The Startup Multiplier[*]

Conor Walsh

Columbia University

March, 2023

Abstract

In US Census microdata, local employment growth is very persistent. The new firm margin plays the dominant role in this persistence. Firms do not exit at different rates across space, and the composition of firm types does not change as areas grow. I embed these findings in a model of spatial growth, and estimate that the new firm margin can act as a significant multiplier for local shocks. New firms raise labor demand, increase local spending and attract new workers, who in turn increase the demand for new firms. This process working in reverse appears crucial to explain the decline of the US Rust Belt.

Keywords: Growth, Employment, Firm Dynamics, Startups

JEL Codes: R12, L26, R32, R13

1. INTRODUCTION

There is a large amount of variation in long-run employment and wage growth at the local level.[1] Some places grow quickly for many years, and some stagnate. This paper offers new evidence that a "startup multiplier" is crucial to understanding these persistent differences.

I begin by using U.S. Census microdata to document three facts about firm creation and local growth. *First*, places with high initial startup rates see long periods of growth in employment and wages. *Second*, most of this growth is accounted for by the continued entry of new firms. Startup rates are very persistent, and local incumbents uniformly shed employment as a group. *Third*, as places grow larger, they do not become more likely to host fast-growing firms. Firms and establishments in dense areas do not grow faster, exit faster, or see divergence in wages or sales per worker. The firm lifecycle is mostly invariant across space.

Rationalizing persistent differences in firm creation is then the primary task in explaining local growth. In this paper, I study the role new firms play in propagating local shocks. I present a simple dynamic model of monopolistic competition and specialization in space. In the model, a new firm brings a new idea or product to the area, which raises local productivity. This increases wages, and attracts more workers to the area. In turn, these new workers increase demand for all local firms, spurring the creation of new entrants.

The feedback between firm creation and labor mobility helps explain how both growth and entry at the local level can be so persistent. To an extent, local booms can feed themselves. New firms create externalities for other entrants through their effect on local demand. In this way, shocks to locations are multiplied and propagated by the process of firm creation.

As an example, consider the growth of Deschutes County, in central Oregon. Formed in 1916, Deschutes became known in the early 1990's for its outdoor lifestyle and environmental amenities, attracting a host of new residents.[2] A larger population needed new restaurants, retail outlets, and medical services, all of which saw a 300% increase in new establishments between 1990 and 2015. All this new business activity raised the demand for professional and technical services: architects, lawyers, and accounting firms saw a six-fold increase in new business formation, with finance and insurance not far behind.

---

[1]From 1975-2015, mean cumulative employment growth at the commuting zone level was 108%, with a standard deviation of 105%. More detail is given in Appendix A.1.

[2]The population grew from 76,000 in 1990 to 174,000 in 2015. The Economist (2007, Jan 25th) credits the end of the county's logging industry for the initial boom in those searching for outdoor lifestyles.

Construction, specialized manufacturing, and transportation all saw entry rates far above state and national trends. Wages rose abruptly, and employment growth rates among the highest in the nation have been sustained for many years.

However, the same process can run in reverse, particularly when a place is exposed to macroeconomic shifts. Gary, Indiana, was incorporated by the U.S. Steel Company in 1906, and experienced strong population growth for decades on the back of rising steel output. In the 1970's, falling global demand for steel and heightened foreign competition led to large reductions in local steel-making employment (Tarr, 1988). This decline was accompanied by a collapse in business formation across a broad range of industries. The population began to shrink, and Gary continued to experience losses in employment and population long after steel-making employment stabilized.

The theory I develop provides a qualitative account of such persistent, creation-fueled propagation. I then estimate the strength of propagation using microdata for the full U.S. geography. In order to estimate the model I face a central identification challenge. Firms are forward looking, and unobserved shocks that affect future local growth will also affect their startup decisions. This confounds a naive causal interpretation of the correlation between entry and future growth.

To gain identification, I expand on Gary's experience, and isolate a series of demand shifters for firm creation driven by aggregate structural change. The U.S. has seen a large decline in manufacturing employment in recent decades. This decline had a significant local dimension: areas specialized in manufacturing activity in 1975 saw much slower growth in *total* employment over the subsequent 40 years. I show new evidence that these regions also saw reduced rates of firm creation across all sectors, and particularly in sectors that catered to local demand.

I use the structure of the model to provide a set of moment conditions for estimation, exploiting a counterfactual prediction for local demand were the only changes since 1975 to have come from an aggregate shock to manufacturing. These conditions are used to estimate the key structural parameters of the model (see Adao et al. (2019) , Allen et al. (2019), and Faber and Gaubert (2019) for similar strategies).[3]

There are two key results. First, the estimated model attributes an important role for the startup multiplier in explaining local employment and wage growth. The effect of a 1% shock to local productivity is multiplied almost two-fold over the medium term by startup

---

[3]This strategy also relates to the widespread use of Bartik industry shocks in reduced form work. However, it is distinct in that it fully specifies the mechanisms by which aggregate structural change impacts the macroeconomy and local activity, and is transparent about what the moment conditions require for identification of structural parameters.

dynamics for both employment and wages. Moreover, there is substantial feedback between labor mobility and startup dynamics: shutting down movement of people across space lowers the local wage growth attributable to firm creation by almost half. Newcomers generate local demand that spurs new firm creation, productivity and wage growth, and more labor demand.

Second, I apply the model to understand the decline of the US Rustbelt. The model reveals that a start-up deficit has amplified the employment gap between the Rustbelt commuting zones and the growing regions of the U.S. by almost 50% in the last four decades. This indirect effect has been almost as important as the direct shock to manufacturing that started the decline of the Rustbelt. The relative decline of employment growth and the resulting distressed state of former manufacturing areas is in part due to a collapse in business formation in the service sector, which has greatly multiplied the initial shock.

Related Literature.

A small number of papers have studied the contribution of firm creation to local growth (see Glaeser et al. (2010); Gourio et al. (2016); Carlino and Drautzburg (2017)). Glaeser et al. (2015) instrument for the presence of startups with the locations of historical mines, and find an important role for startups in explaining cross-sectional patterns of growth. This paper is also related to the work of Peters (2016), who considers the relationship between shifts in factor supplies and firm entry using data on the large-scale reallocation of German refugees after World War 2. A similar mechanism is at work there: inflows of people into a location will act as a spur to the demand for new firms, and via increasing returns raise wages in the long run.

The dynamic theory I propose has a long tradition in economic geography. Myrdal (1957) discusses a process of "circular causation", where local shocks are amplified by investment and local demand, in what is perhaps the earliest verbal description of the core idea in this paper. A similar logic is at work in the canonical geography model of Krugman (1991). However, explicit modeling of growth dynamics has been limited outside of stylized, stationary environments[4] (notable exceptions are Desmet et al. (2018) and Nagy (2017), who develop rich models of spatial innovation and production). This has constrained the extent to which one could discuss the dynamic effects of firm creation, and just as importantly, the ability to measure these effects with new micro-data sets. The theory here models non-stationary environments with a rich geography, and is more suited to empirical analysis.

---

[4]See influential early work in Black and Henderson (1999), Fujita and Thisse (2003) and reviews in Duranton and Puga (2014). A recent paper by Greaney (2019) studies the distributional consequences of local growth in quantitative settings with asset accumulation.

My work also relates to a broad literature dealing with the decline of manufacturing in the U.S. (see Fort et al. (2018) for a review). While incorporating aggregate structural change into the model disciplines the quantitative estimates, it also provides novel insights into how structural change operates. The firm creation margin amplifies and propagates the local decline of manufacturing, hastening the reallocation of workers across space. This indirect effect is common in stories in the popular press, but absent from most economic models.

2. DATA: STARTUP DYNAMICS AND LOCAL GROWTH

In this section, I document three central facts about local growth and firm creation. These facts show that variation in the firm creation margin is central in accounting for differences in growth across space.

