill-specific residual productivities, and an amenity index. Workers of either type must choose where to locate: upon having located somewhere, they consume a local good in fixed supply and they compete in providing a subset of tasks used in production, receiving a wage in return for their services. Firms produce a homogeneous good that combines tasks supplied by both worker types and there is free-entry. The condition needed for a reduction in the measure of tasks to have a positive effect on production and wages is that an expansion in production along the intensive margin (through learning by doing) must outweigh a contraction along the extensive margin (through variety).

Cities where production involves a narrower set of tasks (higher specialization-driven productivity) attract a larger number of workers: this triggers a second round improvement in worker productivity through a standard division of labor argument, which further reinforces in-migration. I allow for other competing explanations for productivity differences across cities and workers, such as market externalities on local employment and generic labor demand shocks as in Bartik (1991).

The model is estimated via indirect inference. The estimation is based on a set of moment conditions: the model counterparts are constructed by solving for two distinct cross-sections, where the 2000 cross-section is fed with the Bartik shocks measured in the data, and subsequently employing a few statistics that take into account the strength of comovement between wage/employment growth and specialization growth.

The model captures well a number of features in the data, both in first-differences and levels. Its performance measured by its ability to match wage variation is always better compared to employment variation, regardless of the variable transformation, year or education groups: this is a consequence of the fact that I voluntarily abstract from supply-related determinants of location choice, such as variation in housing prices or amenities.

The estimated parameters are used to gauge the impact of specialization growth on workers location choice and wage growth and construct a counterfactual scenario where specialization in 2000 is held constant at the 1980 level. I find that specialization growth has been an important force behind the rising concentration of college workers in 1980 skill-abundant cities. The high degree of cross-city dispersion in college workers specialization growth amplifies this dynamic. In the absence of specialization growth, a less adverse relative supply channel compensates for the negative impact on college wage growth.

If there had been no specialization growth between 1980 and 2000, non-college employment growth would have been higher in cities where both types of workers actually benefited from an increase in specialization, since lack of in-migration of college workers tempered housing prices growth. Owing to the presence of a sizeable relative supply channel on wages, non-college workers choosing to locate in cities where actual specialization growth has been high would have experienced lower wages in the counterfactual scenario.

Finally, average specialization growth for non-college workers has been several orders of magnitude higher than for college workers: hence, if specialization growth did not prove successful at counteracting regional divergence, at least it attenuated the rise of the college premium between cities in 1980-2000. This finding also echoes the intuition behind the smithian pin-factory example that specialization might be a more suitable driver of productivity among manual, low-skilled workers who do not leverage much on interactive or creative skills.

Related literature This paper is related to the urban and labor economics literature. A few notable old papers dealing with the issue of specialization in the context of the labor economics literature are Rosen (1983), who highlights the importance of the cost of skill acquisition, Baumgardner (1988), who analyzes a spatial economy in which specialization depends on the presence of cooperation, Kim (1989), who deals with an assignment problem with specialization of firms and workers, Becker and Murphy (1992), who stress the connection between specialization and coordination costs/accumulation of knowledge.

Duranton (1998) and Becker and Henderson (2000) consider the division of labor as a mechanism driving local agglomeration economies. Duranton and Jayet (2011) provide an empirical assessment of the division of labor among cities using French data. They find that specialist occupations are overrepresented in large cities, where rarer occupations are defined as "specialist". Kok (2013) is closely related to this paper: she exploits a unique German database asking respondents to provide information on the tasks and, respectively, the intensity performed at work, and finds that workers employed in the same occupation tend to be more specialized in larger cities.

The choice to model the returns to specialization by means of an increasing returns to scale function of task production in the time devoted to each task is reminiscent of the work in Erosa, Fuster, Kambourov and Rogerson (2017), who use this assumption to help replicating the negative correlation between the log of mean annual hours in an occupation and the standard deviation of log annual hours within that occupation. Alternative models of learning by doing typically assume that the productivity premium matures over time, e.g., de la Roca and Puga (2017). Importantly, common to all these models is the notion that workers become more productive as the time they spend exposed to the production of a certain activity increases.

This work complements an increasing body of literature that has turned to the task-content of occupations to understand recent developments of the labor market, such as the increase in the wage premium (Autor, Levy and Murnane (2003)), or the wage polarization (Autor and Dorn (2013)).

Finally, this research brings a contribution to the theoretical urban economics literature researching the causes of differences in the wage distribution across cities (e.g., Behrens et al. (2015), Davis and Dingel (2014), Eeckhout, Pinheiro and Schmidheiny (2014)).

The rest of the paper is structured as follows: Section 2 describes the data and the stylized facts, Section 3 outlines the model, Section 4 discusses the estimation methodology and presents the results, Section 5 concludes.

3.2 Empirical Analysis

3.2.1 Data Description

The empirical analysis relies on two sources of data. The first one are 5% samples of U.S. Censuses (1980, 1990, 2000) from Integrated Public Use Microdata Series (IPUMS) (Ruggles et al. (2010)). These data provide individual-level information on a wide range of economic and demographic variables, including wages, housing costs, and geographic location of residence. All analysis is restricted to 25-55 year-old workers working at least 35 hours per week and 52 weeks per year. The geographical unit of analysis is the metropolitan statistical area (MSA) of residence. Individuals residing in rural areas are excluded from the sample.[2] Individuals are classified into low (less than college degree) or high-skilled (college degree or more).

The second dataset employed in this analysis is occupation-level information contained in the O*NET (18.1 release). I proceed according to the following steps. First, I match (1-to-many) the SOC 2010 taxonomy adopted by the 2010 American Community Survey (474 occupations) into O*NET-SOC codes (974 occupations). I assign to each SOC 2010 occupational title the simple average of specialization across the corresponding O*NET-SOC codes. Next, I map SOC 2010 occupational codes into the OCC 2005 classification. Finally, I match the OCC 2005 classification with the time-consistent occupation scheme developed by Dorn (2009) that provides a balanced panel of occupations spanning the years 1980-2000.

Being an expert-based classification system, the information contained in the O*NET has the advantage of not being subject to self-reporting biases, unlike specialization measures derived from surveys, e.g., the BIBB database employed in Spitz-Oener (2006). On the other hand, survey-based measures of specialization have the advantage of allowing to gauge specialization at the worker-level, thus providing an additional layer of variation. While surveys of this kind would thus provide an interesting cross-check, they are unavailable for the U.S.

The first occupation-level information recorded from the O*NET is from the *Task Statements* file that contains information on title, ID, and type for each of the tasks associated with a specific occupational title, as exemplified in Table 3.1.

[2] The alternative way to include workers residing in rural areas in the analysis is by treating them as if they lived all together in a fictitious city.

Intuitively an occupation should be less specialized if it involves a larger set of tasks. Not all tasks, however, are performed with the same frequency. Take the example of an accountant: he performs a wide variety of tasks but only a narrow subset is carried out on a regular basis, since many tasks, e.g., filing taxes or budget planning, take place only once in a year. Accounting for the frequency of tasks is important: if we did not account for it, we would wrongly infer that an accountant is a very poorly specialized occupation, while in reality her everyday work is really devoted to a much more limited set of tasks that can, therefore, be dealt with more effectively. For this reason, I include a second piece of information derived from the O*NET on the average frequency with which each task is performed. This information is contained in the *Task Ratings* file. In this file each task is featured alongside with a distribution for task intensity ranging from 1 to 7: the probability of each value corresponds to the fraction of respondents having selected it as the appropriate value to describe the degree of recurrence of a given task. The legend for the different scores is: 1 = "Yearly or less", 2 = "More than yearly", 3 = "More than monthly", 4 = "More than weekly", 5 = "Daily", 6 = "Several times a day", 7 = "Hourly or more".

Specialization for occupation o should intuitively be inversely proportional to the sum of tasks, where tasks characterized by higher frequency receive more weight. I calculate the sum of frequency-adjusted tasks for occupation o as $\sum_{t\in\mathcal{I}_o}\omega_t$, where $\mathcal{I}_o$ corresponds to the set of tasks characterizing occupation o and ω_t is the weighted average frequency of task t.[3] Fig.3.3 reports the empirical distribution of the sum of frequency-adjusted tasks. Next, specialization for occupation o is defined as:

$$\text{specialization}_o = \frac{\max_{o'\in O}\left\{\sum_{t\in\mathcal{I}'_o}\omega_t\right\}}{\sum_{t\in\mathcal{I}_o}\omega_t} \tag{3.1}$$

where the numerator corresponds to the highest sum of frequency-adjusted tasks recorded among occupations in the occupation set O. Table 3.2 provides an example of both sets of information - list of tasks and corresponding frequency - for the "economist" occupation.

In Table 3.3 are reported a set of statistics for each of the 23 major occupational groups characterizing the OCC 2005 classification scheme: the within-group average and standard deviation for the number of tasks, both the plain count and the frequency-adjusted count; the within-group average log deviation from the av-

[3]As a robustness check, I also convert the 1/7 score into the following weights: 1/365 = "Yearly or less", .5(1/365+12/365) = "More than yearly", .5(12/365+52/365) = "More than monthly", .5(52/365+1) = "More than weekly", 1 = "Daily", 4 = "Several times a day", 8 = "Hourly". The evidence on the specialization patterns across cities and education groups is robust to this alternative weighting scheme.

erage occupation, both in terms of the number of tasks and the wage.[4] Starting from the average number of tasks by major occupational groups, there is a substantial amount of variation, with Legal occupations standing at the lower end of the spectrum with an average of 16 tasks (22% less than the average occupation), and Education, Training and Library occupations occupying the upper end with an average of 31 tasks (44% more than the average occupation across all groups). In terms of within-group heterogeneity, Production occupations feature the highest degree, while Community and Social Science occupations the lowest one. Moving to the frequency-adjusted count of tasks, Education, Training and Library occupations keep being the group featuring the lowest degree of specialization. Due to the frequency-adjustment, however, the group featuring the highest degree of specialization is now Computer and Mathematical occupations. Most groups characterized by higher (lower) than average number of tasks (both in terms of plain count and the frequency-adjusted count) also feature lower (higher) wages: this evidence points in favor of a positive connection between the wage received by a worker and her degree of specialization, which will be dealt with in greater detail in the next subsection.

Finally, there are several other prominent occupational characteristics that have been looked at especially in the literature on wage polarization. For the sake of comparability and to show that the specialization measure defined in this paper does not overlap with those measures, I report in Table 3.4 the regression coefficients obtained from regressing the log of occupational specialization on the routine, abstract, manual and offshorable content of occupations, as identified in Autor and Dorn (2013). Specialized occupations tend to be more abstract and to have higher offshorable content: however, the Person pairwise correlation coefficient is 33% and 3% respectively, which shows that, while there is some overlap, especially with the information contained in the abstract task content, the correlation is far from perfect. Hence, the newly developed occupational characteristic described in this paper captures a few features of the labor market that have so far been unexplored.

3.2.2 Cross-sectional Stylized Facts

Specialization Patterns by City and Education Group

I use IPUMS weights to construct a measure of average specialization for workers living in a given Metropolitan Area and characterized by a certain educational status. High-skilled workers are workers with at least a college degree, and low-skilled workers are those that report less than a college degree. In Fig.3.1 the average specialization is plotted against the log size of the Metropolitan Area for

[4]Wage is measured by yearly pre-tax wage and salary income.

both high and low-skilled and three distinct decades. The following set of findings applied to each decade:

1. College graduates tend to be employed on average in more specialized occupations relative to non-college graduates.

2. Individuals working in large cities tend to be employed on average in more specialized occupations, regardless of educational attainment.

3. The rate at which average specialization rises in city size is remarkably higher for college graduates.

These facts are robust across years and no tendency is observed concerning the strengthening or weakening of any of the above reported stylized facts over time.

Next, I test the robustness of the specialization patterns just uncovered to industry composition. It is possible that these findings are being driven by a few industries hiring very specialized workforce that have a tendency to agglomerate in large cities. In Fig.3.2 the analysis portrayed in Fig.3.1 is repeated for eight distinct macro sectors: the patterns observed in Fig.3.2 confirm that industry composition is not behind the empirical findings. Average specialization is higher in larger cities and especially so for high-skilled workers in Manufacturing, Retail and Wholesale Trade, Utilities, Communications and Transportation, Finance and Professional Services. These three sectors together account for 61% of aggregate employment, and an even larger share if one excludes public sector jobs, such as Health and Education.

Specialization Wage Premium

Next, I investigate whether more specialized occupations earn a wage premium, and if so, in which local labor market they do so. First, I lump together occupations characterized by similar degree of specialization. Specifically, I construct 40 occupation bins corresponding to 40 percentiles of the specialization distribution. Subsequently, I regress the within occupation bin o, city j average wage on the log of city size, the log of the number (adjusted by frequency) tasks, and their interaction:

$$\ln \overline{w}_{jo} = \alpha + \beta_0 \ln \text{tasks}_o + \beta_1 \ln \text{size}_j^d + \beta_2 (\ln \text{tasks}_o \times \ln \text{size}_j^d) + e_{jo} \quad (3.2)$$

where size_j^d denotes demeaned city size.[5] The estimation output of eq.3.2 is

[5]This choice simplifies the interpretation of the estimated coefficients since now β_0 stands for the specialization premium earned in a mid-sized city and β_2 represents the incremental specialization premium obtained by moving to a larger city.

reported in the first column of Table 3.6 for Census 2000 data and in the first column of Table 3.5 for Census 1980 data.

In column 2 I add a set of controls based on demographic characteristics. Specifically, I run the following regression at the individual i level:

$$\ln w_i = \delta_0 + \delta_1 \text{Female}_i + \delta_2 \text{White}_i + \delta_3 \text{Experience}_i + \delta_4 \text{Experience}_i^2 + e_i \quad (3.3)$$

Next, I calculate $\hat{d}_{jo}$ as the average residual within each occupation bin/city cell. Finally, I run eq.3.2 on $\hat{d}_{jo}$ instead of the average log of the raw wage. The parameter estimates are reported in column 2 of Table 3.6 (Census 2000 data) and Table 3.5 (Census 1980 data).

In column 3 I add a set of industry fixed effects (2-digit) as an additional set of controls and run the following regression at the individual i level:

$$\ln w_i = \delta_0 + \delta_1 \text{Female}_i + \delta_2 \text{White}_i + \delta_3 \text{Experience}_i + \delta_4 \text{Experience}_i^2 + \gamma_{industry} + e_i \quad (3.4)$$

I calculate $\hat{d}_{jo}$ following the same logic as in the previous specification and run eq.3.2 on the newly estimated $\hat{d}_{jo}$. The parameter estimates are reported in column 3 of Table 3.6 (Census 2000 data) and Table 3.5 (Census 1980 data).

Moving from the least to the most restrictive specification all coefficients of interest - β_0 and β_2 - shrink in absolute value. The only coefficient making the exception is the incremental specialization premium for high-skilled workers, thus implying an understatement of the specialization premium for high-skilled workers living in larger cities when demographics and industry characteristics are not controlled for.

Overall, more specialized occupations (lower number of frequency-adjusted tasks) earn a wage premium and this increases in the size of the Metropolitan Area. This is true independently of the cross-section taken into consideration. While the high-to-low skilled specialization premium ratio for a mid-sized city does not show a clear-cut pattern, by being smaller than 1 in 1980 and larger than 1 in 2000, the high-to-low skilled *incremental* - i.e., city-size dependent - specialization premium ratio is around 3 in 1980, and 2 in 2000.

In order to understand whether the just uncovered specialization premium is driven by unobservable worker characteristics, I use the Survey on Income and Program Participation (SIPP) for years from 1990 to 1996 and after having applied the same sample restrictions as with IPUMS data. Next, I run the following specification:

$$\ln w_{ijot} = \delta X_{it} + \beta_0 \ln(\text{tasks}_{ot}) + \beta_1 \ln(\text{size}_{jt}) + \beta_2 (\ln(\text{tasks}_{ot}) \times \ln(\text{size}_{jt})) + \gamma_{industry} + \gamma_t + e_{it} \quad (3.5)$$

where i indexes the individual, j the city, o the occupation and t the year: X_{it} corresponds to a large set of individual controls, such as a dummy for being of white ethnicity, a dummy for being female, a dummy for being of age between 25 and 40/between 40 and 55, a quadratic in on-the-job experience and a quadratic in on-the-occupation experience. The estimated coefficients are reported in Table 3.7. In column 1 I exclude individual-level fixed effects: while slightly smaller in absolute value than the estimates obtained in the IPUMS sample, both coefficients of interest are negative and statistically significant. When I add individual-level fixed effects to fully leverage the longitudinal dimension of the panel, the relevant coefficients are still statistically significant, although they are now halved in absolute value.