The analysis uses the U.S. Census' Longitudinal Business Database (LBD), which contains information on every private employer establishment in the U.S. since 1978. In particular, it contains information on annual employment, annual payroll, 6-digit industry and the address of each establishment, along with longitudinal links at the firm level that allows me to track a record of ownership. I add new information on sales from the Business Register to this dataset, following Moreira (2015) and Haltiwanger et al. (2016). The Business Register is the Census' master establishment list for every establishment in the U.S., and contains information on the revenues of the parent companies from IRS administrative tax data. These sales are only available at the firm level, and I discuss in the Appendix several methods of apportioning sales among establishments for multi-establishment firms.

I focus on two primary local geographic units throughout the paper: commuting zones and counties.[5] Both form a complete partition of the U.S. geography, and cover large and small cities, towns, and rural areas (in contrast to administrative boundaries such as Metropolitan Statistical Areas). The documented facts and results do not depend materially on the choice of geography.

---

[5]First constructed by Tolbert and Sizer (1996), commuting zones are now widely used in empirical work for studying economic activity across space (see e.g. Autor and Dorn (2013)). They are defined as groups of contiguous counties which see high proportions of cross-county commuting flows between the group. I use the 703 commuting zones defined by the U.S. Census for the year 2000 throughout, though results are robust to using 1980 and 1990 borders.

Figure 1: Future Employment Growth and Initial Start-up Percentage
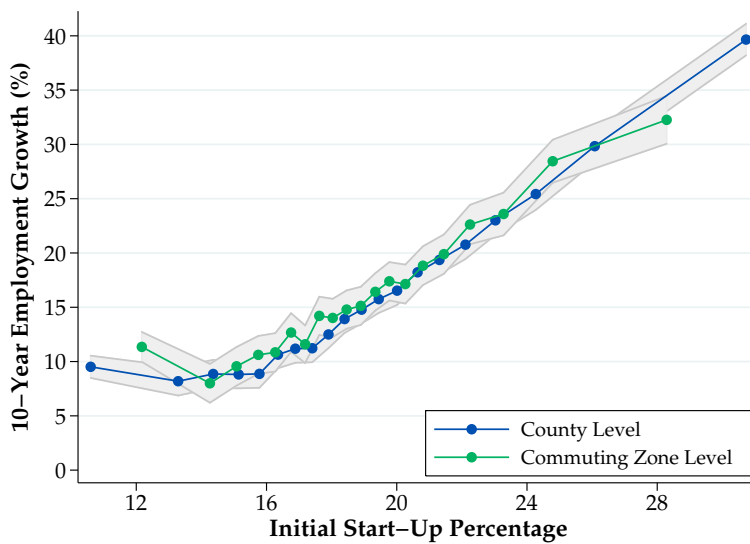
Note: This figure plots the coefficients of regressions of 10-year county-level (blue) and commuting zone-level (green) employment growth on dummies for 20 quantiles of the initial start-up percentage, for the period 1980-2015. Employment growth is calculated as total percentage change in headcount at establishments in the area between any two years. Start-up percentage is the fraction of local establishments of age zero or one. 95% confidence intervals in grey. $N = \{9,200; 2,100\}$ for the county and commuting zone level respectively, where these counts have been rounded to accord with U.S. Census disclosure rules.

Fact 1: *Places with high initial start-up rates experience high long-term growth in employment and wages.*

I first use the LBD to calculate local-level measures of employment and total wage bill by aggregating establishment-level headcount and payroll in each year. This gives both a county and commuting zone panel for the full U.S. geography. In Figure 1, I regress 10-year forward log changes in employment at the local level on 20 quantiles of the startup percentage at year t, for the decades beginning in 1980, 1990 and 2000. I define the start-up percentage as the fraction of establishments in that location who are of age 0 or age 1 in year t. Employment growth varies substantially among these quantiles. In areas with high numbers of startups, employment growth is predicted to be almost 30% over the subsequent ten years. This drops to around 10% for areas with a low proportion of startups.

In Table 2 in Appendix A.2 I reestimate this relationship with various controls, including local industry shares and local fixed effects. I then restrict the analysis to true new firms by excluding new establishments from existing, multi-establishment firms. Furthermore, I show an estimate of the conditional distribution using quantile regressions. The conclusion is unchanged: there is a strong positive relationship between future employment

Figure 2: Entrant Contributions to 10-Year Employment Growth
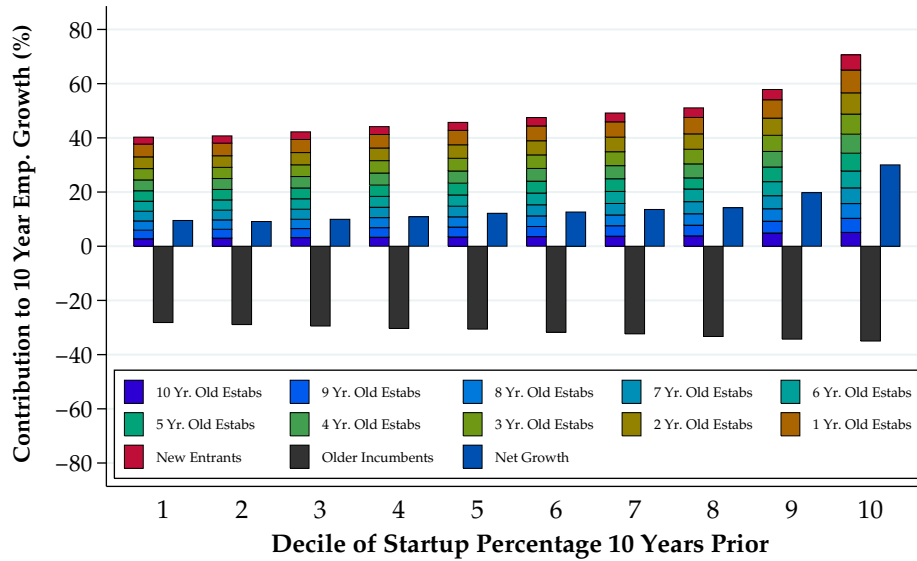
Note: This figure plots the contributions to future 10-year employment growth at the commuting zone level by successive cohorts of entrants, as well as current incumbents. Areas are assigned in each year to 10 deciles of the startup percentage over all years. Decomposition is as in equation (1). Years are from 1990-2015 (for example, 10-year growth for 1990 is taken for the period 1980-1990). $N = \{24,000\}$, where this count has been rounded to accord with U.S. Census disclosure rules.

growth and the initial start-up percentage.

I finally show that the same is true for average wage growth in Table 3 in Appendix A.3. Over a ten year horizon, commuting zones with an initial startup rate in the upper quartile of the distribution will see average wage growth that is 8 percentage points faster than commuting zones in the lower quartile.

Similar findings have also been documented by Faberman (2011), who uses a more restricted dataset from the BLS and focuses on variation in employment growth at the MSA-level. My analysis here is complementary, extending the results to a more comprehensive set of geographies and firms, and including measures of wages.

Fact 2: *All growth in employment comes from the continued entry of new establishments.*

One question that immediately arises in connection with Fact 1 is how much the initial startups are themselves contributing to the 10-year periods of growth documented above. I first answer this question in an accounting sense. I decompose 10-year forward em-

ployment growth rates at the commuting zone level into separate contributions from each cohort of new establishments that enters during these ten years, as well as a contribution from incumbent establishments. Formally, we can write the percentage change in employment at the commuting zone level between year $t - 10$ and t as

$$(1) \qquad \frac{Emp_{i,t} - Emp_{i,t-10}}{Emp_{i,t-10}} = \frac{I_{i,t} - I_{i,t-10}}{I_{i,t-10}} + \sum_{a \in \{0,1,..,10\}} \frac{N_{i,t}^a}{Emp_{it-10}},$$

where $N_{i,t}^a$ is the total employment at time t in commuting zone i of firms of age a, and $I_{i,t}$ is the total employment at time t in commuting zone i of incumbents who existed 10 years prior to time t. By this definition, incumbents may add or subtract from 10-year employment growth, but since new firms do not exist 10 years prior, their employment at the end of the period only adds to total growth. I then average each of these contributions to employment growth within deciles of the start-up percentage at the commuting zone level 10 years beforehand. The results are plotted in Figure 2.