3.2.3 Divergence in Location Choices over 1980-2000

To what extent has the rise or decline in the average specialization of the workforce in a given city contributed to faster or slower local employment growth? A body of literature has analyzed the determinants of city growth in the U.S. between 1980 and 2000.[6] and found that cities that were abundant with college workers in 1980 further increased their share of college workers over the subsequent decades.

The divergence in the location choices of college vs. non-college workers has been accompanied by an increase in the college premium in cities that were becoming increasingly more skill-abundant, thus pointing in the direction of productivity shifts as the likely driver of the changing geography of jobs. Urban economists, in particular, have taken this pattern as evidence of the existence of human capital spillovers: places that by some luck had been initially able to attract more college relative to non-college workers managed to further increase their share of college workers since a high college share generated productivity gains that disproportionately benefited college workers.

However, higher average income in skill-abundant cities has been coming along with higher housing prices: the increase in real wage inequality associated with increasing spatial segregation by education group has been therefore much more subdued than the increase in nominal wage inequality (Moretti (2013)).[7]

In this paper, specialization acts as an amplification mechanism for diverging productivity growth across cities and types. As it will become clearer in the model section, specialization is assumed to be the byproduct of two elements: on the

[6] See Moretti (2013) for a review.

[7] A more recent strand of literature has emphasized the importance of skill-specific agglomeration externalities on the consumption side (see Diamond (2016) and Handbury (2013)). According to these studies college workers have concentrated in initially skill-abundant cities because of a higher provision of consumption-related amenities (ranging from the availability of leisure activities to higher green areas density, to a wider variety of local goods).

one hand, a smaller required set of tasks for firm-level production raises average individual specialization within the firm; on the other, a broader workforce at the firm-level also raises average individual specialization for a given set of tasks. Conditional on firm size increasing in city size, average individual specialization is higher in larger cities for a given set of tasks employed in production.

When the set of tasks changes, there are two effects on production: on the one hand, a narrower set of tasks raises specialization and therefore productivity in a given task; on the other, it reduces firm-level production by a variety effect. In the model section, the conditions for the first effect to be dominating are described. Since the positive effect along the intensive margin dominates the negative effect along the extensive one, a *reduction* in the number of tasks is interpreted as an exogenous *increase* in productivity: I refer to this component of worker productivity as *specialization-driven*.

Specialization-driven productivity may have changed across cities and skill types between 1980 and 2000 for several reasons. An increased division of labor within the firm can be considered as the consequence of a more efficient management style concerning the assignment of tasks to workers. Firms more exposed to international competition might have been compelled to update their internal organization, thus seeking a more efficient use of their human capital. By the existence of technology spillovers firms located nearby would have then adopted the newer more efficient production technology.

Shocks to specialization-driven productivity are amplified by the existence of agglomeration externalities. Cities where wages become more attractive following a shift in the production technology experience faster growth: this fosters a second round of improvement in the division of labor, thanks to an expansion in the workforce at the firm-level for a given set of tasks that generates a further increase in worker specialization and wages.

Are shocks to specialization-driven productivity entirely driven by different exposure across cities to industry wide shocks? Table 3.8 provides a negative answer. I construct a popular measure of city-level shock to labor demand that hinges on the local industry mix and the evolution of industry groups at the national level, i.e., the Bartik shock (Bartik (1991)). Specifically:

$$Bartik_{H,j} = \sum_k \frac{L_{H,j,k,1980}}{L_{H,j,1980}} \Delta w_{H,-j,k,1980-2000} \qquad Bartik_{L,j} = \sum_k \frac{L_{L,j,k,1980}}{L_{L,j,1980}} \Delta w_{L,-j,k,1980-2000} \tag{3.6}$$

In column 1 and 4 I regress wage growth across cities and education groups between 1980 and 2000 on college share growth: human capital externalities as proxied by college share growth are an important determinant of wage growth especially for college workers, as the nearly identical coefficient for college and non-college wage growth shows. Industry-level aggregate changes in productivity as proxied by the Bartik shock also play an important role, and the R-squared rises significantly for both education groups once it is included. The evidence presented in column 3 and 6 allows us to confirm that shocks to specialization-driven productivity are not entirely explained by industry-level shifts in labor demand: this finding reinforces what already discussed in the previous section (see Fig.3.2), namely that differences in worker specialization across cities are not driven by the industry-mix. While the Bartik shock by construction captures the contribution to local wage growth of specific industries that have risen (or declined) during the observation period, the specialization shock operates *across industries*.

Finally, I find that while college employment grew in cities that saw an improvement in average college worker specialization over this period, non-college employment actually declined in local labor markets where specialization grew the most for non-college workers. Specifically, in Table 3.9 I regress employment growth on the Bartik shock and specialization growth. The cities where specialization has grown the most for both education groups were initially very human capital abundant: hence, by the disproportionate impact of human capital externalities on the productivity and wages of college workers, housing costs went up in these cities thus discouraging non-college workers from migrating to such locations.[8] Furthermore, the inclusion of specialization growth raises significantly the explanatory power of the set of regressors for non-college workers, as captured by the R-squared.

3.3 Model

The stylized facts presented in the previous section suggest that the evolution of worker specialization over time across cities and worker types provides an explanation for rise and fall of U.S. cities and the divergence in location choices of college workers relative to non-college ones. The aim of the next section is to develop a structural model of location choice that can be taken to the data and used to quantify the extent to which specialization-driven productivity shocks alone accounted for the observed location choices of workers in 1980-2000. The only non-standard part of the model is the labor demand side of the economy: workers get hired by a measure of firms producing the same homogenous good and using an exogenous measure of tasks as inputs in production.

[8] The correlation between specialization growth for college and non-college workers across cities is 25%.

3.3.1 Description

Workers An individual z characterized by education group or skill $i \in \{H, L\}$ living in city $j \in J$, where J designates the set of locations available, solves the utility maximization problem:

$$\begin{array}{c} \max_{C_{ij},H_{ij}} U_{ij}(z) = A_j + \ln\left(C_{ij}^{1-\alpha}\right) + \ln\left(H_{ij}^{\alpha}\right) + \epsilon_{ij}^{z} \\ s.t. \quad W_{ij} \geq R_j H_{ij} + C_{ij} \end{array} \tag{3.7}$$

where C_{ij} corresponds to consumption of a national good having price normalized to 1, H_{ij} indicates consumption of a local good, e.g., housing, with price R_j, and W_{ij} is the wage received in exchange for labor supplied by individuals to firms inelastically. Individuals derive utility also from local amenities A_j: in addition to it, each individual experiences an idiosyncratic taste shock, ϵ_{ij}^{z}.

With indirect utility of living in city j given by $V_{ij}(z) = \ln\left(\alpha^{\alpha}(1-\alpha)^{1-\alpha}\right) + a_{ij} + w_{ij} - \alpha r_j + \beta\epsilon_{ij}^{z}$, individual z prefers location j to location j' if:

$$(a_j + w_{ij} - \alpha r_j) - (a_{j'} + w_{ij'} - \alpha r_{j'}) > -\left(\epsilon_{ij}^{z} - \epsilon_{ij'}^{z}\right) \tag{3.8}$$

It is assumed that ϵ_{ij}^{z} are independent and identically distributed random variables following an Extreme Value Type I distribution with location parameter $\mu = 0$ and scale parameter β, such that $F\left(\epsilon_{ij}^{z} \leq x\right) = \exp\left(-\exp\left(-x/\beta\right)\right)$. By the distributional properties of the taste shock the fraction of individuals of a certain type locating in city j is given by:

$$\frac{L_{H,j}}{L_H} = \frac{\exp\left(\frac{1}{\beta}a_{ij} + \frac{1}{\beta}w_{H,j} - \frac{\alpha}{\beta}r_j\right)}{\sum_{j'}\exp\left(\frac{1}{\beta}a_{ij'} + \frac{1}{\beta}w_{H,j'} - \frac{\alpha}{\beta}r_{j'}\right)}; \quad \frac{L_{L,j}}{L^L} = \frac{\exp\left(\frac{1}{\beta}a_{ij} + \frac{1}{\beta}w_{L,j} - \frac{\alpha}{\beta}r_j\right)}{\sum_{j'}\exp\left(\frac{1}{\beta}a_{ij'} + \frac{1}{\beta}w_{L,j'} - \frac{\alpha}{\beta}r_{j'}\right)} \tag{3.9}$$

Local amenities are assumed to be exogenous, $A_j = \exp\left(\sigma^{\eta}\eta_j\right)$ with $\eta_j \sim N(0,1)$.

Firms The production structure is a slightly more elaborate version of the specialization model presented in Duranton and Puga (2003), which has been modified to allow for multiple skill types and free-entry of firms.

The national good is homogenous and sold by an endogenous mass N_j of firms indexed by n. Firms production results from the aggregation of a fixed and exogenous measure τ_{ij} of symmetrically differentiated tasks performed by different worker types, with elasticity of substitution among tasks performed by the same worker type set to $1/(1-\varepsilon)$. Worker types are endowed with efficiency units Z_{ij}.

Firms choose how much labor $L_{ij}(n)$ to hire taking as given the wages according to the maximization problem:

$$\begin{array}{ll} \max_{L_{ij}(n)\forall i\in\{H,L\}} \Pi_j(n) = Y_j(n) - \sum_i W_{ij} L_{ij}(n) \\ s.t. \quad Y_j(n) = \sum_i Z_{ij} \left(\int_0^{\tau_{ij}} X_{ij}^{\varepsilon}(t)dt\right)^{\frac{\rho}{\varepsilon}}; \quad \varepsilon, \rho \in (0,1) \end{array} \tag{3.10}$$

A task $X_{ij}(t)$ with $t \in [0, \tau_{ij}]$ is produced according to the technology:

$$X_{ij}(t) = h_{ij}^{\frac{1}{\theta}}; \quad \theta \in (0,1) \tag{3.11}$$

where h_{ij} indicates the number of hours employed in the production of one task, and θ captures the strength of increasing returns to scale in the production of one task. Intuitively the gain from working an additional hour on one task is greater when you have already spent sufficient time working on it: in other words, the assumption of $\theta \in (0,1)$ is a synonim for learning by doing on the workplace.

None of the tasks is performed in equilibrium by more than one worker: whenever two workers want to perform the same task, they become Bertrand competitors and get zero revenues from the execution of the task. On the other hand, they obtain compensation equal to $P_{ij}(t)X_{ij}(t) = \rho Z_{ij} \left(\int_0^{\tau_{ij}} X_{ij}^{\varepsilon}(t)dt\right)^{\frac{\rho}{\varepsilon}-1} X_{ij}^{\varepsilon}(t)$ when they perform an even number of tasks.

Each worker is assumed to have time endowment $\overline{h} = 1$, so that by the symmetry of tasks:

$$h_{ij} = \frac{L_{ij}(n)}{\tau_{ij}} \Leftrightarrow Y_j(n) = \sum_i Z_{ij} \tau_{ij}^{\nu-\phi} L_{ij}(n)^{\phi} \tag{3.12}$$

with $\rho/\theta = \phi$ and $\rho/\varepsilon = \nu$.

Firm labor demand and profits are respectively:

$$L_{ij}(n) = \phi^{\frac{1}{1-\phi}} \left(Z_{ij}\tau_{ij}^{\nu-\phi}\right)^{\frac{1}{1-\phi}} W_{ij}^{\frac{1}{\phi-1}} \tag{3.13}$$

$$\Pi_j(n) = \left(\phi^{\frac{\phi}{1-\phi}} - \phi^{\frac{1}{1-\phi}}\right) \sum_i \left(Z_{ij}\tau_{ij}^{\nu-\phi}\right)^{\frac{1}{1-\phi}} W_{ij}^{\frac{\phi}{\phi-1}} \tag{3.14}$$

Conditional on $\nu < \phi < 1$ ($\theta < \varepsilon < \rho$), optimal firm size is a decreasing function of the wage and the measure of tasks needed to produce. The interpretation for this condition is that output must increase more with an expansion along the intensive margin as opposed to an expansion along the extensive margin, given the same amount of resources employed: having to choose whether to allocate an additional hour of working time to one of the existing tasks as opposed to the execution of a new task, it is more profitable to invest this time on an existing task.

The equilibrium is closed by a labor market clearing for each worker type and the free-entry condition for firms. Each firm must in fact pay a sunk cost κ

in terms of capital. The unit capital requirement is location invariant and, being capital perfectly mobile, so it is for the cost of capital.

$$L_{ij} = N_j L_{ij}(n) \quad i \in \{H, L\} \tag{3.15}$$

$$\Pi_j(n) = \kappa \tag{3.16}$$

Using eq.3.15 and eq.3.13, the wage for worker type i located in city j is:

$$W_{ij} = \phi Z_{ij} \tau_{ij}^{\nu-\phi} L_{ij}^{\phi-1} N_j^{1-\phi} \tag{3.17}$$

Using eq.3.17 and eq.3.14, individual firm profits are:

$$\Pi_j(n) = (1-\phi) \left(\sum_i Z_{ij} \tau_{ij}^{\nu-\phi} L_{ij}^{\phi} \right) N_j^{-\phi} \tag{3.18}$$

Using eq.3.18 and eq.3.16, the equilibrium number of firms is:

$$N_j = \left(\frac{1-\phi}{\kappa} \sum_i Z_{ij} \tau_{ij}^{\nu-\phi} L_{ij}^{\phi} \right)^{1/\phi} \tag{3.19}$$

Equilibrium specialization is given by the ratio between total labor demand per education group and the measure of tasks produced across N_j firms:

$$S_{ij} = \frac{L_{ij}}{\tau_{ij} N_j} = \frac{1}{\tau_{ij}} \left(\frac{L_{ij}^{\nu}}{\frac{1-\phi}{\kappa} \sum_{i'} Z_{i'j} S_{i'j}^{\phi-\nu} L_{i'j}^{\nu}} \right)^{1/\nu} \tag{3.20}$$

Using eq.3.20, eq.3.19 and eq.3.17, the equilibrium wage for each worker type as a function of specialization, labor demand and efficiency units is:

$$W_{ij} = \phi Z_{ij} S_{ij}^{\phi-\nu} L_{ij}^{\nu-1} \left(\frac{1-\phi}{\kappa} \sum_{i'} Z_{i'j} S_{i'j}^{\phi-\nu} L_{i'j}^{\nu} \right)^{(1-\nu)/\nu} \tag{3.21}$$

To give the model a greater degree of realism, I assume that efficiency units depend on city-level employment: in other words, wages depend on local employment also through market externalities apart from the traditional supply channel. Market externalities are allowed to differ based on the education group and take the log-linear form:

$$Z_{H,j} = \delta_H \left(L_{H,j}\right)^{\gamma_{HH}} \left(L_{L,j}\right)^{\gamma_{HL}} \exp\left(\sigma_H^{\varepsilon} \varepsilon_{H,j}\right); \quad Z_{L,j} = \delta_L \left(L_{H,j}\right)^{\gamma_{LH}} \left(L_{L,j}\right)^{\gamma_{LL}} \exp\left(\sigma_L^{\varepsilon} \varepsilon_{L,j}\right) \tag{3.22}$$

where $\begin{bmatrix} \varepsilon_{H,j} & \varepsilon_{L,j} \end{bmatrix} \sim N(0, I)$. [9]

Substituting in efficiency units into eq.3.21 and log-linearizing yields the log-linearized labor demand equations of the model:[10]

$$\begin{aligned} w_{H,j} &= \ln(\phi\delta_H) + (\phi-\nu)s_{H,j} + (\gamma_{HH}+\nu-1)\, l_{H,j} + \gamma_{HL} l_{L,j} + (1-\nu)\ln N_j + \sigma^{\varepsilon}_{H}\varepsilon_{H,j} \\ w_{L,j} &= \ln(\phi\delta_L) + (\phi-\nu)s_{L,j} + (\gamma_{LL}+\nu-1)\, l_{L,j} + \gamma_{LH} l_{H,j} + (1-\nu)\ln N_j + \sigma^{\varepsilon}_{L}\varepsilon_{L,j} \end{aligned} \tag{3.23}$$

with the number of firms given by:

$$N_j = \left(\frac{1-\phi}{\kappa}\right)^{1/\nu} \left(\sum_i Z_{ij} S_{ij}^{\phi-\nu} L_{ij}^{\nu}\right)^{1/\nu} \tag{3.24}$$

The log-linearized labor supply equations of the model are:

$$\begin{aligned} l_{H,j} &= c_H + \tfrac{1}{\beta}a_j + \tfrac{1}{\beta}w_{H,j} - \tfrac{\alpha}{\beta}\ln R_j + \sigma^{\eta}\eta_j \\ l_{L,j} &= c_L + \tfrac{1}{\beta}a_j + \tfrac{1}{\beta}w_{L,j} - \tfrac{\alpha}{\beta}\ln R_j + \sigma^{\eta}\eta_j \end{aligned} \tag{3.25}$$

with housing prices given by:

$$R_j = \left(\frac{\alpha\phi}{H}\right)\left(\frac{1-\phi}{\kappa}\right)^{(1-\nu)/\nu} \left(\sum_i Z_{ij} S_{ij}^{\phi-\nu} L_{ij}^{\nu}\right)^{1/\nu} \tag{3.26}$$

where H is the amount of common land supply available in each location. Together, eq.3.23, eq.3.25, eq.3.24 and eq.3.26 fully characterize the equilibrium of the model.

3.3.2 Equilibrium

Definition *The equilibrium for an economy characterized by exogenous city-skill specific productivities $\tau_{ij}^{\nu-\phi}\exp(\sigma^{\varepsilon}_{i}\varepsilon_{ij})$, exogenous city specific amenities*

[9] The assumption underlying eq.3.22 is that different types of workers might benefit asymmetrically from density. For instance, Diamond (2016) finds that college workers benefit from human capital spillovers more than non-college workers, where human capital spillovers are increasing in the relative mass of high to non-college workers.

[10] The Simulated Method of Moments algorithm chosen to structurally estimate the model parameters requires us to take a stance on the structural form of efficiency units in eq.3.22 to be able to construct simulated moments to be compared with the actual ones: this approach can be suboptimal and expose the model to misspecification. An alternative would be to estimate the model via simultaneous equation non-linear Generalized Method of Moments as in Diamond (2016): in this case no assumption on functional forms is required since the moments compared to the actual ones are obtained by evaluating the model at the data for given parameters combination, thus skipping the model solution step.

$\exp(\sigma^\eta \eta_j)$, *and structural parameters* $\Omega = \{\Gamma, \Delta, \Sigma, \phi, \nu, \alpha, \beta, \kappa, H\}$, *is a sequence of prices* $\{\mathbf{R}, \mathbf{W_H}, \mathbf{W_L}\}$*and factor allocations* $\{\mathbf{L_H}, \mathbf{L_L}\}$ *such that:*

- *the housing market clears (eq.3.26);*
- *individuals choose optimally where to live (eq.3.9);*
- *the labor market clears (eq.3.13);*
- *there is free-entry of firms (eq.3.16),*

where:

$$\Gamma = \begin{bmatrix} \gamma_{LL} & \gamma_{LH} \\ \gamma_{HL} & \gamma_{HH} \end{bmatrix} \quad \Delta = \begin{bmatrix} \delta_L & \delta_H \end{bmatrix} \quad \Sigma = \begin{bmatrix} \sigma^\varepsilon_L & \sigma^\varepsilon_H & \sigma^\eta \end{bmatrix}$$

The equilibrium does not need to be regular, i.e., such that all available locations are inhabited. In Appendix 3.5.1 I sketch the conditions for the existence of a non-degenerate equilibrium in a simplified setup where $Z_{H,j} = \delta_H L_j^\gamma$, $Z_{L.j} = \delta_L L_j^\gamma$ and all shocks $\varepsilon_{ij} = \eta_j = 0$: γ - the parameter governing the strength of market externalities - must be sufficiently small, while β (α) - the parameter governing the elasticity of labor supply to local conditions (fraction of expenditure on housing) - must be sufficiently high.

Intuitively, if agglomeration externalities in production are too strong only a site will be inhabited in equilibrium. Equivalently, if the elasticity of labor supply to wages is too high there will be an increasing number of workers that given higher wages in large cities wishes to locate there: given the existence of agglomeration economies, this puts upward pressure on wages that further feeds into labor supply. The equilibrium must consist once again of only one inhabited site. Similarly, if the fraction of income spent on housing is too small, rents are not high enough to compensate for higher wages in large cities, which translates, once again, into a degenerate spatial equilibrium.

3.4 Estimation

3.4.1 Methodology

I adopt a Simulated Method of Moments (SMM) approach and estimate the model via indirect inference (Gourieroux et al. (1993)). To be consistent with existing literature (Beaudry et al., (2012), (2017)) the estimation is based on first-differenced outcome variables. In our model changes in the allocation of workers across cities between 1980 and 2000 are due to three underlying forces:

- Specialization-driven labor demand shocks. These are proxied by the specialization measure described in Section 3.2.
- Other labor demand shocks. These are proxied by the Bartik shock as follows:

$$e_{i,j,1980} = \sigma_i^{\varepsilon}\varepsilon_{i,j,1980} \quad e_{i,j,2000} = e_{i,j,1980} + Bartik_{i,j} \quad \forall i \in \{L, H\} \tag{3.27}$$

  with $\varepsilon_{i,j,1980} \sim N(0,1)\ \forall i$.

- Labor supply shocks $a_{j,t}$. These are i.i.d. across cities and distributed according to $a_{j,t} = \sigma^{\eta}\eta_{j,t}$, with $\eta_{j,t} \sim N(0,1)\ \forall t$.

The setup features a shock structure $\mathbf{e} = \left[\begin{array}{cccc} \varepsilon_{H,1980} & \varepsilon_{L,1980} & \eta_{1980} & \eta_{2000} \end{array}\right]$ and a set of exogenous variables $\widetilde{\mathbf{x}} = \left[\begin{array}{cccccc} \mathbf{s_{H,1980}} & \mathbf{s_{H,2000}} & \mathbf{s_{L,1980}} & \mathbf{s_{L,2000}} & \mathbf{Bartik_H} & \mathbf{Bartik_L} \end{array}\right]$, where the lower case letter denotes the log of the variable corresponding upper case letter. A SMM-based estimation routine consists of simulating data from a structural model $\mathbf{y}\left(\Omega, \widetilde{\mathbf{x}}, \mathbf{e}\right)$, computing the implied moments $m\left(\mathbf{y}\left(\Omega, \widetilde{\mathbf{x}}, \mathbf{e}\right)\right)$, and determining the set of structural parameters Ω that minimize an objective function of the type:

$$\min_{\Omega} \left[m\left(\mathbf{y}(\Omega, \widetilde{\mathbf{x}}, \mathbf{e}\right) - m^{data}\right]' W \left[m\left(\mathbf{y}(\Omega, \widetilde{\mathbf{x}}, \mathbf{e}\right) - m^{data}\right]$$

Specifically, I minimize:

$$\min\nolimits_{\Omega} \mathbf{m}(\Omega, \widetilde{\mathbf{x}}, \mathbf{e})' W \mathbf{m}(\Omega, \widetilde{\mathbf{x}}, \mathbf{e}) \quad \vee \quad \mathbf{m}(\Omega, \widetilde{\mathbf{x}}, \mathbf{e}) = \begin{bmatrix} \mathbf{M}_1' \\ \mathbf{M}_2' \\ \mathbf{M}_3' \end{bmatrix} \tag{3.28}$$

where:

$$\begin{array}{r} \mathbf{M}_{1,(1\times4)} = \frac{1}{I}\sum_{i=1}^{I}\frac{1}{J}\iota'\mathbf{y}_i^m - \frac{1}{J}\iota'\mathbf{y}_j^d \\ \mathbf{M}_{2,(1\times4)} = \text{trace}\left[\frac{1}{I}\sum_{i=1}^{I}\left(\frac{1}{J}\left(\mathbf{y}_i^m - \overline{\mathbf{y}}_i^m\right)'\left(\mathbf{y}_i^m - \overline{\mathbf{y}}_i^m\right)\right) - \frac{1}{J}\left(\mathbf{y}^d - \overline{\mathbf{y}}^d\right)'\left(\mathbf{y}^d - \overline{\mathbf{y}}^d\right)\right] \\ \mathbf{M}_{3,(1\times8)} = \text{vec}\left[\frac{1}{I}\sum_{i=1}^{I}\left(\frac{1}{J}\left(\mathbf{y}_i^m - \overline{\mathbf{y}}_i^m\right)'\left(\mathbf{x} - \overline{\mathbf{x}}\right)\right) - \frac{1}{J}\left(\mathbf{y}^d - \overline{\mathbf{y}}^d\right)'\left(\mathbf{x} - \overline{\mathbf{x}}\right)\right] \end{array} \tag{3.29}$$

and:

$$\mathbf{y}_{(J\times4)} = \left[\widehat{\mathbf{w}}_H \quad \widehat{\mathbf{w}}_L \quad \widehat{\mathbf{l}}_H \quad \widehat{\mathbf{l}}_L\right] \qquad \mathbf{x}_{(J\times2)} = \left[\widehat{\mathbf{s}}_H \quad \widehat{\mathbf{s}}_L\right] \tag{3.30}$$

i.e., respectively, the set of first-differenced outcome variables and a restricted set of first-differenced exogenous variables, where the hat notation stands for the 1980-2000 growth rate.

The first set of moments $\mathbf{M_1}$ corresponds to the average wage/employment growth rate per each education group group across cities: the model counterpart is obtained via averaging across I simulations, with $I = 500$. The second set of moments $\mathbf{M_2}$ measures the variance of of the wage/employment growth rate per each education group across cities. Finally, the third set of moments $\mathbf{M_3}$ lists the covariances of each outcome variable in $\mathbf{y}$ with each exogenous variable listed in $\mathbf{x}$, both taken in first-differences.

Two parameters are borrowed externally: α is set equal to .24 as the median estimate obtained in Davis and Ortalo-Magné (2007), and β equal to .33 as the average estimate across education groups obtained in Diamond (2016).

Three sets of parameters are calibrate internally. I set the housing supply, H_t, and the city-independent components of worker productivity, $\delta_{i,t}$, in order to match exactly the average rent, and the average wage by skill group across cities in both periods, respectively.[11] Finally, I set the sunk cost of entry, κ, in order to match exactly the average number of establishments across cities in 2000.[12]

The final set of parameters to be estimated is:

$$\Omega = \{\phi, \nu, \gamma_{HH}, \gamma_{HL}, \gamma_{LL}, \gamma_{LH}, \sigma_H^{\varepsilon}, \sigma_L^{\varepsilon}, \sigma^{\eta}\} \tag{3.31}$$

Being the model overidentified I use a weighting matrix W corresponding to the inverse of the variance-covariance matrix of bootstrapped empirical moments. The exact estimation routine followed is described in greater detail in Appendix 3.5.2.

Before presenting the estimation results, it is useful to understand how the structural parameters shape the targeted moments, so as to get a hint of what are the restrictions that the estimated parameters must satisfy in order to replicate the patterns observed in the data. I describe the intuition with the help of a simplified 1-skill setup. In this case, the analytical expression for log rents, log labor supply and the inverse of log labor demand is:

$$\begin{aligned} r_j &= \rho_R + w_j + l_j \\ l_j &= \rho_L + \tfrac{1}{\beta} w_j - \tfrac{\alpha}{\beta} r_j + \sigma^{\eta}\eta_j \\ w_j &= \rho_W + \left(\tfrac{\phi-\nu}{\nu}\right) s_j + \left(\tfrac{\gamma}{\nu}\right) l_j + \sigma^{\varepsilon}\varepsilon_j \end{aligned} \tag{3.32}$$

with $\rho_R = \ln\left(\frac{\alpha}{H}\right)$, $\rho_L = c(\mathbf{L})$, $\rho_W = \ln(\phi\delta) + ((1-\nu)/\nu)\ln((1-\phi)/\kappa)$.

Substituting in log rents, the equilibrium is summarized by a $2J \times 2J$ system:

$$\begin{aligned} \text{Supply:} \quad l_j &= \tfrac{\beta\rho_L - \alpha\rho_R}{\alpha+\beta} + \tfrac{1-\alpha}{\alpha+\beta} w_j + \tfrac{\beta}{\alpha+\beta}\sigma^{\eta}\eta_j \\ \text{Demand:} \quad w_j &= \rho_W + \left(\tfrac{\phi-\nu}{\nu}\right) s_j + \left(\tfrac{\gamma}{\nu}\right) l_j + \sigma^{\varepsilon}\varepsilon_j \end{aligned} \tag{3.33}$$

[11] The structural interpretation of the change in these parameters across cross-sections is undermined by the inflation component that is built into them.

[12] For this information I rely on information provided in the County Business Patterns database (2000), which delivers an average number of establishments in 2000 $N = 18000$.

Notice that the labor supply elasticity to the wage is always decreasing in α and β, while the wage elasticity to labor demand is increasing in γ, the parameter governing the strength of agglomeration economies, and decreasing in ν: the latter stems from the fact that when ν is high, firm entry is weaker and this exerts downward pressure on wages. Substituting the supply into the demand equation, wages can be rewritten as a function of the exogenous variables only:

$$w_j\left[\frac{\nu(\alpha+\beta)-\gamma(1-\alpha)}{\nu(\alpha+\beta)}\right]=\left[\rho_W+\left(\frac{\gamma}{\nu}\right)\left(\frac{\beta\rho_L-\alpha\rho_R}{\alpha+\beta}\right)\right]+\left(\frac{\phi-\nu}{\nu}\right)s_j+\left(\frac{\gamma}{\nu}\right)\left(\frac{\beta}{\alpha+\beta}\right)\sigma^{\eta}\eta_j+\sigma^{\varepsilon}\varepsilon_j \tag{3.34}$$

According to eq.3.34 the structural parameters should be such that $\frac{\nu}{\gamma}>\frac{1-\alpha}{\alpha+\beta}$ in order for the model to replicate the features observed in the data, i.e., $Cov(\widehat{w}_{ij},\widehat{s}_{ij})>0$ and $Cov(\widehat{l}_{ij},\widehat{s}_{ij})>0$, as seen in Section 3.2. The economic interpretation for this condition is that the slope of labor supply with respect to the wage must be less steep than the one for the labor demand in a $\{w,L\}$ diagram. If the opposite were true, an increase in specialization would lead to a direct increase in wages and, through the labor supply feedback, to an indirect self-reinforcing increase in wages via agglomeration externalities, so that the path would be explosive. It follows that when the inequality is reversed, wages must depend negatively on specialization, which is the opposite of what is seen in the data.

3.4.2 Results

Parameter estimates

The estimated parameters are reported in Table 3.11. I estimate a negative elasticity of wages with respect to the measure of tasks employed in production at the individual firm-level $\nu-\phi=-.48$: a one percent decline in the measure of tasks leads to a .48% increase in wages. To put things into perspective, consider the definition of specialization provided in eq.3.20: given information on individual specialization and local employment, I calculate an average growth rate in $\tau_{ij}N_j$ between 1980 and 2000 of -41% and -78% for non-college and college workers, respectively. Assuming a yearly pace of establishment creation in the order of .5%, I infer that the measure of tasks must have declined between 1980 and 2000 by -51% and - 88%, respectively. At an annual inflation rate of 2% real wages have increased by 30% and 39% for non-college and college workers, respectively: an elasticity of -.48% therefore suggests that gains in specialization-driven productivity accounted for the totality of real wage growth registered on average across U.S. cities between 1980 and 2000 for workers with a college degree, and approximately 80% of real wage growth for workers without a college degree.