Two key points are apparent. First, areas that have high startup rates in a given year continue seeing strong employment growth from new cohorts of startups in each one of the subsequent 10 years. While the contribution of any one year's startups to 10-year employment growth is minor, they accumulate over time. This illustrates a key feature of the startup rate at the local level: it is persistent, and this persistence generates large differences between high and low startup areas over time. To highlight this point, in Appendix A.4 I show estimates of the autocorrelation of the local startup rate. For commuting zones, this yields values of 0.70 at a 1 year lag and 0.54 at a 10-year lag.

Second, incumbent firms who existed at the start of the period uniformly subtract from 10-year growth as a group. While this masks a large amount of heterogeneity at the micro level (some of these incumbent establishments exit, some grow fast and some plateau), what matters for employment growth is the total change in employment for incumbent firms. This is robustly negative, recalling Haltiwanger et al. (2013).

Just as importantly, this negative contribution varies only slightly over the deciles of the startup fraction 10 years beforehand. Essentially, most of the difference between high startup areas and low startup areas is that the former continue seeing strong employment growth from the creation of new establishments, while incumbent employment shrinks at roughly the same rate across areas. In particular, the exit rates of incumbent establishments across space are unrelated to both the growth rate of the local area, and the initial startup percentage. In Figure 3 I plot the average 10-year employment growth of com-

Figure 3: Local Growth and Local Exit
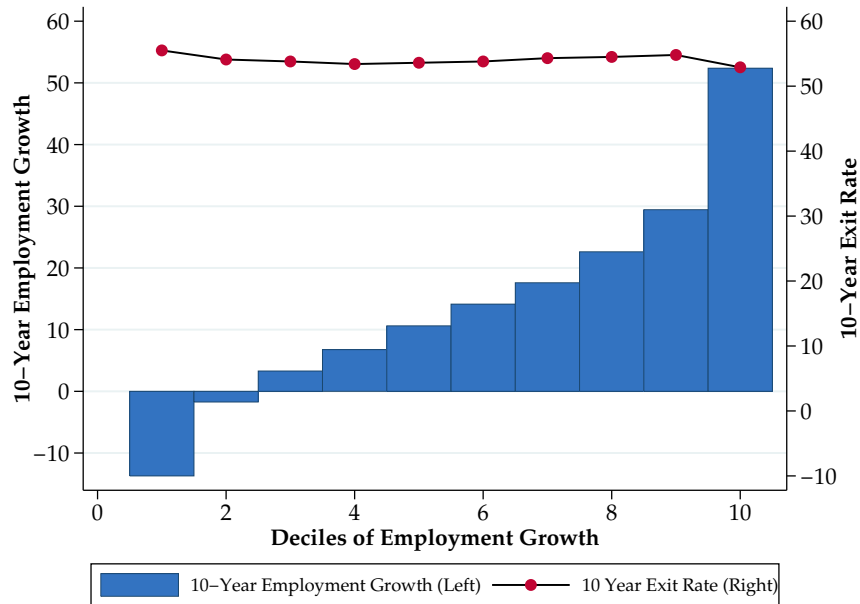
Note: This figure compares 10-year employment growth at the commuting zone level to the 10-year exit rate of establishments. Deciles of employment growth split commuting zones into 10 ranked bins of 10-year employment growth. The 10-year exit rate is the fraction of establishments present in the commuting zone at year t which are no longer present in year $t + 10$. Years are from 1990-2015 (for example, 10-year growth for 1990 is taken for the period 1980-1990). $N = \{24,000\}$, where this count has been rounded to accord with U.S. Census disclosure rules.

muting zones average within 10 deciles of Employment growth. The dispersion is large: 10-year employment growth ranges from over 50% in the highest decile, to -10% in the lowest. Nevertheless, across all 10 cells, the 10-year exit rate of establishments is essentially constant, ranging between 53% and 55%. Establishments exit at almost the same rate across space.

In unpacking these dynamics, one might suspect that as areas become more populated, they attract or produce different, potentially faster-growing establishments and firms. Jacobs (1970) argued in her seminal work that large cities facilitate idea spillovers, and most novel ideas come from dense urban areas. This hypothesis is canvassed by a variety of influential work in urban economics, including Glaeser et al. (1992), Duranton and Puga (2001) and Davis and Dingel (2019). One might then expect that firms and establishments born in a large city like New York would outperform those born in Cleveland OH, conditional on industry.

Fact 3: *The firm lifecycle is invariant across large and small areas*

I break this statement into two parts. I first show that neither firms nor establishments grow their employment faster in larger areas, conditional on survival. I then show that differences in survival patterns are negligible between large and small areas. In the Appendix I show that average wages and average sales mirror employment patterns.

Fact 3a: *Conditional on survival, establishments and firms grow no faster in larger areas.*

To show this, I calculate employment growth as log changes in employment from the previous year, for the full sample of establishments in the LBD. Note that this is only observed if the establishment survives from one year to the next. I then estimate the following equation:

$$(2) \qquad \Delta y_{it} = \sum_{a=1}^{A} \gamma_a \mathbb{I}_{Age_{it}=a} + \sum_{a=1}^{A} \beta_a \mathbb{I}_{Age_{it}=a} \times log(Pop_i) + Ind_i + \mu_t + X'_{it}\delta + \epsilon_{it},$$

where Δy_{it} is log change in employment of the establishment from year $t-1$ to t, and $\mathbb{I}_{Age_{it}=a}$ is a full set of age dummies, up to a maximum age A which I set at 15 years.[6] The vector of coefficients $\{\gamma_a\}_{i=1}^{A}$ documents how establishment growth varies with age, conditional on survival.

To measure systematic growth differences across space, the second set of coefficients $\{\beta_a\}_{i=1}^{A}$ measures the interaction of age with the log of population at the local level ($log(Pop_i)$) in the year the establishment hires its first employee. It thus describes whether establishment growth rates vary systematically with area population size across the firm life cycle.[7]

Finally, Ind_i are industry fixed effects at the 4-digit NAICS level, exploiting the longitudinally consistent industry codes for the LBD made available by Fort and Klimek (2018). These fixed effects capture systematic mean growth rate differences across detailed industries. μ_t are aggregate time fixed effects and X_{it} are controls for multi-establishment firm and state fixed effects.

The results are shown in Table 4 in Appendix A.5. The first column restricts β_a to be equal for all ages, and finds no significant differences in establishment growth rates across space. Figure 4 uses these estimates to plot the growth rates by age for establishments

---

[6]After 15 years, more than 80% of the establishments in the LBD have exited. Varying the maximum age *A* does not change the conclusions of this section.

[7]Using population density (people per square mile) in the local area instead of population size does not affect the conclusions of this section.

Figure 4: Establishment Employment Growth Rates by Age and Location Size
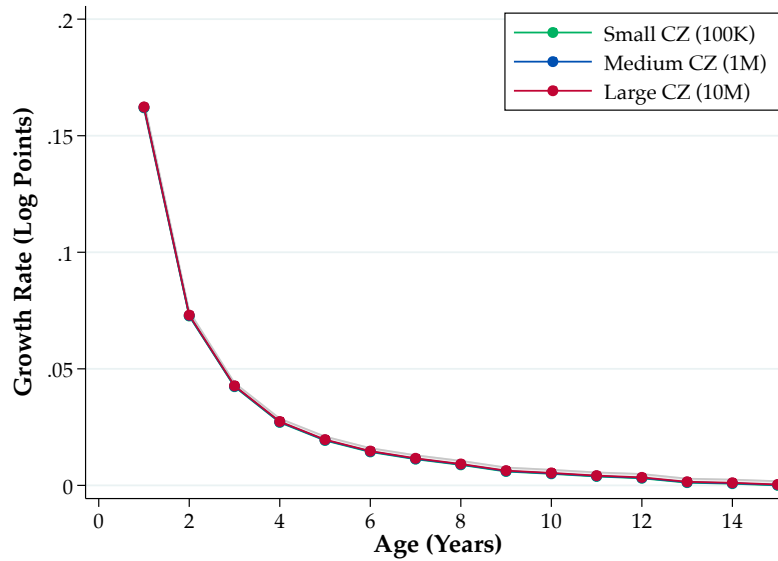

Note: This figure plots the estimated employment growth rates of establishments by age above the baseline of age 15 in three different categories of population size, using the coefficients obtained from estimating equation (2) in the LBD. 95% confidence intervals in grey. $N = 104,000,000$, where this count has been rounded to accord with U.S. Census disclosure rules. For reference, the average growth rate of 15 year old establishments conditional on survival across industries is -1.3%.