I estimate an elasticity of labor substitution equal to $1/(1-\nu) = 1.69$ and between the estimates of Card (2009) of 2.5 and Diamond (2016) of 1.6.

I do not find evidence of human capital spillovers as in Moretti (2004): the wage elasticity is -.81 for college workers and -.91 for non-college workers, thus requiring a market externality in the order of -.7 and -.8, respectively, on the same worker type local employment. Diamond (2016) estimates a statistically significant non-college wage elasticity with respect to non-college employment of -.55 - close to my estimated value of -.7 - and a not statistically significant college wage elasticity with respect to college employment of .22. She also estimates a statistically significant non-college wage elasticity with respect to college employment of .69, and a non statistically significant college wage elasticity with respect to non-college employment of .31 - close to my estimated value of .25. Hence, she finds evidence in favor of human capital spillovers on non-college workers only, which I do not find evidence for in the current setup.

The results in Diamond (2016) are in contrast with recent work in Beaudry et al. (2017). In their research, the authors estimate the wage elasticity to be -1 for non-college workers and - .75 for college workers. These numbers are not directly comparable with mine since they are based on city/industry-level variation. The authors' main contribution is however to show that the wage elasticity to local employment is much more negative when it is based on city-level variation compared to city/industry-level variation, and they ascribe the difference to the relevance of congestion externalities that are arise in the search literature. An increase in employment by one firm represents a negative externality on other firms via local tightness: by an increase in the market tightness the job filling probability declines, thus depressing firm creation and putting downward pressure on wages.

While my results seem to confirm the presence of congestion externalities as in Beaudry et al. (2017), I refrain from pushing further the interpretation of the market externalities estimates obtained in this work for three reasons. First, my interest is on the value of ν and ϕ, more than on the parameters governing market externalities, which were introduced with the objective of giving an otherwise very stylized model a better chance of matching the data and producing therefore realistic estimates for the parameters of interest. Second, the estimated shape of market externalities might suffer from the strong parametrization imposed on them. Last, I experimented with different specifications and found that the estimated absence of agglomeration externalities hinges on a realistic calibration for the average number of establishments per city: after setting κ equal to unit, hence an arbitrary value, and re-estimating the model I find an externality parameter on college employment of .15 on average. This reinforces the story put forward by Beaudry et al. (2017) on the negative elasticity of wage growth to employment growth via the tightness externality imposed on firm creation: hence, I tentatively conclude that a more careful empirical re-assessment of the magnitude and sign of market externalities in the presence of free-entry is needed.

Finally, the quantitative assessment also implies estimates for the standard deviation of the three structural shocks: I find that the standard deviation of shocks to non-college labor demand is twice as high as the one to college labor demand. This confirms the patterns observed in Table 3.8, namely that the R-squared of a regression of college wage growth on the determinants formally included in the just presented model is 3 percentage points higher than the R-squared of a regression of non-college wage growth on the same regressors.

Model Assessment

The targeted moments are reported in Table 3.10. Fig.3.7 and Fig.3.8 provide visual support by plotting the growth rate of specialization against either wage or employment growth between 1980 and 2000 for both education groups.

Overall the model performs quite well in terms of targeted moments. Being the labor demand side of the model much richer than the supply side, changes in the distribution of employment across cities are by construction dependent almost exclusively on nominal wage growth variation. This leads us to overestimate wage growth dispersion to be able to replicate the degree of dispersion in city-level employment growth observed in the data. I am also unable to replicate the negative comovement between specialization growth and non-college employment growth as seen in column 3 of Table 3.9: the introduction of a higher migration elasticity with respect to the real wage for non-college workers compared to the college ones could help match this feature of the data, by providing stronger incentives to non-college workers to move out of cities where specialization-driven productivity has increased the most.

In terms of non-targeted moments, the model provides a better description of location patterns for college workers as Table 3.13 and Fig.3.5 show: the correlation between wage growth in the model and the data is .28 and .40 for non-college and college workers, respectively, while the correlation between employment growth in the model and the data is .27 and .41, respectively.

Fig.3.6 provides an indirect test for the relevance of non-wage determinants in the location choice of workers, which have been mostly overlooked in the current setup: in the data non-college (college) employment growth features a .30 (.48) correlation with non-college wage growth, as opposed to a unit correlation in my model where labor supply shocks are absent.[13]

[13] The model is estimated by drawing $I = 500$ replications of the economy and a series of shocks for each replication. The Tables and Figures reported in Appendix 3.5.2 and 3.5.2 are constructed by solving for the equilibrium at the estimated parameters vector and in the absence of shocks. This choice rules out labor supply shocks in the form of random amenities.

Next, the model is assessed based on the geographical distribution of workers in both cross-sections. The correlation between (log) wages in the model and in the data is higher than one for (log) employment regardless of the education group and year, as Table 3.14 and Fig.3.10: in 2000 the (log) wage correlation is as high as .50 and .65 for low and college workers, respectively. Conversely, the (log) employment correlation is .43 and .41, respectively.[14]

Finally, I assess the elasticity of wages and employment to specialization in log-levels for both cross-sections. Table 3.15 reports the correlation between log wages/employment and log specialization for both education groups and cross-sections: the model performs substantially better in 2000, when location choices of workers are being steered by the Bartik shocks. Instead, the model assigns a too high correlation between the two outcome variables and specialization in 1980. The performance in 2000 with respect to wages is superior to the one with respect to employment: this reflects the non-trivial role of labor supply shocks considerations that I have chosen to abstract from in the present work.

3.4.3 Counterfactual Analysis

The objective of the quantitative exercise was to ultimately be able to gauge to what extent differences in specialization growth across Metropolitan Statistical Areas (MSAs) and education groups have accounted for differences in employment growth experienced by U.S. MSAs between 1980 and 2000. In this section, I turn this question.

The counterfactual exercise consists of comparing the distribution of workers across MSAs in the absence of specialization growth - i.e., keeping specialization constant at 1980 levels - with the actual distribution. This is done in Fig.3.15: for each education group I plot on the y-axis the difference between actual employment growth at the city-level and employment growth predicted by the model in the absence of specialization growth - i.e., exclusively under the influence of the Bartik shocks - and specialization growth on the x-axis. Each data point mark is proportional to MSAs employment by education group in 1980.

Both education groups should feature a positive slope: the discrepancy between actual and counterfactual employment growth should be higher in MSAs where the weight of specialization-driven out of total employment growth is higher due to the stronger increase in worker specialization. This is indeed what I find for college workers: college employment growth in Worcester, MA - which is among the top-5 MSAs in terms of college workers specialization growth - has been 50

[14]Interestingly, in spite of being able to account for a larger fraction of variation in college relative to non-college employment growth, the college/non-college relative performance of the model in log-levels deteriorates between 1980 and 2000, as the last two rows of Table 3.14 show.

percentage points higher than what would have been holding specialization constant.

Conversely, I find a negative slope for non-college workers: this result is due to the negative correlation between non-college workers specialization growth and non-college employment growth observed in the data.

MSAs that have registered higher specialization growth were the largest ones in 1980: hence, specialization growth can be considered as a divergence force akin to those portrayed in the literature, such as skill-biased technological change and agglomeration (Giannone (2018)).[15] However, the divergence implied by specialization growth holds only for college workers due to the negative elasticity of non-college empoyment to specialization.

Further, between-MSAs dispersion in specialization growth is twice as high for college compared to non-college workers, thus magnifying the impact of specialization growth on diverging location choices of college workers.

In Fig.3.16 I repeat the same analysis as in Fig.3.15 with the only exception that I now consider the difference between actual wage growth and the city-level and wage growth predicted by the model in the absence of specialization growth - i.e., holding specialization fixed at the 1980 level.

Specialization growth has accounted for a larger fraction of 1980-2000 wage growth for non-college compared to college workers, being specialization growth equal to 1.2% and .05% on average across cities for the two groups, respectively. Hence, although specialization growth has represented a force towards geographical divergence, it counteracted the rise in the college premium observed between 1980-2000.

Overall, I find that non-college workers specialization growth has been a solid driver of non-college wage growth, as the upper panel in Fig.3.16 shows, while college workers specialization growth is positively associated with college wage growth, but the coefficient is statistically insignificant. My interpretation of this finding is related to the negative college wage elasticity to college employment identified in the previous section: in the absence of specialization growth the productivity-related component of wages stays constant, but this is compensated by the absence of an adverse wage movement that would result from an increase in local college employment coupled with a negative relative labor supply channel.

[15] A linear regression of specialization growth on 1980 log employment yields a positive and statistically significant coefficient for college workers, and a positive but statistically insignificant coefficient for non-college workers.

3.5 Conclusion

In this paper, I operationalize the smithian insight of specialization being an engine of individual productivity and that the payoff of specializing is higher in large cities. I ask the following set of questions: Do specialized occupations earn a higher wage? And do they earn a higher wage especially in large cities? Are specialized occupations more than proportionately represented in large cities? Using data on US workers over three decades, I find a positive answer to all three questions, thus proving that the almost three-centuries-old smithian intuition bears strong ties with the organization of production in the US nowadays.

Furthermore, I provide an application of the specialization-driven wage premium by connecting with the literature that has explored the causes and consequences of the divergence in location choice for college and non-college workers between 1980 and 2000. I estimate a structural model of location choice where I let the traditional generic industry-specific labor demand shocks in the form of Bartik shocks race against the specific specialization-driven productivity shocks defined in this paper, to see how far can specialization go in accounting for the patterns observed in the data.

I find that the higher dispersion in measured specialization growth for college graduates can account for the larger divergence in location choices for this class of workers, and that average specialization growth has been higher for non-college graduates: hence, if specialization growth did not prove successful at counteracting regional divergence, it at least attenuated the rise of the college premium between US cities in 1980-2000.

Finally, the complementarity between specialization and city size highlighted in this paper and the resulting enhanced productivity premium in large cities could apply also to other organizational contexts, aside from cities, e.g., firms. At the same time, we can ask whether the wave of structural change investing the labor market in the form of increasing digitalization will make individual specialization even more valuable, thereby increasing the associated returns. Both the analysis of how employees' specialization affects firm productivity and the interaction between digitalization and worker specialization in the labor market landscape of the near future can both be attractive topics to pursue in future research.

Appendices

3.5.1 Theory

Existence of an equilibrium featuring spatial dispersion. Rewrite eq.3.9 as:

$$C_H^{\beta}(\mathbf{L}) = A_{H,j} W_{H,j} R_j^{-\alpha} L_{H,j}^{-\beta}; \quad C_L^{\beta}(\mathbf{L}) = A_{L,j} W_{L,j} R_j^{-\alpha} L_{L,j}^{-\beta} \quad \forall j \tag{3.35}$$

with $C_i^{\beta}(\mathbf{L}) = \left(\frac{1}{L_i} \sum_{j'} A_{ij} W_{ij'}^{1/\beta} R_{j'}^{-\alpha/\beta} \right)^{\beta}$.
Replacing eq.3.26 and eq.3.21 into the labor supply equation:

$$\begin{aligned} C_H^{\beta}(\mathbf{L}) &= \Upsilon \delta_H L_j^{\gamma} S_{H,j}^{\phi-\nu} L_{H,j}^{\nu-1-\beta} \left(\sum_{i'} \delta_{i'} L_j^{\gamma} S_{i'j}^{\phi-\nu} L_{i'j}^{\nu} \right)^{\frac{1-\nu-\alpha}{\nu}} \\ C_L^{\beta}(\mathbf{L}) &= \Upsilon \delta_L L_j^{\gamma} S_{L,j}^{\phi-\nu} L_{L,j}^{\nu-1-\beta} \left(\sum_{i'} \delta_{i'} L_j^{\gamma} S_{i'j}^{\phi-\nu} L_{i'j}^{\nu} \right)^{\frac{1-\nu-\alpha}{\nu}} \end{aligned} \tag{3.36}$$

with $\Upsilon = \left[\phi \left(\frac{1-\phi}{\kappa} \right)^{\frac{1-\nu}{\nu}} \right]^{1-\alpha} \left(\frac{\alpha}{H} \right)^{-\alpha}$.
The high-low skilled ratio in city j is obtained implicitly by dividing both sides of the eq.3.36 by each other:

$$\chi_j \left(\mathbf{C}(\mathbf{L}), \mathbf{S}_j \right) = \left(\frac{C_L(\mathbf{L})}{C_H(\mathbf{L})} \right)^{\frac{\beta}{1+\beta-\nu}} \left(\frac{\widetilde{S_{H,j}}}{\widetilde{S_{L,j}}} \right)^{\frac{1}{1+\beta-\nu}} \tag{3.37}$$

where $\mathbf{C}(\mathbf{L}) = [C_L(\mathbf{L}), C_H(\mathbf{L})]$ and $\mathbf{S}_j = \left[\delta_L S_{L,j}^{\phi-\nu}, \delta_H S_{H,j}^{\phi-\nu} \right]$.
Since $\nu < 1 + \beta$, the ratio of high-to-low skilled, $\chi_j = \frac{L_{H,j}}{L_{L,j}}$, is increasing in relative specialization, $S_{H,j}/S_{L,j}$. Using local labor market clearing:

$$L_{L.j} = L_j \left(\frac{1}{1+\chi_j} \right); \quad L_{H,j} = L_j \left(\frac{\chi_j}{1+\chi_j} \right) \tag{3.38}$$

Replacing the expression in eq.3.37 into eq.3.36:

$$1 = \Upsilon \left(\widetilde{S_{H,j}} \left(\frac{\chi_j}{1+\chi_j} \right)^{\nu} + \widetilde{S_{L,j}} \left(\frac{1}{1+\chi_j} \right)^{\nu} \right)^{\frac{1-\nu-\alpha}{\nu}} \left(\frac{\widetilde{S_{L,j}}}{C_L^{\beta}(\mathbf{L})} + \frac{\widetilde{S_{H,j}}}{C_H^{\beta}(\mathbf{L})} \right) L_j^{\left(\frac{1-\alpha}{\nu} \right)(\gamma+\nu)-(1+\beta)} \tag{3.39}$$

from which it follows that if $\frac{\gamma+\nu}{\nu} < \frac{1+\beta}{1-\alpha}$, then the equilibrium features spatial dispersion.[16]

[16] While this discussion does not provide a proof of uniqueness, solving for the equilibrium repeated times starting from distinct initial guesses for $L_{H,j}$ and $L_{L,j}$ always yields the same solution, regardless of the combination of structural parameters adopted.

3.5.2 Estimation

We follow the below described estimation routine:

1. We draw $J \times 3$ (J being the maximum number of cities and 3 the number of structural shocks) seeds from a uniform distribution: given the seed, for each city and type of shock we draw a vector of $I = 500$ shocks from a standard normal distribution once and for all;

2. We run the optimization using the Nelder-Mead Method. At each set of structural parameters the algorithm solves for the model according to the following steps:

   (a) rescale the first set of shocks by the relevant standard deviation;

   (b) solve for the model:

      i. start with a uniform spatial distribution as initial guess;
      ii. solve for the labor demand block (wages) of the model by calibrating κ to match exactly the log of the total count of firms in the economy;
      iii. solve for the labor supply block (supply) of the model by calibrating H to match exactly the average (log) of housing prices across MSAs;
      iv. update the factors allocation by taking the half-step between the solution provided in the previous step and the current iteration guess;
      v. repeat i-v until convergence.

   (c) repeat (a)-(b) I times, using each time the i-th set of structural shocks;

   (d) calculate the moments as in eq.3.29.

Housing prices at the MSA-level are estimated by means of a hedonic regression of the type:

$$\ln rent_{ij} = \beta \ln rooms_{ij} + \gamma \mathbf{x_{ij}} + \epsilon_{ij} \tag{3.40}$$

after having excluded farm-type of properties, and where $\mathbf{x_{ij}}$ is a set of dummy variables corresponding to the decade when the house was built.

Tables: Empirical Analysis

Table 3.1: Example from Task Statements file (code: 11-1011.00)

Task ID	Description
8823	Direct or coordinate an organization's financial or budget activities, to fund operations, maximize investments, or increase efficiency.
8824	Confer with board members, organization officials, or staff members to discuss issues, coordinate activities, or resolve problems.
8825	Analyze operations to evaluate performance of a company or its staff in meeting objectives or to determine areas of potential cost reduction, program improvement, or policy change.
8826	Direct, plan, or implement policies, objectives, or activities of organizations or businesses to ensure to continuing operations, maximize returns on investments, or to increase productivity.
8827	Prepare budgets for approval, including those for funding or implementation of programs.
8828	Direct or coordinate activities of businesses or departments concerned with production, pricing, sales, or distribution of products.
...	...

Source: O*NET.

Table 3.2: Occupation-level information: “economist”

Occ. Code	Task ID	Description	Frequency
19-3011.00	7544	Teach theories, principles, and methods of economics.	4.21
19-3011.00	7536	Study economic and statistical data in area of specialization, such as finance, labor, or agriculture.	
19-3011.00	20053	Conduct research on economic issues and disseminate research findings through technical reports or scientific articles in journals.	3.96
19-3011.00	7538	Compile, analyze, and report data to explain economic phenomena and forecast market trends, applying mathematical models and statistical techniques.	4.04
19-3011.00	20052	Study the socioeconomic impacts of new public policies, such as proposed legislation, taxes, services, and regulations.	3.46
19-3011.00	7542	Supervise research projects and students’ study projects.	3.83
19-3011.00	7539	Formulate recommendations, policies, or plans to solve economic problems or to interpret markets.	3.29
19-3011.00	7540	Develop economic guidelines and standards and prepare points of view used in forecasting trends and formulating economic policy.	2.83
19-3011.00	7537	Provide advice and consultation on economic relationships to businesses, public and private agencies, and other employers.	2.88
19-3011.00	7543	Forecast production and consumption of renewable resources and supply, consumption and depletion of nonrenewable resources.	2.45
19-3011.00	7541	Testify at regulatory or legislative hearings concerning the estimated effects of changes in legislation or public policy and present recommendations based on cost-benefit analyses.	1.57
19-3011.00	20051	Provide litigation support, such as writing reports for expert testimony or testifying as an expert witness.	1.53

Source: O*NET.

Table 3.3: Summary statistics by occupational group

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Management, Business, Science, and Arts Occupations	22	3	11	71	12	-14	40
Business Operations Specialists	17	4	-15	59	11	-32	27
Financial Specialists	18	6	-17	65	17	-24	31
Computer and Mathematical Occupations	18	3	-11	56	11	-38	54
Architecture and Engineering Occupations	19	3	-2	57	12	-36	47
Life, Physical, and Social Science Occupations	19	4	-5	59	13	-32	30
Community and Social Services Occupations	20	2	2	75	12	-8	-5
Legal Occupations	16	4	-22	62	11	-26	49
Education, Training, and Library Occupations	31	6	44	118	33	34	-11
Arts, Design, Entertainment, Sports, and Media Occupations	19	3	-5	69	14	-16	16
Healthcare Practitioners and Technical Occupations	19	4	-4	87	22	5	30
Healthcare Support Occupations	17	1	-15	82	3	3	-40
Protective Service Occupations	19	5	-7	68	15	-18	-1
Food Preparation and Serving Occupations	23	5	15	117	28	36	-58
Building and Grounds Cleaning and Maintenance Occupations	23	5	16	92	22	11	-35
Personal Care and Service Occupations	20	3	2	87	12	8	-31
Sales and Related Occupations	19	3	-5	77	14	-6	16
Office and Administrative Support Occupations	19	5	-7	85	23	2	-19
Farming, Fishing, and Forestry Occupations	18	8	-16	75	29	-13	-37
Construction and Extraction Occupations	20	4	1	75	11	-7	-7
Installation, Maintenance, and Repair Workers	22	7	8	88	26	5	5
Production Occupations	22	7	8	88	26	5	5
Transportation and Material Moving Occupations	17	4	-12	83	16	2	-5

Average number of tasks as measured in the *Task Statements* file and standard deviation (column 1-2), average difference between the log number of tasks for each specific occupation and overall mean in % (column 3), average number of frequency-adjusted tasks and standard deviation (column 4-5), average difference between the log number of frequency-adjusted tasks for each specific occupation and overall mean in % (column 6), average difference between the log wage for each specific occupation and overall mean in % (column 7). Source: Census 2000 (IPUMS)/O*NET.

Table 3.4: Specialization vs. alternative occupational attributes

N	Abstract	Routine	Manual	Offshorability
325	0.00299***	-0.000135	0.000990	0.00209**
	(0.000396)	(0.000388)	(0.000665)	(0.000711)

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Estimated coefficients in regression of specialization metric on alternative occupational attributes. Source: O*NET/Autor and Dorn (2013).

Table 3.5: Specialization wage premium: cross-sectional regressions - 1980

Low Skilled			
	w_{jk}	w_{jk}	w_{jk}
Log tasks	-0.330***	-0.293***	-0.219***
	(0.00955)	(0.00891)	(0.00831)
Log of MA size	0.105**	0.124***	0.0763*
	(0.0382)	(0.0357)	(0.0333)
Log tasks × Log MA size	-0.0181*	-0.0209**	-0.0101
	(0.00868)	(0.00810)	(0.00755)
Observations	10684	10684	10684
Industry FE	NO	NO	YES
Residual wage	NO	YES	YES
High Skilled			
	w_{jk}	w_{jk}	w_{jk}
Log tasks	-0.342***	-0.279***	-0.190***
	(0.0140)	(0.0132)	(0.0127)
Log of MA size	0.167**	0.181***	0.182***
	(0.0555)	(0.0526)	(0.0506)
Log tasks × Log MA size	-0.0309*	-0.0325**	-0.0331**
	(0.0126)	(0.0120)	(0.0115)
Observations	9416	9416	9416
Industry FE	NO	NO	YES
Residual wage	NO	YES	YES

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

1980 Results: Log wage regressions on inverse of occupational specialization (log of adjusted number of tasks) and city size. Occupations have been lumped together into 40 groups, corresponding to 40 percentiles of the distribution for the inverse of occupational specialization (down from originally $\approx$ 325). Wage in each MSA and occupation is constructed as follows: within city/occupational group average of log wage (column 1); the city/occupation fixed effect, $\hat{d}_{jk}$, obtained by regressing $w_i = \beta_0 + \beta_1\text{Female}_i + \beta_2\text{White}_i + \beta_3\text{Experience}_i + \beta_4\text{Experience}_i^2 + d_{jk} + e_i$, being w_i the log wage received by individual i and experience calculated as age minus 25 (column 2); the city/occupation fixed effect, $\hat{d}_{jk}$, obtained by regressing $w_i = \beta_0 + \beta_1\text{Female}_i + \beta_2\text{White}_i + \beta_3\text{Experience}_i + \beta_4\text{Experience}_i^2 + \gamma_{industry} + d_{jk} + e_i$, where γ are industry fixed effects at the 2-digit level (column 3). Source: Census 1980 (IPUMS)/O*NET.

Table 3.6: Specialization wage premium: cross-sectional regressions - 2000

Low Skilled			
	w_{jk}	w_{jk}	w_{jk}
Log tasks	-0.236***	-0.202***	-0.193***
	(0.00660)	(0.00588)	(0.00566)
Log of MA size	0.170***	0.185***	0.160***
	(0.0269)	(0.0239)	(0.0230)
Log tasks × Log MA size	-0.0305***	-0.0319***	-0.0267***
	(0.00610)	(0.00543)	(0.00523)
Observations	9486	9486	9486
Industry FE	NO	NO	YES
Residual wage	NO	YES	YES

High Skilled			
	w_{jk}	w_{jk}	w_{jk}
Log tasks	-0.312***	-0.248***	-0.227***
	(0.00990)	(0.00929)	(0.00942)
Log of MA size	0.237***	0.246***	0.246***
	(0.0402)	(0.0377)	(0.0382)
Log tasks × Log MA size	-0.0447***	-0.0451***	-0.0456***
	(0.00914)	(0.00858)	(0.00870)
Observations	9184	9184	9184
Industry FE	NO	NO	YES
Residual wage	NO	YES	YES

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

2000 Results: Log wage regressions on inverse of occupational specialization (log of adjusted number of tasks) and city size. Occupations have been lumped together into 40 groups, corresponding to 40 percentiles of the distribution for the inverse of occupational specialization (down from originally $\approx$ 325). Wage in each MSA and occupation is constructed as follows: within city/occupational group average of log wage (column 1); the city/occupation fixed effect, $\hat{d}_{jk}$, obtained by regressing $w_i = \beta_0 + \beta_1\text{Female}_i + \beta_2\text{White}_i + \beta_3\text{Experience}_i + \beta_4\text{Experience}_i^2 + d_{jk} + e_i$, being w_i the log wage received by individual i and experience calculated as age minus 25 (column 2); the city/occupation fixed effect, $\hat{d}_{jk}$, obtained by regressing $w_i = \beta_0 + \beta_1\text{Female}_i + \beta_2\text{White}_i + \beta_3\text{Experience}_i + \beta_4\text{Experience}_i^2 + \gamma_{industry} + d_{jk} + e_i$, where γ are industry fixed effects at the 2-digit level (column 3). Source: Census 2000 (IPUMS)/O*NET.

Table 3.7: Specialization wage premium: panel regressions

	w_{ikt}	w_{ikt}
Log tasks	-0.188***	-0.0965***
	(0.00318)	(0.00540)
Log of MA size	0.152***	0.0820***
	(0.00755)	(0.0138)
Log tasks × Log MA size	-0.0224***	-0.00973***
	(0.00171)	(0.00307)
Observations	962010	962010
Industry FE	YES	YES
Individual FE	NO	YES

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Minc erian regressions: $\ln w_{ijot} = \delta X_{it} + \beta_0 \ln(\text{tasks}_{ot}) + \beta_1 \ln(\text{size}_{jt}) + \beta_2(\ln(\text{tasks}_{ot}) \times \ln(\text{size}_{jt})) + \gamma_i + \gamma_{industry} + \gamma_t + e_{it}$. Controls include: college, dropout, white, male, married, a set of dummies standing for age groups, experience on the job (quadratic), experience in the occupation (quadratic); industry categories are at the 2-digit level of detail. Source: Survey of Income and Participation Program (1990-1996)/O*NET.

Table 3.8: 1980-2000: Determinants of wage growth by skill group

	$\Delta w_{L,j}$	$\Delta w_{L,j}$	$\Delta w_{L,j}$	$\Delta w_{H,j}$	$\Delta w_{H,j}$	$\Delta w_{H,j}$
$\Delta(l_{H,j}/l_{L,j})$	0.402***	0.387***	0.310***	0.396***	0.347***	0.289***
	(0.0277)	(0.0274)	(0.0290)	(0.0266)	(0.0262)	(0.0266)
$Bartik_{L,j}$		0.175***	0.172***			
		(0.258)	(0.253)			
$\Delta s_{L,j}$			0.203***			
			(0.369)			
$Bartik_{H,j}$					0.250***	0.244***
					(0.121)	(0.118)
$\Delta s_{H,j}$						0.208***
						(0.164)
Observations	196	196	196	196	196	196
R^2	0.162	0.192	0.228	0.156	0.216	0.256

Standardized beta coefficients; Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Bartik shocks at the city/skill-level are constructing by taking a weighted sum of growth in industry wages for a given skill group at the national level (excluding the Metropolitan Area of interest), with weights corresponding to 1980 industry employment share. Source: Census 1980/2000 (IPUMS)/O*NET.

Table 3.9: 1980-2000: Determinants of city size growth by skill group

	$l_{L,j}$	$l_{L,j}$	$l_{H,j}$	$l_{H,j}$
$Bartik_{L,j}$	0.111	0.116		
	(1.140)	(1.136)		
$\Delta s_{L,j}$		-0.117		
		(1.540)		
$Bartik_{H,j}$			0.323***	0.301***
			(0.557)	(0.540)
$\Delta s_{H,j}$				0.255***
				(0.731)
Observations	196	196	196	196
R^2	0.012	0.026	0.104	0.169

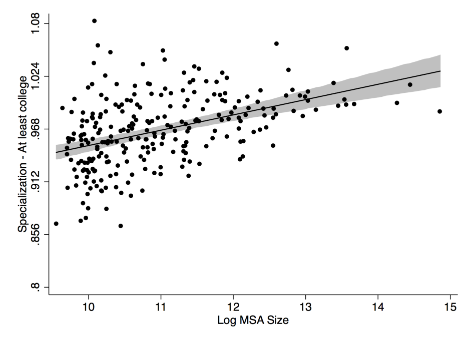
Standardized beta coefficients; Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

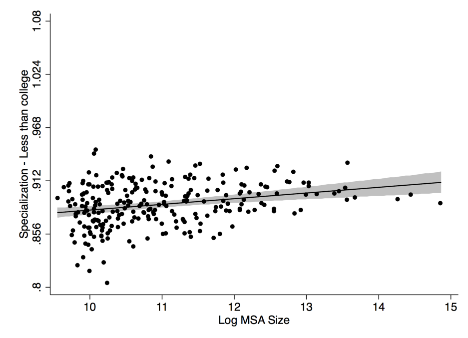
Bartik shocks at the city/skill-level are constructing by taking a weighted sum of growth in industry wages for a given skill group at the national level (excluding the Metropolitan Area of interest), with weights corresponding to 1980 industry employment share. Source: Census 1980/2000 (IPUMS)/O*NET.

Figures: Empirical Analysis

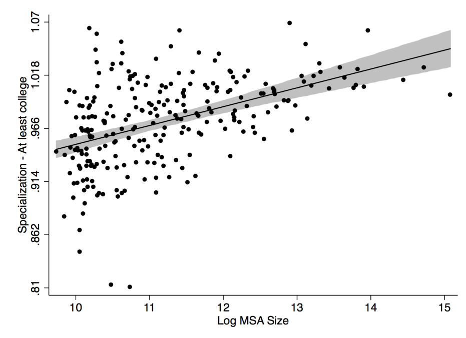
Figure 3.1: Average specialization per city/educational attainment group



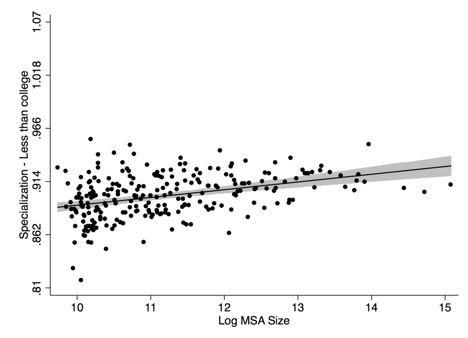
(a) Census 1980: High Skill



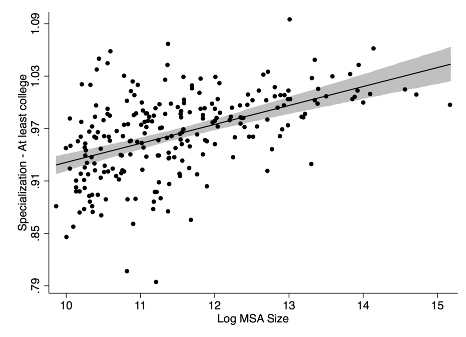
(b) Census 1980: Low Skill



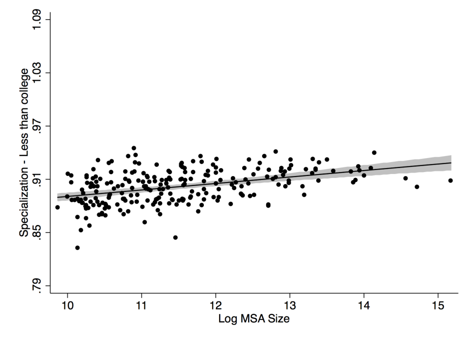
(c) Census 1990: High Skill



(d) Census 1990: Low Skill



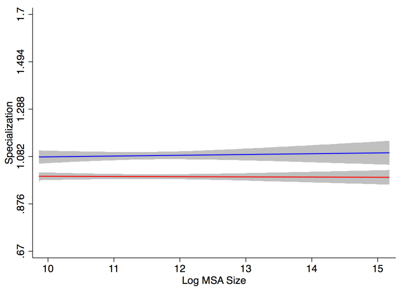
(e) Census 2000: High Skill



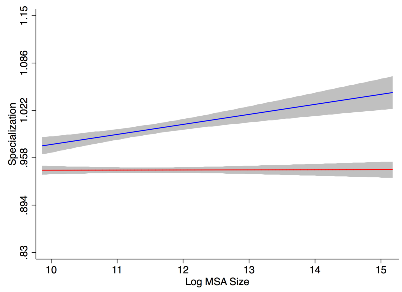
(f) Census 2000: Low Skill

Red line - less than college, blue line - college degree or higher. Source: Census 1980/1990/2000 (IPUMS)/O*NET.