Figure 5: Exit Rates by Age and Location Size
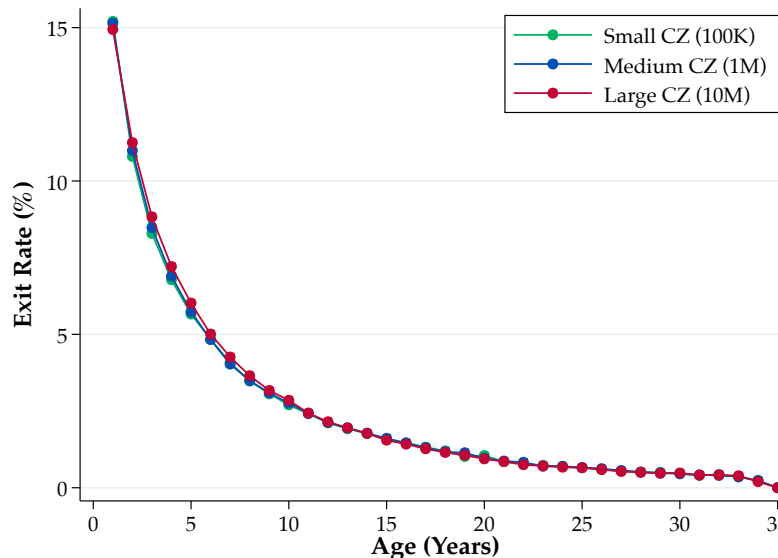
Note: This figure plots the (unconditional) probability of exiting by age for all establishments born between 1980 and 1995, split by being born in three different categories of CZ population size. These probabilities are increments of Kaplan-Meier survival functions estimated for 10 deciles of commuting zone size (full output is reported in Appendix A.8) The three categories reported here correspond to decile 2, 5 and 10. $N = 70,000,000$, where this count has been rounded to accord with U.S. Census disclosure rules.

located in three size classes of commuting zone. Growth in employment is rapid for young firms, before quickly leveling out with age, recalling results presented in Haltiwanger et al. (2013). However, growth is not correlated with the population of the area in which the establishment is located.

The second and third columns of Table 4 in Appendix A allow the interaction terms in β_a to vary by age. While this introduces some noise into the parameter estimates, the estimated growth rates are not systematically higher in denser areas. Figure 23 in Appendix A uses these estimates to again plot the growth rates by age for establishments located in three size classes of commuting zone. No systematic difference over an establishment lifetime is detectable between large and small areas, conditional on survival.

The analysis for firms is complicated by the fact that firms can own several establishments, and exist at multiple points in space. Defining the current location of such multi-establishment firms is not straightforward. Though relatively few in number, these firms account for large fractions of employment and output, and cannot be ignored.

Nevertheless, we can ask a related question. Most multi-establishment firms begin life with a single establishment, allowing a well-defined notion of the birthplace of the firm. I test whether firm growth over the lifecycle (conditional on survival) is correlated with the population of the firm's birthplace. I follow standard practice by defining firm age as the age of the oldest establishment the firm owns at birth, from which point the firm ages naturally. The results and further detail are given in Appendix A.5. As with establishments, I find that firms born in more populous locations do not grow systematically faster over their lifecycle.

Fact 3b: *Differences in exit rates by age across large and small areas are negligible.*

To show this, I examine the exit decisions of establishments non-parametrically as a function of age. I take all establishments born between 1980 and 1995, and follow them until 2015. I estimate Kaplan-Meier survival functions for these establishments, split by the local size of the area in which the establishment is born. Exit rates by age are taken from the increments in the estimated survival functions.

In Figure 5, I plot the exit rates for establishments, split by the same size classes of commuting zone as Figure 4. These are almost identical over the sample period; exit rates are on average less than 0.1% higher for each age for establishments in a commuting zone above the median size. In Appendix A.8 I show the estimated survival functions, as well as further detailed breakouts by size category. The conclusion continues to hold: entry rate differences across areas of different population sizes are negligible. Over a 20 year

Figure 6: Establishment Employment By Age and Location Size
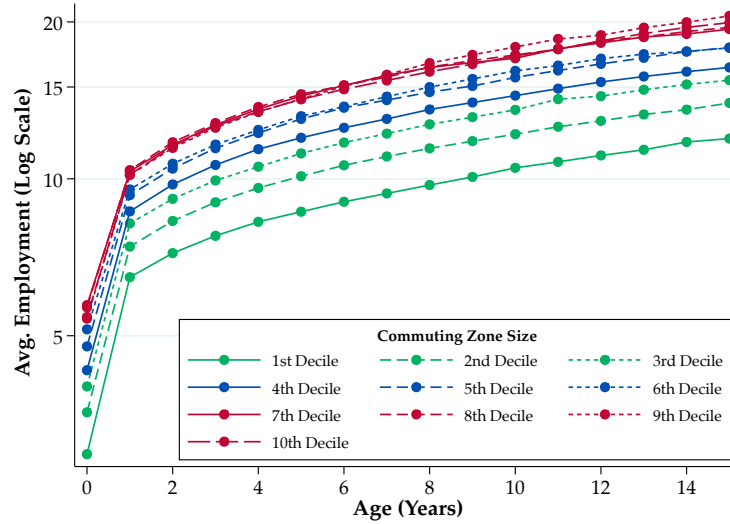
Note: This figure plots average employment by age for establishments, split by deciles of the population of the commuting zone in which they were born. Commuting zone decile is calculated using data for commuting zone populations from the NBER intercensal county estimates and commuting zone definitions from the U.S. Census for the year 2000. $N = 125,000,000$, where this count has been rounded to accord with U.S. Census disclosure rules.

time span, differences in exit probability are less than 2% between the most populous areas and the least, out of an average exit rate of 81%.

Facts 3a and 3b combined suggest establishment scale within industries evolves in parallel across areas of different sizes. A sense of this can be gained from Figure 6, where I plot the log of average establishment employment by age across 10 deciles of commuting zone size.[8] Without additional controls, it indeed appears as if growth in establishment scale by age is unrelated to the population of the local area (though establishments are systematically larger in more populous areas). I show this more rigorously for employment, average wages and sales per worker in Appendix A.6, using the same controls and specification as in equation (2).

3. A MODEL OF FIRM CREATION AND LOCAL GROWTH

I now describe a simple model of local growth, where firm creation generates persistent growth differences across areas. In the model, firms are created in response to local profit opportunities. A new entrant pays a fixed cost to bring a new product to an area, and in

---

[8]Each of these deciles contains 10% of the U.S. population in the year 2000. A summary is provided in Table 5 in Appendix A.7.

doing so raises local productivity. High productivity increases labor demand and wages locally, and attracts new workers to the area. In turn, higher spending drives demand for new local firms.

This logic generates a dynamic interaction between the decisions of entrants and the location decisions of workers, causing persistent differences in growth across space. The strength of propagation depends crucially on the availability of land supply. Places where it is easier to build new housing for workers will see greater inflows of workers in response to an increase in wages, and hence a stronger impact on local demand.

Though stylized, the model is rich enough to be estimated directly on the microdata, and I use it to quantitatively assess the importance of firm creation and new varieties to local growth. In this section I present the model, and in the following one I discuss the estimation strategy.

3.1 General Environment

Time is continuous. There are a discrete number of fixed locations $j = \{1, ..., J\}$. Regions are characterized by an amenity level $A_{j,t}$ and a productivity $B_{j,t}$.

Consumer Preferences. There are three classes of agents in this economy. First, there is a mass L_t of workers. These workers have preferences over final goods and housing services, given by a Cobb-Douglas period utility function

$$(3) \qquad U_{j,t}^{W} = C_t^{\alpha} H_t^{1-\alpha} A_{j,t} \,,$$

where C_t is consumption of a homogenous final good, H_t is housing services and $A_{j,t}$ is a location specific amenity. Workers are freely mobile in each instant, and so choose the location that maximizes instantaneous utility. Workers cannot save. They supply their labor inelastically for a competitive wage $w_{j,t}$ when in location j at time t.

Second, in each location there is a unit mass of identical capitalists. These capitalists cannot move. They consume only the final consumption good, and have intertemporal preferences given by

$$(4) \qquad U_t^{E} = \mathbb{E}_t \int_t^{\infty} e^{-\rho s} \frac{(C_s^{E})^{1-\gamma}}{1 - \gamma} ds,$$

for $\gamma \geq 0$. These capitalists own all the firms in their location, and receive dividends from their operations. They also finance the creation of new firms in their location in order to

14

maximize their intertemporal utility. They may not invest in other locations.[9]

Third, there is a mass of landlords living in each location. They own an amount of housing real estate, h_j, and an amount of commercial real estate K. They too cannot work or move, and consume only the income from their property holdings through purchases of the final good.

Production. In each location, there are firms producing intermediate inputs used in producing the final good. I assume each firm produces a single intermediate variety, and a firm only exists in one location fixed at the time of birth. After entering, firms cannot move. Final good output is produced by competitive firms in each location, which aggregate all intermediate inputs in that location according to

$$(5) \qquad Y_{j,t} = \left(\int_0^{N_{j,t}} y_{j,t}^{\frac{\sigma-1}{\sigma}}(i)di \right)^{\frac{\sigma}{\sigma-1}},$$

where $y_{j,t}(i)$ is the use of variety i in location j at time t, and $N_{j,t}$ is the number of intermediates firms in the local economy. This final good is freely traded across locations, and so commands the same price everywhere. It serves as the numeraire.

Each intermediate firm i is characterized by an idiosyncratic efficiency term $\tilde{z}$. Its productivity depends both on a local productivity shifter $B_{j,t}$ and its efficiency $\tilde{z}$. It produces according to $q_{j,t}(i) = B_{j,t}\tilde{z}_t(i)l(i)$, where $l(i)$ is employment of labor. Standard aggregation results for monopolistic competition imply that the wage in location j depends on three local state variables, according to

$$(6) \qquad w_{j,t} = \frac{\sigma-1}{\sigma} B_{j,t} \left(\bar{Z}_{j,t} N_{j,t} \right)^{\frac{1}{\sigma-1}},$$

where $\bar{Z}_{j,t}$ measures the average efficiency of the firms in location j at time t. For ease of notation and without loss, I work with a scaled version of firm efficiency, defined as $z \equiv \tilde{z}^{\sigma-1}$. Then $\bar{Z}_{j,t}$ is defined as $\bar{Z}_{j,t} \equiv \int_0^\infty z dM_{j,t}(z)$, where $M_{j,t}(z)$ is the probability measure of firms with efficiency z at time t in location j.

Note from equation (6) that there are increasing returns to scale in the number of local firms $N_{j,t}$, due to the specialization effect common to many economic geography models (see e.g. Krugman (1991)).

Firm Dynamics. Starting a firm requires paying a fixed set up cost in the form of purchasing a building. Upon paying the cost, the new firm will draw its efficiency from a

---

[9]In Appendix C I consider an alternative formulation with a representative capitalist who owns all firms in all locations. The long run implications are identical, but short run dynamics feature some differences.

distribution $G(z)$ that is independent of location, and commence operation. The building cannot be sold once the firm begins operation, and as such there is no endogenous exit choice to consider.

Firms grow over time through shocks to their idiosyncratic efficiency. If a firm receives a shock, they see their efficiency change proportionally from z to $\Delta_i z$, where $\Delta_i \in \{\Delta_u, \Delta_d\}$. Their efficiency either improves when receiving $\Delta_u > 1$, in which case they expand production, or falls with $\Delta_d < 1$. Shocks arrive at a constant Poisson rate $\phi_i \in \{\phi_u, \phi_d\}$. Lastly, firms die at constant rate δ, independent of size or location. The mass of local intermediates firms evolves according to

$$\dot{N}_{j,t} = N^E_{j,t} - \delta N_{j,t}$$

where $N^E_{j,t}$ is the flow of new entrants.

Land Development. In this model, a location consists of two non-overlapping zones: a commercial district and a residential district. Production by firms takes place in the commercial district, and workers live in the residential district.

A building in the commercial district is required in order to operate an intermediates firm. New buildings can be produced by a competitive construction sector by combining the final good with commercial land, according to $N^E_{j,t} = (X^E_{j,t})^{1-\tilde{\zeta}} K^{\tilde{\zeta}}$ where $N^E_{j,t}$ is the flow of new buildings for entrants (and hence equal to the mass of new entrants in equilibrium), $X^E_{j,t}$ is use of the final good in commercial construction, and K is a fixed amount of commercial land, common across areas. This introduces a congestion friction into the construction of new buildings in the commercial district; when demand for new buildings is high because the city is undergoing a boom, the cost of entering will be higher. It also introduces a dynamic element to entry decisions, since the mass of firms cannot immediately adjust to local shocks, and investment takes time to play out.

Workers rent housing in the residential district.[10] Housing services are produced competitively using the final good and the fixed land supply h_j owned by the residential landlords, according to $H_{j,t} = c_{v_j}(X^H_{j,t})^{1-v_j} h_j^{v_j}$, where c_{v_j} is a combination of model constants, and $X^H_{j,t}$ is use of the final good in residential services. As with commercial land, the cost of renting land in region j will depend on how difficult it is to expand development of this land when demand rises. This depends on the parameter v_j, since residential land h_j is in fixed supply. A higher value for v_j will result in a less elastic supply of housing services in location j.

---

[10]I model residential land as a produced service so as to abstract from durability in the residential sector. This can be accommodated at some cost in complexity.

3.2 Equilibrium Characterization

I begin by characterizing the behavior of an intermediate-producing firm. I assume that capitalists have perfect foresight over the paths for location fundamentals, so that the evolution of economy wide state variables can be captured through the dependence of the value of the firm on time and location only. The value of the firm with efficiency z, at time t and in location j, denoted $V_{j,t}(z)$, is given by the HJB

$$(7) \qquad (r_{j,t} + \delta)V_{j,t}(z) = \pi_{j,t}(z) + \dot{V}_{j,t}(z) + \sum_{i \in \{u,d\}} \phi_i(V_{j,t}(\Delta_i z) - V_{j,t}(z)),$$

where $r_{j,t}$ is the interest rate on the local capitalists portfolio. I provide a full treatment of the determination of this interest rate in equilibrium in Appendix **??**. The discounted value of any firm involves three separate parts. First is the flow profits from operation $\pi_{j,t}(z)$. Second is the appreciation in value that results from evolving state variables in the economy, including changing wages and populations across all regions. These state variables enter only through their impact on current local profits. Last is the appreciation in value that arises from stochastic improvements of firm productivity.