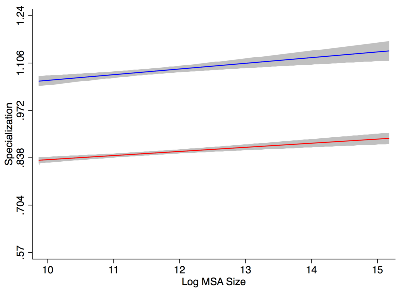
Figure 3.2: Average specialization per city/educational attainment group by sector



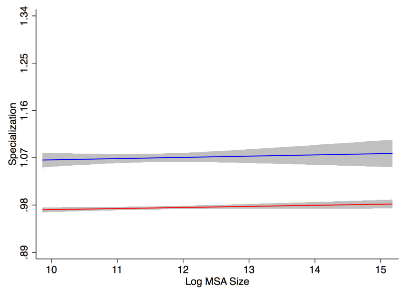
(a) Agriculture and Mining



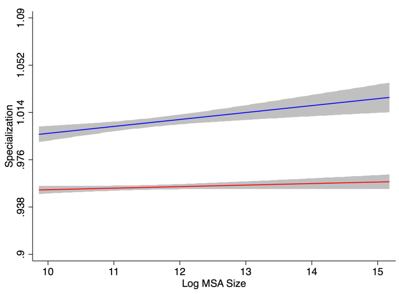
(b) Utilities, Transportation and Communications



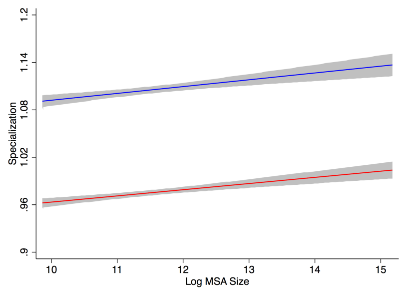
(c) Manufacturing



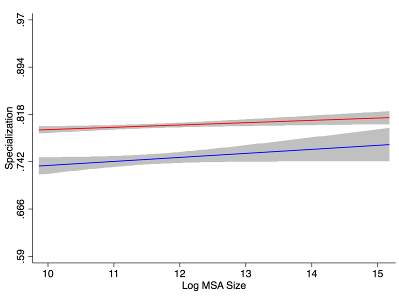
(d) Construction



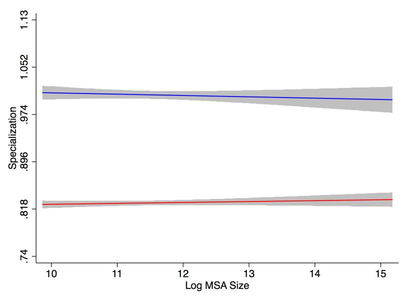
(e) Retail and Wholesale



(f) Finance and Professional Services



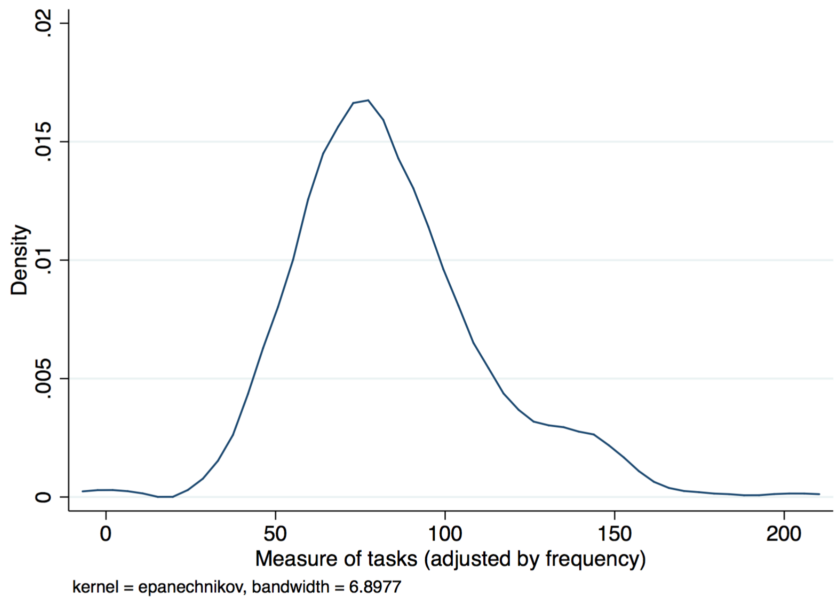
(g) Education and Health



(h) Other Services

Red line - less than college, blue line - college degree or higher. Source: Census 2000 (IPUMS)/O*NET.

Figure 3.3: Kernel density of sum of frequency-adjusted tasks



Occupational categories as in Dorn (2009). Source: O*NET and Dorn (2009).

Tables: Model Estimation

Table 3.10: Targeted moments

Description	Model	Data
Mean wage growth rate L	.76	.78
Mean wage growth rate H	.84	.87
Variance wage growth rate L	.02	.01
Variance wage growth rate H	.05	.01
Covariance wage L/specialization growth L	2.69×10^{-4}	3.3×10^{-4}
Covariance wage H/specialization growth L	3.41×10^{-4}	3.0×10^{-4}
Covariance wage L/specialization growth H	1.84×10^{-3}	$.56 \times 10^{-3}$
Covariance wage H/specialization growth H	2.68×10^{-3}	$.62 \times 10^{-3}$
Mean employment growth rate L	.34	.43
Mean employment growth rate H	.76	.77
Variance employment growth rate H	.06	.09
Variance employment growth rate H	.07	.10
Covariance employment L/specialization growth L	4.12×10^{-4}	-4.62×10^{-4}
Covariance employment H/specialization growth L	3.51×10^{-4}	4.65×10^{-4}
Covariance employment L/specialization growth H	2.65×10^{-3}	1.08×10^{-3}
Covariance employment H/specialization growth H	3.01×10^{-3}	2.50×10^{-3}

Table 3.11: Parameter estimates

Description	Parameter	Estimate
Elasticity of wage to measure of tasks	$\nu - \phi$	-.48
Elasticity of labor substitution	$1/(1-\nu)$	1.69
Elasticity of H wage to L employment	γ_{HL}	.25
Elasticity of H wage to H employment	$\phi - 1 + \gamma_{HH}$	-.81
Elasticity of L wage to L employment	$\phi - 1 + \gamma_{LL}$	-.91
Elasticity of L wage to H employment	γ_{LH}	.00
Standard deviation L demand shock	σ^{ε}_{L}	.90
Standard deviation H demand shock	σ^{ε}_{H}	.30
Standard deviation supply shock	σ^{η}	.61

Table 3.12: Calibrated parameters

Description	Parameter	Calibration
Expenditure share on housing	α	Davis and Ortalo-Magné (2007)
Inverse of migration elasticity	β	Diamond (2016)
Housing supply	H_t	avg. rent in each year
City-independent worker productivity	$\delta_{i,t}$	avg. wage per worker type in each year
Sunk cost of entry	κ	avg. # establishments per city in 2000

Table 3.13: Non-targeted moments on 1980-2000 changes: correlation outcomes between data and model

	Correlation
Wage growth L	.282
Wage growth H	.397
employment growth L	.273
employment growth H	.415

Pearson coefficient of correlation between 1980-2000 wage/employment growth in the model and in the data by skill group.

Table 3.14: Non-targeted moments on 1980/2000 levels: correlation outcomes between data and model

	Correlation: 1980	Correlation: 2000
Log wage L	.330	.500
Log wage H	.562	.646
Log employment L	.252	.438
Log employment H	.286	.430

Pearson coefficient of correlation between log wage/log employment in the model and in the data by skill group for different years.

Table 3.15: Non-targeted moments on 1980/2000 levels: correlation between specialization and wage/employment according to the data vs. model

	Correlation: 1980	Correlation: 2000
Log wage L/log specialization L	.147 **.568**	.351 **.527**
Log wage H/log specialization H	.559 **.989**	.617 **.773**
Log employment L/log specialization L	.123 **.584**	.244 **.537**
Log employment H/log specialization H	.285 **.992**	.370 **.773**

Pearson coefficient of correlation between log wage and log employment with log specialization by skill group for different years: first row for each cell is correlation in the data, second row (bold font) is correlation in the model.

Figures: Model Estimation

Figure 3.4: The division of labor across US cities

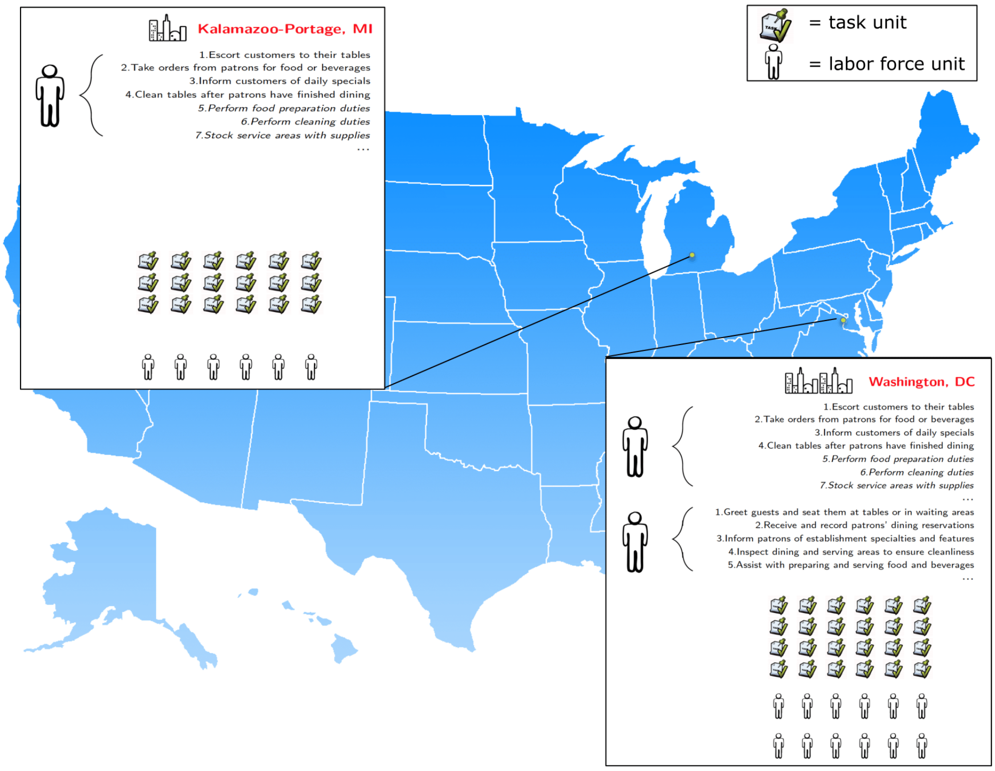
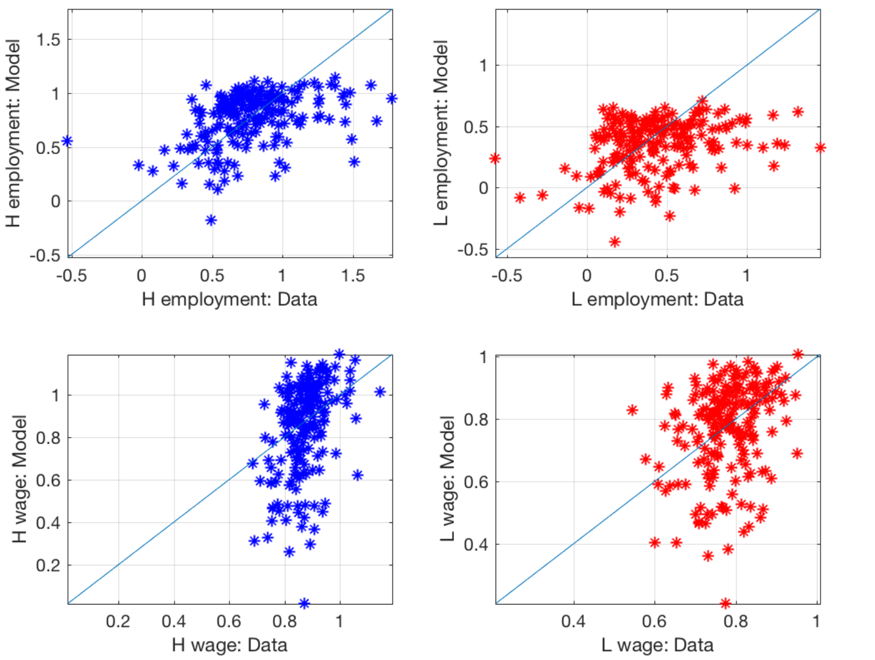
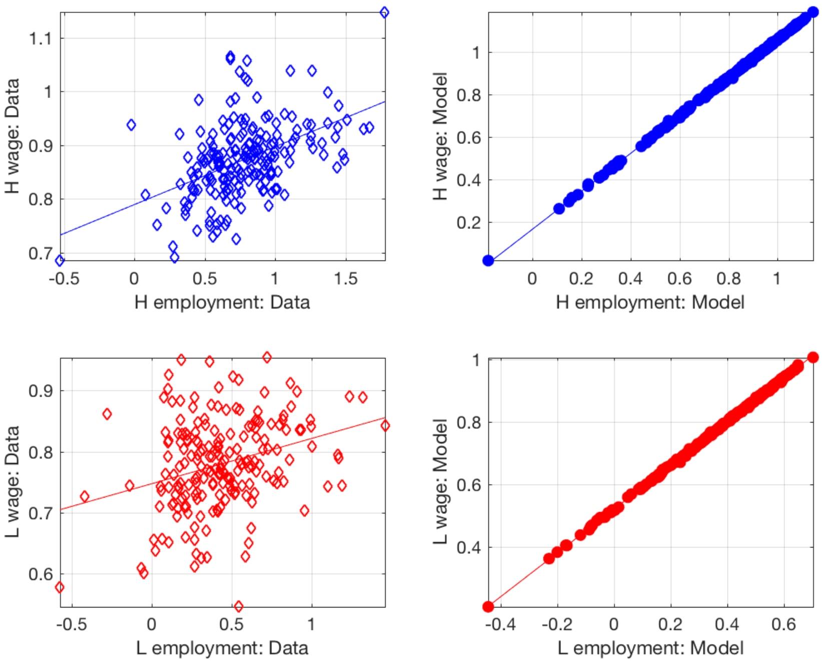


Figure 3.5: Non-targeted moments: employment/wage growth in data vs. model



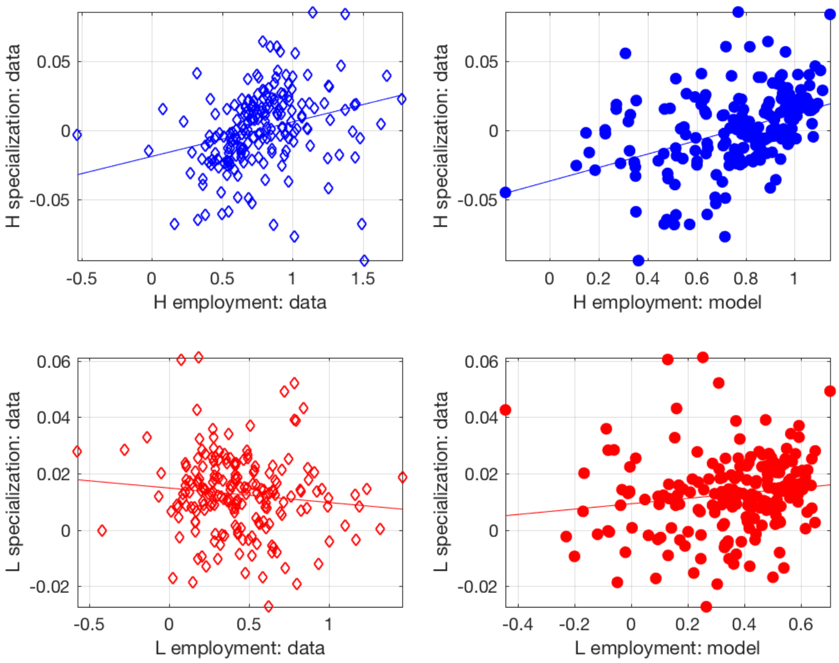
employment growth in the model (y-axis) vs. data (x-axis); wage growth in the model (y-axis) vs. data (x-axis) by skill group (red is low skilled, blue is high skilled), together with 45-degree line.