Under the CES production structure for the final good, profits are a constant fraction of revenues. We can exploit the fact that sales of the final good within each location have to equal the income flowing to labor and intermediate firm profits, since the final good sector is competitive. This allows us to write the profits of the firm with mean efficiency $\bar{Z}_{j,t}$ as

$$(8) \qquad \pi_{j,t}(\bar{Z}_{j,t}) = \frac{w_{j,t}L_{j,t}}{(\sigma - 1)N_{j,t}}.$$

This is a classic result in the standard model of monopolistic competition. Independent of demand, firms charge a constant markup over marginal cost, which itself is constant due to the linear production function. As a result, when spending rises firms will respond by increasing output, making a constant profit on each extra unit sold. Profits are then increasing in total spending, and decreasing as the number of firms in a location rises due to competition. Moreover, the profit of any firm is linear in efficiency z, and can be written $\pi_{j,t}(z) = \pi_{j,t}(\bar{Z}_{j,t})\bar{Z}_{j,t}^{-1}z$.

Now given that the probability of death is also independent of z, the value of the firm takes a particularly tractable form. Assuming that the value of the firm is bounded (which is proved in Appendix B.1 under simple conditions on primitives), we have the following result:

Lemma 1. *The value of the firm is linear in z at all times, and in all locations.*

This result aids substantially in characterizing the equilibrium. As we will see below, it implies that it is unnecessary to keep track of the distribution of firm-level productivity over time when computing equilibrium dynamics. Instead, the value at the mean entrant efficiency in a location is a sufficient statistic to characterize the entry and investment decisions of firms.

These decisions will depend on the profitability of the average firm in any given location. To see how this is determined, note that from (8), profits depend on a balance between two local scale variables. The first is the total wage bill for production, $w_{j,t}L_{j,t}$, which is a measure of local economic activity and spending. The second is the number of firms $N_{j,t}$ in that location. Places that see growth in local spending, perhaps due to a growing population, will incentivize the creation of new firms. These new firms can profitably use the local labor force, due to the decreasing returns to scale each firm faces on the demand side. Entry of these new firms will lower spending per firm, decreasing the attractiveness of entry in the future and stabilizing the local economy.

More precisely, in equilibrium, free entry will require that the cost of creating a new firm equals the expected value of creating a firm, or

$$(9) \qquad\qquad \tau(N_{j,t}^E)^{\zeta} = \int_0^{\infty} V_{j,t}(z)dG(z),$$

where $\zeta \equiv \frac{\tilde{\zeta}}{1-\tilde{\zeta}}$ and τ is a combination of model constants, and I have derived the cost of creating a new firm on the LHS of equation (9) from market clearing for commercial land in location j at time t. Thus for free entry to hold, areas with high profitability and high firm values will induce higher rates of entry, and drive up the costs of building construction.[11] an

In the simple linear setting adopted here, we can solve analytically for the entry decisions of firms in every location. In combination with Lemma 1, free entry will pin down the value of the firm at the mean entrant efficiency, such that $\tau(N_{j,t}^E)^{\zeta} = V_{j,t}(\bar{z}^E)$ and $\bar{z}^E \equiv \int_0^{\infty} zdG(z)$ is the mean efficiency of entrants. In addition, free entry will ensure that appreciation in this value depends only on changes in the mass of entrants through time. Thus, while new firms have to solve a potentially complex forecasting problem involving

---

[11]Note that in this model the free entry condition is never slack, and entry is always positive. Since firms pay no fixed operating costs, expected profits over the firm's life must always be positive. However, as entry goes to zero, the cost of creating firms also goes to zero, so in equilibrium there will always be some firm creation. Net firm creation (after accounting for exit), on the other hand, may be negative.

the entire future evolution of activity across space, in general equilibrium they know that changes in the value of their firm will reflect only the number of entrants.

Inserting the expression for profits in (8) into the HJB gives us another expression for this value in

$$(r_{j,t} + \delta)V_{j,t}(\bar{z}^E) = \frac{\bar{z}^E}{(\sigma-1)} \frac{w_{j,t}L_{j,t}}{N_{j,t}\bar{Z}_{j,t}} + \dot{V}_{j,t}(\bar{z}^E) + \sum_{i \in \{u,d\}} \phi_i(V_{j,t}(\Delta_i \bar{z}^E) - V_{j,t}(\bar{z}^E)),$$

Combining the two expressions for the mean value, along with another application of Lemma 1, gives us

$$r_{j,t} - \Phi = \frac{\bar{z}^E}{(\sigma-1)\tau} \frac{w_{j,t}L_{j,t}}{N_{j,t}\bar{Z}_{j,t}(N_{j,t}^E)^\zeta} + \zeta \frac{\dot{N}_{j,t}^E}{N_{j,t}^E},$$

where $\Phi = \sum_{i \in \{u,d\}} \phi_i(\Delta_i - 1) - \delta$ and $g_{j,t}^{ne} = \dot{N}_{j,t}^E / N_{j,t}^E$ is the growth in the flow of entrants. Rearranging, this discussion results in the following proposition:

Proposition 1. *Equilibrium entry in each location depends only on local state variables, and solves the differential equation*

(10)
$$(N_{j,t}^E)^\zeta = \frac{w_{j,t}L_{j,t}}{(\sigma-1)\tau N_{j,t}\bar{Z}_{j,t}/\bar{z}^E} \frac{1}{r_{j,t} - \zeta \dot{N}_{j,t}^E / N_{j,t}^E - \Phi},$$

Proposition 1 is the key result that allows me to solve the model. The result says that the equilibrium amount of entry depends positively on the local market size, captured in the amount of available labor $L_{j,t}$, and is falling in the number of existing firms. Moreover, population inflows will raise the entry rate, and cause the number of incumbent firms to rise, eating away at the profitability of new incumbent firms. The second term in (10) reflects the appreciation in value due to rising entry costs. This appreciation itself depends directly on the change in the amount of entry, and as such the entry rate solves a differential equation in time in every location which depends only on current state variables in that location. This substantially simplifies equilibrium dynamics.

The second part of the theory's feedback loop comes from the location choices of workers. Since workers are freely mobile at any instant, their welfare must equalize across all locations. Thus, given worker preferences from (3), wages and rental rates must satisfy

(11)
$$\frac{w_{j,t}}{p_{j,t}^{1-\alpha}} A_j = \frac{w_{k,t}}{p_{k,t}^{1-\alpha}} A_k$$

for all pairs of $(j,k) \in \{1,...,J\}$, where $p_{j,t}$ is the price of a unit of housing services in

location j. Given that land is in fixed supply, we can use market clearing for residential land to draw out the relationship between population and wage level. First, the price of a unit of housing services in location j will be given by

$$(12) \qquad p_{j,t} = ((1-\alpha)v_j \frac{w_{j,t}L_{j,t}}{h_{j,t}})^{v_j}.$$

The presence of a fixed supply of residential land for development implies that housing services become more costly the more people move to a location, and the more people earn in this location. Combining (12) and (11) yields an expression that relative populations must satisfy, given by

$$(13) \qquad \frac{L_{j,t}^{\bar{v}_j}}{L_{k,t}^{\bar{v}_k}} = \frac{w_{j,t}^{1-\bar{v}_j} A_{j,t} \bar{h}_j^{\bar{v}_j}}{w_{k,t}^{1-\bar{v}_k} A_{k,t} \bar{h}_k^{\bar{v}_k}},$$

where $\bar{v}_j \equiv (1-\alpha)v_j$ and $\bar{h}_j \equiv h_j/\bar{v}_j$. As in standard models of economic geography, relative populations are increasing in relative wages, amenities and residential land supply, recalling the static structure of Allen and Arkolakis (2014).[12] Dynamically, an increase in wages from entry of new firms will draw in new workers. The greater spending on local firms will partially offset declining profits, allowing further entry in the future.

The final piece to characterize an equilibrium concerns movements in $\bar{Z}_{j,t}$, or average-firm level efficiency. This efficiency evolves over time in response to firm level growth and movements in the entry rate $n_{jt}^E \equiv N_{j,t}^E / N_{j,t}$, according to

$$(14) \qquad \dot{\bar{Z}}_{j,t}/\bar{Z}_{j,t} = \sum_{i \in \{u,d\}} \phi_i(\Delta_i - 1) + n_{j,t}^E(\bar{z}^E/\bar{Z}_{j,t} - 1).$$

The first term in this expression reflects the growth in average firm level efficiency due to the stochastic firm-level shocks. The second is the contribution to average productivity due to entry. Entrants begin with an average productivity $\bar{z}^E$, and as such the speed at which they enter will also influence the evolution of average productivity.