Figure 3.6: Non-targeted moments: employment vs. wage growth in the data/model



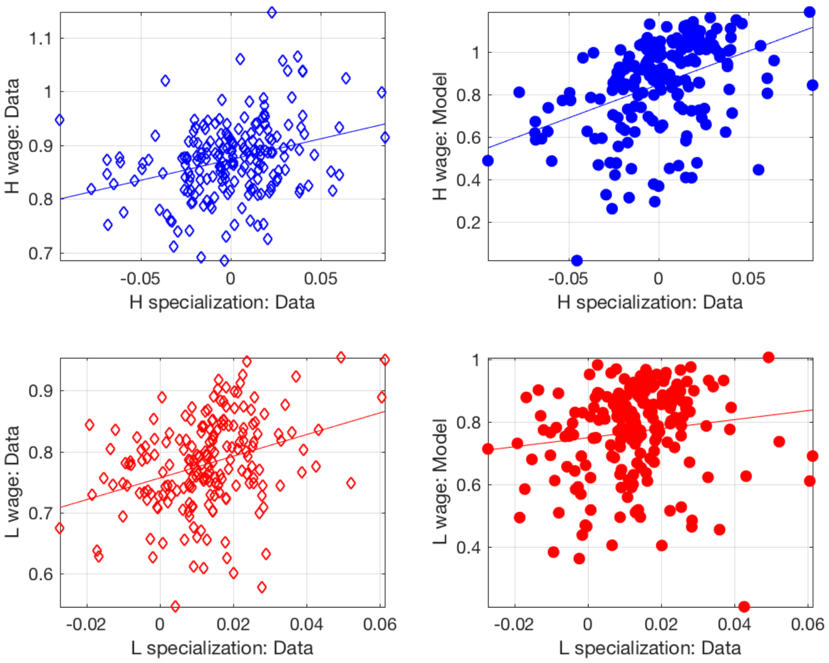
Wage growth (y-axis) vs. employment growth (x-axis) in the model (filled circles) and in the data (empty diamonds) by skill group (red is low skilled, blue is high skilled), together with least squares fitting line.

Figure 3.7: Targeted moments: specialization vs. employment growth in the data/model



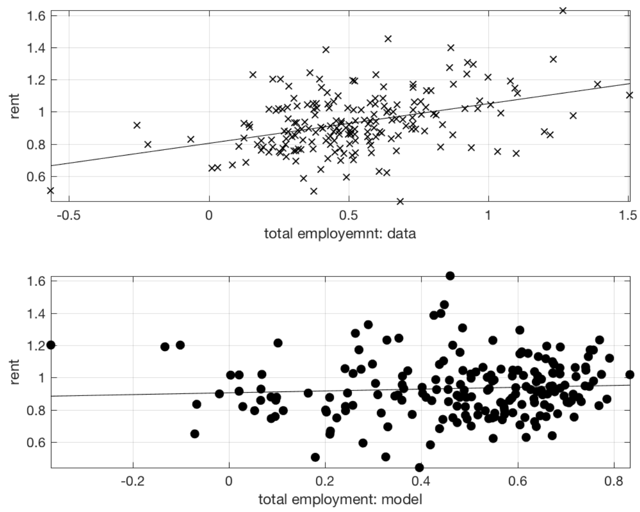
Specialization growth (y-axis) vs. employment growth (x-axis) in the model (filled circles) and in the data (empty diamonds) by skill group (red is low skilled, blue is high skilled), together with least squares fitting line.

Figure 3.8: Targeted moments on 1980-2000 changes: wage vs. specialization growth in the data/model



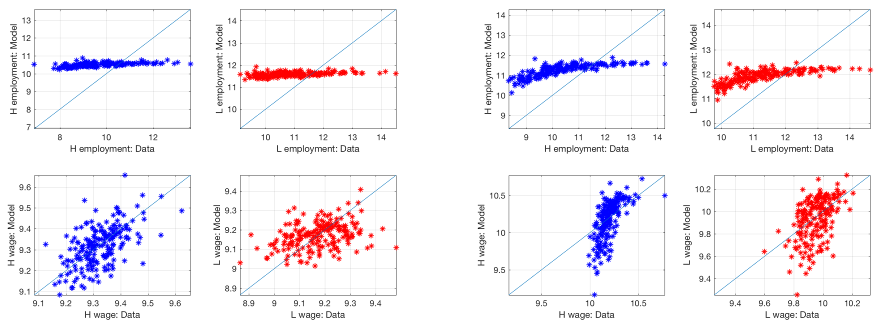
Wage growth (y-axis) vs. specialization growth (x-axis) in the model (filled circles) and in the data (empty diamonds) by skill group (red is low skilled, blue is high skilled), together with least squares fitting line.

Figure 3.9: Non-targeted moments: rent vs. employment growth in the data/model



Rent growth (y-axis) vs. employment growth (x-axis) in the model (filled circles) and in the data (empty diamonds), together with least squares fitting line. Metropolitan Area average rents in a given year are derived by running for each year a hedonic regression of rental rates on the number of bedrooms and a set of dummy variables corresponding to the period of construction.

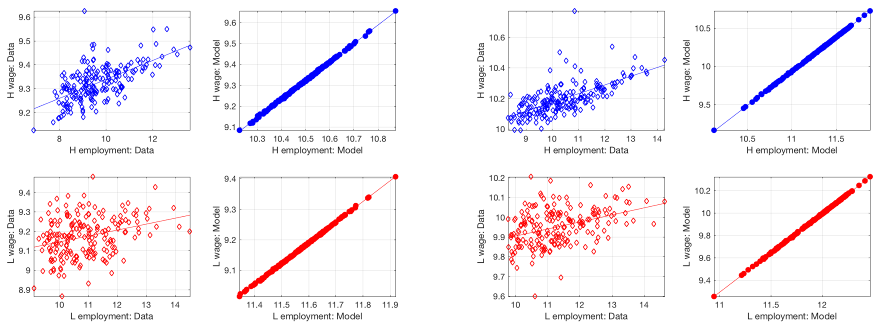
Figure 3.10: Non-targeted moments: employment/wage in data vs. model



(a) 1980 (b) 2000

Log employment in the model (y-axis) vs. data (x-axis); log wage in the model (y-axis) vs. data (x-axis) by skill group (red is low skilled, blue is high skilled), together with 45-degree line.

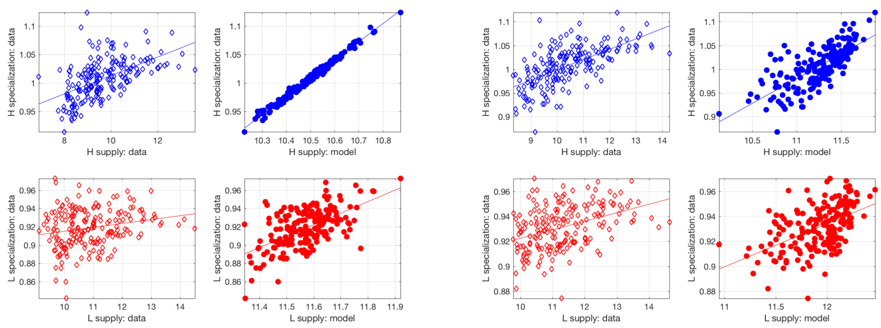
Figure 3.11: Non-targeted moments: employment vs. wage in the data/model



(a) 1980 (b) 2000

Log wage (y-axis) vs. log employment (x-axis) in the model (filled circles) and in the data (empty diamonds) by skill group (red is low skilled, blue is high skilled), together with least squares fitting line.

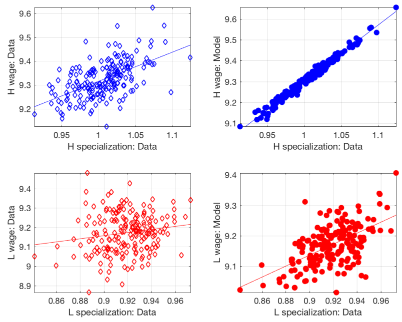
Figure 3.12: Non-targeted moments: specialization vs. employment in the data/model



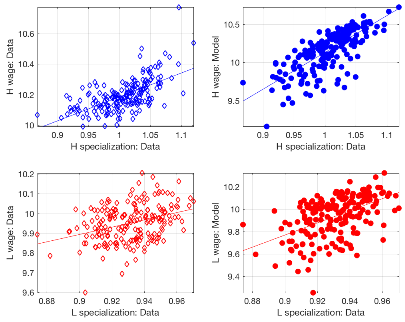
(a) 1980 (b) 2000

Log specialization (y-axis) vs. log employment (x-axis) in the model (filled circles) and in the data (empty diamond) by skill group (red is low skilled, blue is high skilled), together with least squares fitting line.

Figure 3.13: Non-targeted moments: wage vs. specialization in the data/model



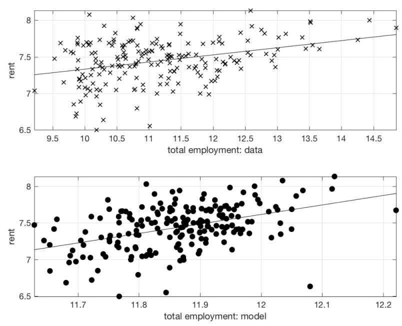
(a) 1980



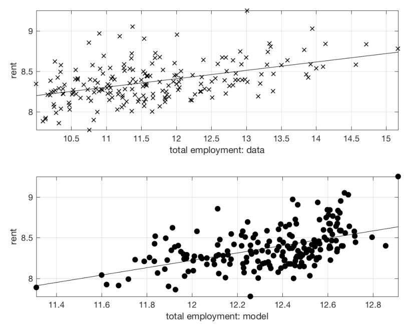
(b) 2000

Log wage (y-axis) vs. log specialization (x-axis) in the model (filled circles) and in the data (empty diamonds) by skill group (red is low skilled, blue is high skilled), together with least squares fitting line.

Figure 3.14: Non-targeted moments: rent vs. employment in the data/model



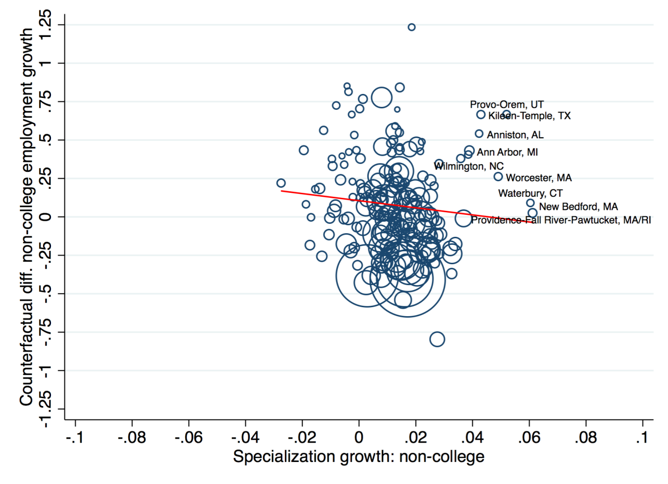
(a) 1980



(b) 2000

Log rent (y-axis) vs. log employment (x-axis) in the model (filled circles) and in the data (empty diamonds), together with least squares fitting line. Metropolitan Area average rents in a given year are derived by running for each year a hedonic regression of rental rates on the number of bedrooms and a set of dummy variables corresponding to the period of construction.

Figure 3.15: Counterfactual employment growth in the absence of specialization growth

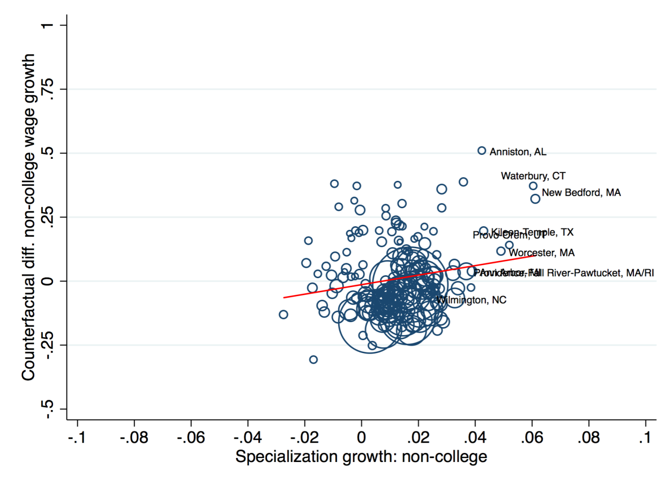


(a) Non-college

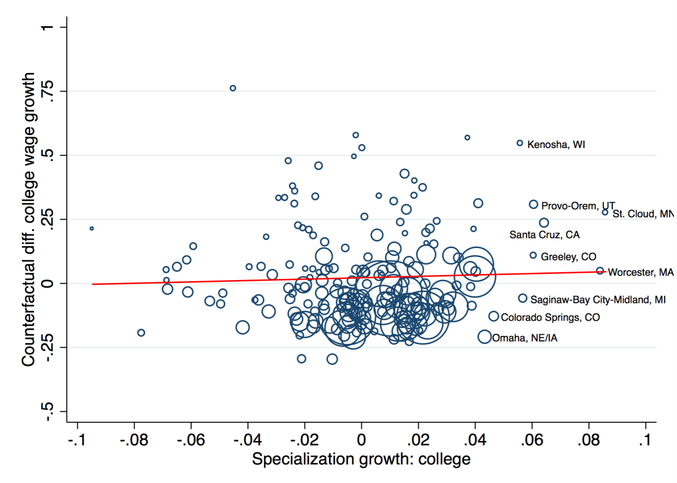
(b) College

1980-2000 difference between actual employment growth and model predicted employment growth in the absence of a variation in the degree of specialization (y-axis) against specialization growth (x-axis) by skill group, together with a least squares fitting line. Marker size is proportional to employment by education group in 1980. Labels are reported only for Metropolitan Areas that have experienced above the 95^{th} percentile in specialization growth for the relevant skill group.

Figure 3.16: Counterfactual wage growth in the absence of specialization growth



(a) Non-college



(b) College

1980-2000 difference between actual wage growth and model predicted wage growth in the absence of a variation in the degree of specialization (y-axis) against specialization growth (x-axis) by skill group, together with a least squares fitting line. Marker size is proportional to employment by education group in 1980. Labels are reported only for Metropolitan Areas that have experienced above the 95^{th} percentile in specialization growth for the relevant skill group.

Bibliography

[1] Andrew Abel. Optimal investment under uncertainty. *American Economic Review*, 73(1):228–233, 1983.

[2] Treb Allen and Costas Arkolakis. Trade and the Topography of the Spatial Economy. *The Quarterly Journal of Economics*, 129(3):1085–1140, 2014.

[3] Costas Arkolakis. Market Penetration Costs and the New Consumers Margin in International Trade. *Journal of Political Economy*, 118(6):1151–1199, 2010.

[4] David H. Autor and David Dorn. The growth of low-skill service jobs and the polarization of the US Labor Market. *American Economic Review*, 103(5):1553–1597, 2013.

[5] David H. Autor, Frank Levy, and Richard J. Murnane. The Skill Content of Recent Technological Change: an Empirical Exploration. *Quarterly Journal of Economics*, 118(4):1279–1333, 2003.