This expression highlights the importance of including a description of post-entry dynamics in the model. From equation (6), wages are determined both by the number of firms and by their average efficiency. In the data, most entrants start small, and grow over time.

---

[12]It is worth noting that the (potential) presence of heterogenous land supply elasticities in the model have an important implication: relative rankings of places for workers are not invariant to the choice of units for $A_{j,t}$ and h_j. However, given the paper is largely concerned with responses to *relative changes* in fundamentals (which are unaffected by base units), this concern is immaterial for my purpose.

If size is a proxy for productivity, then a surge of entry will actually *lower* the average efficiency in an area, creating a partial drag on wages that only reverses as these new firms grow.

We now define an equilibrium under perfect foresight over location fundamentals.

Definition 1. *Given paths for fundamentals $\{\{A_{j,t}, h_{j,t}, B_{j,t}\}_{j=1}^{J}, L_t\}$, a perfect-foresight equilibrium is a path for worker populations $\{L_{j,t}\}_{j=1}^{J}$, wages $\{w_{j,t}\}_{j=1}^{J}$, interest rates $\{r_{j,t}\}_{j=1}^{J}$, the mass of firms in each location $\{N_{j,t}\}_{j=1}^{J}$ and average efficiency $\{\bar{Z}_{j,t}\}_{t=1}^{J}$ such that*

1. *Firm values are given by* (7) *and capitalist investment decisions maximize their intertemporal utility.*

2. *Worker welfare is equalized across space, such that* (13) *holds.*

3. *Wages across regions satisfy* (6).

4. *Free entry holds in* (10).

5. *Average efficiency $\bar{Z}_{j,t}$ evolves according to* (14).

6. *The labor market clears, such that*

$$\sum_{j=1}^{J} L_{j,t} = L_t.$$

Stationary Equilibrium. To study the equilibrium growth dynamics in this model, we begin by analyzing a stationary equilibrium where location fundamentals are held constant. The number of firms is constant in all locations, as are wages and populations. As such, the entry rate will be everywhere constant at δ (the exogenous exit rate of firms). Using the entry equation (10), this implies

(15)
$$\frac{w_{j,t} L_{j,t}}{N_{j,t}^{1+\zeta}} = \frac{w_{k,t} L_{k,t}}{N_{k,t}^{1+\zeta}}.$$

Then long-run populations in the stationary equilibrium can be solved from (13) and the total amount of labor in the economy. To characterize this stationary equilibrium in terms of objects which appear in the data, suppose that realizations of location fundamentals A_j, B_j and h_j are drawn from some stochastic distribution at time 0. Then we can show the following.

Proposition 2. *There is a unique stationary equilibrium in this economy if and only if $\min_{j}\{\frac{\bar{v}_j}{1-\bar{v}_j}\} > \frac{1}{(1+\zeta)(\sigma-1)-1}$. This equilibrium is locally stable. In this equilibrium:*

a) Wages increase with population according to

(16)
$$\frac{d}{d\log(L_j)}\mathbb{E}[log(w_j)|log(L_j)] = \frac{1}{(1+\zeta)(\sigma-1)-1} + \tilde{\sigma}\frac{d}{d\log(L_j)}\mathbb{E}[log(B_j)|log(L_j)],$$

where $\tilde{\sigma} \equiv \frac{(1+\zeta)(\sigma-1)}{(1+\zeta)(\sigma-1)-1}$.

b) Average firm size changes with population according to

(17) $$\frac{d}{d\log(L_j)}\mathbb{E}[log(L_j/N_j)|log(L_j)] = \frac{1}{1+\zeta}\left(\zeta - \frac{d}{d\log(L_j)}\mathbb{E}[log(w_j)|log(L_j)]\right).$$

where these conditional expectations are taken before time 0. As is standard in many spatial models, existence and uniqueness of a stationary equilibrium in this model depends on the balance of dispersion forces (represented through house price elasticities $\bar{v}_j$, and entry cost elasticity ζ) and attraction forces (embedded in the increasing returns to scale due to firm creation, embedded in σ). Local stability is demonstrated via linearizing the dynamics of the model around this steady state, and showing convergence.

The increasing returns from variety embedded in the wage equation in (6) leads to an equilibrium urban wage premium, with larger places paying higher wages. This can be seen in the first term in equation (16), where the scale elasticity $(\sigma-1)^{-1}$ shows up. This term is moderated by ζ, which determines the elasticity of the entry cost to the scale of entry, since in a stationary equilibrium places with more firms will see higher entry to replace those that exogenously die.

Without further structure on the location fundamentals, less can be said about the second term in (16), which captures how local TFP varies with population size in the stationary equilibrium. It will in general depend on the correlation between B_j, housing supplies h_j and local amenities A_j (and could in principle even be negative). To gain some intuition, we can derive a closed form for this expression under a set of simple assumptions, given in the following Lemma.

Lemma 2. *Suppose that* $log(A_j)$ *and* $log(B_j)$ *are i.i.d across space and drawn from a joint multivariate normal* $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ *at time 0. Moreover suppose that the housing supply elasticities* v_j *and shifters* h_j *are constant across space. Then*

$$\frac{d}{d\log(L_j)}\mathbb{E}[log(B_j)|log(L_j)] = \frac{(\frac{1}{\bar{v}-1})\Sigma_{BA} + \tilde{\sigma}^2\Sigma_B^2}{(\frac{1}{\bar{v}-1})^2\Sigma_A^2 + \frac{\tilde{\sigma}}{\bar{v}-1}\Sigma_{BA} + \tilde{\sigma}^2\Sigma_B^2}.$$

From this we can see that if highly productive places are also desirable places to live

22

Figure 7: Local Dynamics of Firm Creation
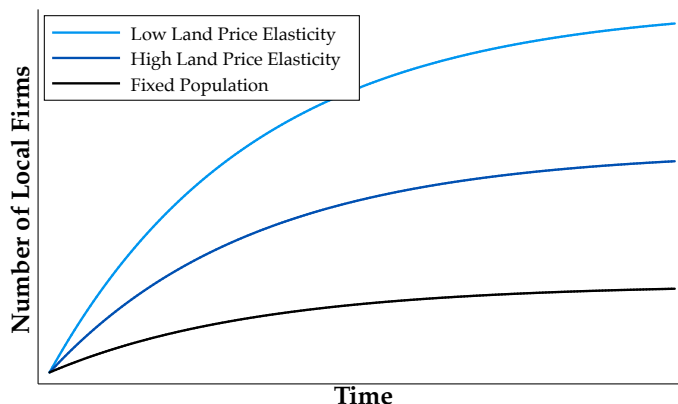
Note: This figure plots representative paths for equation (18) for a low land price elasticity, a high land price elasticity and a fixed population, beginning from the same initial number of firms in each location.

(such that $\Sigma_{BA} > 0$), then we will tend to observe higher wages in populous areas in the long-run. If the variance of amenities across space is large relative to local TFP (such that Σ_A^2 is large relative to Σ_B^2), then the relationship between wages and population will be weaker, and determined mostly by endogenous firm creation. Though stylized, this expression shows the difficulty in inferring the strength of endogenous productivity forces from cross-sectional data.

Note also that in this model, average firm size will tend to be rising in population size due to the congestion in firm creation from scarce land if this congestion is sufficiently large. This is consistent with the facts presented in Figure 6. At birth establishments are larger in more populous areas, but grow no faster. I use the expression in point b) of Proposition 2 directly in the estimation, exploiting the fact that average firm sizes across space give us information about the costs involved in setting up a new firm across space.