[6] Rudiger Bachmann, Steven Elstner, and Eric R. Sims. Uncertainty and Economic Activity: Evidence from Business Survey Data. *American Economic Journal: Macroeconomics*, 5(2):217–249, 2013.

[7] Regis Barnichon. Vacancy posting, job separation and unemployment fluctuations. *Journal of Economic Dynamics and Control*, 36(3):315–330, 2012.

[8] Timothy J. Bartik. *Who Benefits from State and Local Economic Development Policies?* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, 1991.

[9] James R. Baumgardner. The Division of Labor, Local Markets, and Worker Organization. *Journal of Political Economy*, 96(3):509–527, 1988.

[10] Paul Beaudry, David A. Green, and Benjamin M. Sand. Does Industrial Composition Matter for Wages? A Test of Search and Bargaining Theory. *Econometrica*, 80(3):1063–1104, 2012.

[11] Paul Beaudry, David A. Green, and Benjamin M. Sand. In Search of Labor Demand. *American Economic Review, forthcoming*, 2018.

[12] Gary S. Becker and Kevin M. Murphy. The Division of Labor, Coordination Costs, and Knowledge. *The Quarterly Journal of Economics*, 107(4):1137–1160, 1992.

[13] Randy Becker and J. Vernon Henderson. Intraindustry specialization and urban development. *In Jean Marie Huriot and Jacques-François Thisse (eds.) Economics of Cities: Theoretical Perspectives.*, pages 138–166, 2000.

[14] Kristian Behrens, Gilles Duranton, and Frédéric Robert-Nicoud. Productive Cities: Sorting, Selection, and Agglomeration. *Journal of Political Economy*, 122(3):507–553, 2015.

[15] Nicholas Bloom. The Impact of Uncertainty Shocks. *Econometrica*, 77(3):623–685, 2009.

[16] Nicholas Bloom. Fluctuations in Uncertainty. *Journal of Economic Perspectives*, 28(2):153–176, 2014.

[17] Tito Boeri and Pietro Garibaldi. Graded Security and Labor Market Mobility Clean Evidence from the Italian Jobs Act. *INPS Working Paper Series, ISSN 2532 -8565*, 2018.

[18] Karl Brenke, Ulf Rinne, and Klaus F. Zimmermann. Short-Time Work: The German Answer to the Great Recession. *IZA DP No. 5780*, 2011.

[19] Arthur F. Burns and Wesley C. Mitchell. Measuring Business Cycles. *National Bureau of Economic Research*, 1946.

[20] David Card. Immigration and inequality. *American Economic Review*, 99(2):1–21, 2009.

[21] Vasco M. Carvalho and Basile Grassi. Large Firm Dynamics and the Business Cycle. *Cambridge Working Papers in Economics 1556*, 2017.

[22] Gabriel Chodorow-Reich and Johannes Wieland. Secular Labor Reallocation and Business Cycles. *Working Paper*, 2018.

[23] Pierre-Philippe Combes, Gilles Duranton, Laurent Gobillon, Diego Puga, and Sébastien Roux. The Productivity Advantages of Large Cities: Distinguishing Agglomeration From Firm Selection. *Econometrica*, 80(6):2543–2594, 2012.

[24] Juan Carlos Cordoba. A Generalized Gibrat' s Law. *International Economic Review*, 49(4):1463–1468, 2008.

[25] Victor Couture. Valuing the Consumption Benefits of Urban Density. *Working Paper*, 2015.

[26] Donald R. Davis and Jonathan I. Dingel. The Comparative Advantage of Cities. *NBER Working Papers 20602*, 2014.

[27] Morris A. Davis and François Ortalo-Magné. Household expenditures, wages, rents. *Review of Economic Dynamics*, 14(2):248–261, 2011.

[28] Steve Davis, John Haltiwanger, Ron Jarmin, and Javier Miranda. Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms. *NBER Macroeconomics Annual*, 2007.

[29] Steve Davis, John Haltiwanger, and Scott Schuh. *Job Creation and Destruction*. The MIT Press, 1998.

[30] Jorge de la Roca and Diego Puga. Learning by working in big cities. *Review of Economic Studies*, 84(1):106–142, 2017.

[31] Ryan Decker, John Haltiwanger, Ron Jarmin, and Javier Miranda. Changes in Business Dynamism: Volatility of vs. Responsiveness to Shocks? *Working Paper*, 2016.

[32] Klaus Desmet and Esteban Rossi-Hansberg. Spatial Development. *The American Economic Review*, 104(4):1211–1243, 2014.

[33] Julian di Giovanni and Andrei Levchenko. Country Size, International Trade, and Aggregate Fluctuations in Granular Economies. *Journal of Political Economy*, 120(6):1083–1132, 2012.

[34] Julian di Giovanni, Andrei A. Levchenko, and Isabelle Mejean. Firms, Destinations, and Aggregate Fluctuations. *Econometrica*, 82(4):1303–1340, 2014.

[35] Rebecca Diamond. The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill : 1980-2000. 106(3):479–524, 2016.

[36] Francis X. Diebold and Glenn D. Rudebusch. Measuring Business Cycles: A Modern Perspective. *The Review of Economics and Statistics*, 78(1):67–77, 1996.

[37] Avinash K. Dixit and Joseph E. Stiglitz. Monopolistic Competition and Optimum Product Diversity. *The American Economic Review*, 67(3):297–308, 1977.

[38] Juan J. Dolado, Carlos Garcia-Serrano, and Juan F. Jimeno. Drawing Lessons from the Boom of Temporary Jobs in Spain. *Economic Journal*, 112:F270–F295, 2002.

[39] David Dorn. *Essays on Inequality, Spatial Interaction, and the Demand for Skills*. PhD thesis, 2009.

[40] Gilles Duranton. Labor specialization, transport costs, and city size. *Journal of Regional Science*, 38(4):553– 573, 1998.

[41] Gilles Duranton and Jayet Hubert. Is the division of labour limited by the extent of the market? Evidence from French cities. *Journal of Urban Economics*, 69(1):56–71, 2011.

[42] Gilles Duranton and Diego Puga. Micro-Foundations of Urban Agglomeration Economies. *Handbook of Regional and Urban Economics*, 2003.

[43] Jan Eeckhout. Gibrat's Law for (All) Cities. *The American Economic Review*, 94(5):1429–1451, 2004.

[44] Jan Eeckhout, Roberto Pinheiro, and Kurt Schmidheiny. Spatial Sorting. *Journal of Political Economy*, 122(3):554–620, 2014.

[45] Glenn Ellison and Edward L. Glaeser. Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy*, 105(5):889–927, 1997.

[46] Andres Erosa, Luisa Fuster, Gueorgui Kambourov, and Richard Rogerson. Hours, Occupations, and Gender Differences in Labor Market Outcomes. *NBER Working Papers 23636*, 2017.

[47] Wilfred J. Ethier. National and international returns to scale in the modern theory of international trade. *The American Economic Review*, 72(3):389–405., 1982.

[48] Robert C. Feenstra and David E. Weinstein. Globalization, Markups, and US Welfare. *Journal of Political Economy*, 125(4):1040–1074, 2017.

[49] Ana Fernandes, Klenow Pete, Sergii Meleshchuk, Martha Denisse Pierola, and Andres Rodriguez-Clare. The Intensive Margin in Trade. *Working Paper*, 2017.

[50] Lucia Foster, John Haltiwanger, and C. J. Krizan. Market Selection, Reallocation, and Restructur- ing in the U.S. Retail Trade Sector in the 1990s. *The Review of Economics and Statistics*, 88(4):748–758, 2006.

[51] Lucia Foster, John Haltiwanger, and Chad Syverson. Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? *The American Economic Review*, 98(1):394–425, 2008.

[52] Xavier Gabaix. Zipf's Law for Cities: An Explanation. *Quarterly Journal of Economics*, 114(3):739–767, 1999.

[53] Xavier Gabaix. The Granular Origins of Aggregate Fluctuations. *Econometrica*, 79(3):733–772, 2011.

[54] Luis Garicano, Claire Lelarge, and John Van Reenen. Firm size distortions and the productivity distribution: Evidence from France. *American Economic Review*, 106(11):3439–3479, 2016.

[55] Cecile Gaubert. Firm Sorting and Agglomeration. *American Economic Review, forthcoming*, 2018.

[56] Cecile Gaubert and Oleg Itskhoki. Granular Comparative Advantage. *NBER Working Paper 24807*, 2018.

[57] Elisa Giannone. Skill-Biased Technical Change and Regional Convergence. *Working Paper*, 2018.

[58] Xavier Giroud and Holger M. Mueller. Firms' Internal Networks and Local Economic Shocks. *NBER Working Papers 23176*, 2017.

[59] Edward L. Glaeser, Jed Kolko, and Albert Saiz. Consumer city. *Journal of Economic Geography*, 1(1):27–50, 2001.

[60] William L. Goffe, Gary D. Ferrier, and John Rogers. Global optimization of statistical functions with simulated annealing. *Journal of Econometrics*, 60(1):65–99, 1994.

[61] Christian Gourieroux and Alain Monfort. *Simulation-based Econometric Methods*. OUP/CORE Lecture Series, OUP Oxford, 1997.

[62] Christian Gourieroux, Alain Monfort, and Eric Renault. Indirect inference. *Journal of applied econometrics*, 8 (S1):S85–S118, 1993.

[63] Basile Grassi. IO in I-O: Size, Industrial Organization, and the Input-Output Network Make a Firm Structurally Important. *Working Paper*, 2018.

[64] Nezih Guner, Gustavo Ventura, and Yi Xu. Macroeconomic implications of size-dependent policies. *Review of Economic Dynamics*, 11(4):721–744, 2008.

[65] Bronwyn H. Hall. The Relationship between Firm Size and Firm Growth in the U.S. Manufacturing Sector. *Journal of Industrial Economics*, 35(4):583–606, 1987.

[66] John Haltiwanger, Ron Jarmin, and Javier Miranda. Who Creates Jobs? Small versus Large versus Young. *The Review of Economics and Statistics*, 95(2):347–361, 2013.

[67] Jessie Handbury. Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living Across U.S. Cities. *Working Paper*, 2013.

[68] Jessie Handbury and David E. Weinstein. Goods Prices and Availability in Cities. *Review of Economic Studies*, 82(1):258–296, 2015.

[69] Don Harding and Adrian Pagan. Dissecting the cycle: a methodological investigation. *Journal of Monetary Economics*, 49(2):365–381, 2002.

[70] Richard Hartman. The effects of price and cost uncertainty on investment. *Journal of Economic Theory*, 5(2):258–266, 1972.

[71] Fumio Hayashi. *Econometrics*. Princeton University Press, 2000.

[72] Vernon J. Henderson. The sizes and types of cities. *The American Economic Review*, 64(4):640–656, 1974.

[73] Hugo A. Hopenhayn. Entry, exit, and firm dynamics in the long run equilibrium. *Econometrica*, 60(5):1127–1150, 1992.

[74] Colin Hottman. Retail markups, misallocation, and store variety in the US. *Working Paper*, 2016.

[75] Bradford Jensen and Lori G. Kletzer. Tradable Services: Understanding the Scope and Impact of Services Outsourcing. *Working Paper Series WP05-9, Peterson Institute for International Economics*, 2005.

[76] Kyle Jurado, Sydney Ludvigson, and Serena Ng. Measuring Uncertainty, American. *The American Economic Review*, 105(3):1177–1215, 2015.

[77] Sunwoong Kim. Labor specialization and the extent of the market. *Journal of Political Economy*, 97(3):692– 705, 1989.

[78] Sergey Kokovin, Mathieu Parenti, Jacques-François Thisse, and Evgeny Zhelobodko. Monopolistic Competition: Beyond the Constant Elasticity of Substitution. *Econometrica*, 80(6):2765–2784, 2012.

[79] Nicholas Kozeniauskas, Anna Orlik, and Laura Veldkamp. The Common Origin of Uncertainty Shocks. *NBER Working Papers 22384*, 2016.

[80] Paul R. Krugman. Increasing returns, monopolistic competition, and international trade. *Journal of International Economics*, 9(4):469–479, 1979.

[81] Paul R. Krugman. Scale Economies, Product Differentiation, and the Pattern of Trade. *The American Economic Review*, 70(5):950–959, 1980.

[82] Paul R. Krugman. Increasing Returns and Economic Geography. *Journal of Political Economy*, 99(3):483–499, 1991.

[83] Robert Jr. Lucas. On the mechanics of economic development. *Journal of Monetary Economics*, 22(1):3–42, 1988.

[84] Erzo G. J. Luttmer. Selection, Growth and the Size Distribution of Firms. *The Quarterly Journal of Economics*, 122(3):1103–1144, 2007.

[85] Daniel McFadden. A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration. *Econometrica*, 57(5):995–1026, 1989.

[86] Marc J. Melitz. The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, 71(6):1695–1725, 2003.

[87] Marc J. Melitz and Gianmarco Ottaviano. Market Size , Trade , and Productivity. *Review of Economic Studies*, 75(1):295–316, 2008.

[88] Atif Mian and Amir Sufi. What Explains the 2007-2009 Drop in Employment? *Econometrica*, 82(6):2197–2223, 2014.

[89] Enrique Moral-Benito and Francisco Queiros. Firm Distribution, Market Power and Business Dynamism: Evidence from Spain. *Working Paper*, 2018.

[90] Enrico Moretti. Estimating the social return to higher education: Evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics*, 121(1-2):175–212, 2004.

[91] Enrico Moretti. *Local labor markets*, volume 4. Elsevier B.V., 2011.

[92] Enrico Moretti. Real wage inequality. *American Economic Journal: Applied Economics*, 5(1):65–103, 2013.

[93] Walter Oi. The desirability of price instability under perfect competition. *Econometrica*, 29(1):58–64, 1961.

[94] Gianmarco Ottaviano, Takatoshi Tabuchi, and Jacques-François Thisse. Agglomeration and Trade Revisited. *International Economic Review*, 43(2):409–436, 2002.

[95] Gianmarco Ottaviano, Daria Taglioni, and Filippo di Mauro. Deeper, wider and more competitive? Monetary integration, Eastern enlargement and competitiveness in the European Union. *Working Paper Series 847, European Central Bank*, 2017.

[96] Mathieu Parenti, Philip Ushchev, and Jacques-François Thisse. Toward a theory of monopolistic competition. *Journal of Economic Theory*, 167(C):86–115, 2017.

[97] Benjamin Pugsley and Aysegul Sahin. Grown-Up Business Cycles. *Federal Reserve Bank of New York Staff Reports 707*, 2015.

[98] Sherwin Rosen. Specialization and human capital. *Journal of Labor Economics*, 1(1):43–49, 1983.

[99] Stuart Rosenthal and William Strange. Evidence on the nature and sources of agglomeration economies. *Handbook of Regional and Urban Economics*, 4:2119–2171, 2004.

[100] Steven Ruggles, Alexander Trent, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. Integrated Public Use Microdata Series: Version 5.0. Technical report, 2010.

[101] Albert Saiz. The Geographic Determinants of Housing Supply. *The Quarterly Journal of Economics*, 125(3):1253–1296, 2010.

[102] Edouard Schaal. Uncertainty and Unemployment. *Econometrica*, 85(6):1675–1721, 2017.

[103] Stefan Seth and Stüber Heiko. The Public Release Data of the Administrative Wage and Labor Market Flow Panel. *Economics and Statistics Yearbook*, 2017.

[104] Stefan Seth and Heiko Stüber. Administrative Wage and Labor Market Flow Panel (AWFP) 1975-2014. *FAU Discussion Papers in Economics*, 2017.

[105] Ina Simonovska. Income Differences and Prices of Tradables: Insights from an Online Retailer. *Review of Economic Studies*, 82(4):1612–1656, 2015.

[106] Adam Smith. *An Inquiry into the Nature and Causes of the Wealth of Nations*. London: Printed for W. Strahan, and T. Cadell, 1776.

[107] Alexandra Spitz-Oener. Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure. *Journal of Labor Economics*, 24(2):235–270, 2006.

[108] Chad Syverson. Product Substitutability and Productivity Dispersion. *The Review of Economics and Statistics*, 86(2):534–550, 2004.