Non-Stationary Dynamics. Even in this simple setting, much of the dynamics must be analyzed numerically; the non-linear ODE in (10) cannot be solved in closed form. However, considering the linearized dynamics around the stationary equilibrium helps gain intuition for the central growth channel of the model. In particular, it reveals the crucial role of labor mobility in propagating the firm creation process through local demand. In this process, the land price elasticity will play a key role. Places where it is easy to build new housing will see greater amplification and propagation of changes to location fundamentals.

Consider the simplest case, with no post entry dynamics of firms (so $\Phi = -\delta$), and linear

utility for the capitalists (such that $\gamma = 0$). I show in the Appendix that beginning from all $N_{j,0}$ sufficiently close to their stationary values of $\bar{N}_j$, the mass of firms in each location will satisfy

$$N_{j,t} = (N_{j,0} - \bar{N}_j)e^{x_j t} + \bar{N}_j,$$

where

(18)
$$x_j = 0.5(\rho - \sqrt{\rho^2 + 4\delta \frac{(\rho + \delta)}{\zeta}(1 + \zeta - \frac{1}{\bar{v}_j(\sigma - 1)})}),$$

and it can be readily verified that $x_j < 0$ given $1 + \zeta - \frac{1}{\bar{v}_j(\sigma-1)} > 0$. To illustrate the main mechanism of feedback from labor mobility, let us consider how the mass of firms evolves in two identical locations j and k, save that one has a lower land price elasticity, such that $\bar{v}_k < \bar{v}_j$. In such a case, it can be verified that $\bar{N}_k > \bar{N}_j$, so that the long run mass of firms will be higher in the less price elastic location. Beginning from the same point $N_{j,0} = N_{k,0} < \bar{N}_j$, their respective time derivatives satisfy

$$\frac{\dot{N}_{k,t}}{\dot{N}_{j,t}} = \frac{x_k(N_{j,0} - \bar{N}_k)}{x_j(N_{j,0} - \bar{N}_j)}e^{(x_k - x_j)t},$$

and since $x_k - x_j > 0$, this ratio increases without bound. Moreover, if $N_{j,0}$ is sufficiently close to $\bar{N}_j$, this ratio is above one at time 0. As such, while both locations will see positive growth in the number of firms that converges to zero, the rate of increase is always faster in the less price elastic zone. What is happening is that labor is moving into location k at a faster rate than location j at all times, since house prices rise less for a given population inflow. This spurs local demand more in location k than j, and leads to further growth in both population and the mass of firms. I illustrate these dynamics in Figure 7. The limit case of no population dynamics can be considered by taking $\bar{v}_j \to 1$ in (18).

4. MODEL ESTIMATION

The model illustrates a mechanism whereby continued firm creation can generate persistent growth in employment and wages at the local level through variety gains. I now turn to quantifying this mechanism. The aim is simple. Returning to Figure 1, how much are startups generating growth themselves, and how much does their presence merely reflect other sources of local growth?

In answering these questions, I face a fundamental identification challenge. Firms are

forward looking, and expected changes to future fundamentals will affect firm creation today.

The strategy I employ isolates a series of demand shifters for local firm creation which are assumed to be uncorrelated with changes in local TFP. In particular, I leverage variation from the aggregate decline in manufacturing, which induces changes in spending at the local level. These changes depend on the specialization of a local area in manufacturing in 1975. I then use a baseline version of the model to generate a counterfactual equilibrium, where the only local changes are driven by optimal responses to an economy-wide manufacturing decline. This idea builds on a more general model instrumental variables approach developed in Allen et al. (2019) and Adao et al. (2019), who use the general equilibrium structure of trade models to generate instruments for the estimation of key parameters such as trade and labor supply elasticities.

Doing so generates spending shifters at the local level, and hence predictions for firm entry and employment over time. I use these predictions as instruments for the changes actually observed, recovering estimates of local scale elasticities, land supply elasticities and substitution patterns between manufacturing and service employment.

In the following section, I describe the setting and variation used for identification. In section 4.2 I modify the theory to accommodate multiple sectors. Section 4.3 describes the estimation.

4.1 Structural Change and Firm Creation

In recent decades, the U.S. has experienced a significant decline in manufacturing employment. 30.3% of the U.S. workforce were employed in manufacturing in 1975. By 2015, this number had fallen to 12%. Even during a period of significant national employment growth, manufacturing employment in levels actually fell, from 18 million in 1975 to 12 million in 2015. Figure **??** in Appendix **??** displays these patterns.

The causes of this decline explored in the literature embody a combination of trade-based displacement and automation. Trade explanations have emphasized heightened competition from manufacturing in the developing world since at least the 1990s, as well as a role for offshoring by U.S. companies themselves. Several influential studies attribute large losses in manufacturing employment to the entry of China in particular into the world trading system (Autor et al., 2013; Pierce and Schott, 2016). The evidence for the role of automation is so far less clear, but may have been significant in substituting labor for capital in some industries (Acemoglu and Restrepo, 2019; Hubmer, 2018). For a comprehensive review of these issues, see Fort et al. (2018).

Figure 8: Regional Growth and Manufacturing Specialization

(a) Employment Growth 1975-2015

(b) Manufacturing Specialization in 1975
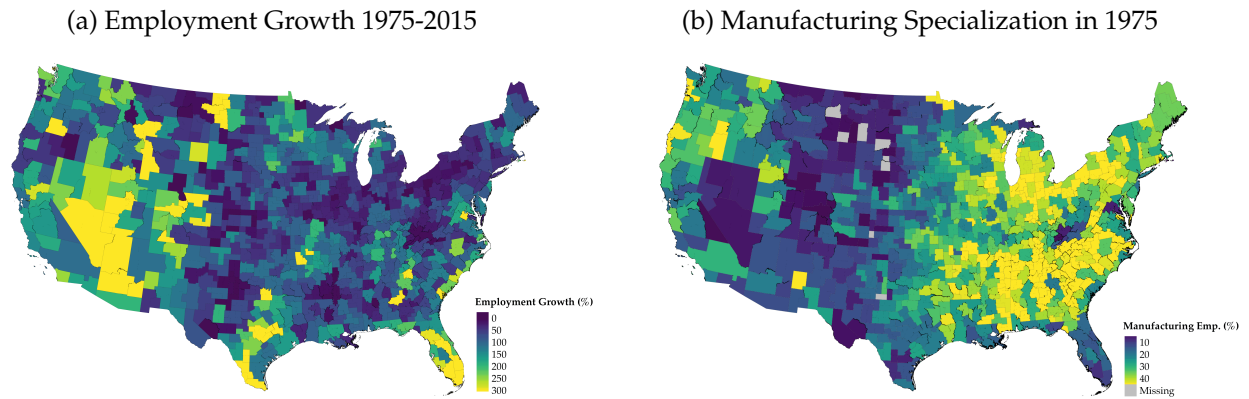
Note: This figure compares total employment growth at the commuting zone level from 1975 to 2015 (Panel (a)) with the fraction of employment in manufacturing within that commuting zone in 1975 (Panel (b)). Employment growth is top-coded at 300%, and manufacturing employment shares in 1975 are top-coded at 40%. Each polygon corresponds to a commuting zone. The source data is the public-use Quarterly Census of Employment and Wages (QCEW) produced by the Bureau of Labor Statistics, with aggregation of county data to the commuting zone level done by the author. Data for manufacturing specialization in South Dakota and North Dakota uses the 1980 QCEW as a proxy for 1975 due to total censoring in the 1975 files.

The effects of this aggregate structural change had an important local dimension. In Panel (a) of Figure 8 I plot local employment growth since 1975 across U.S. commuting zones. The fastest growing areas during this time have been Florida, Texas and and the sparsely populated Mountain West (light colors). The old industrial heartlands in the North-east and Midwest have grown the slowest (dark colors). In Panel (b) I overlay each commuting zone's manufacturing employment share in 1975; clearly, manufacturing-heavy commuting zones experienced slow growth over the past 40 years. This spatial structural change has left significant traces on American economic geography.

The time series dimension of this pattern is shown in Figure 9. From 1990, manufacturing-heavy areas began to diverge from the rest of the nation, with those areas doing little to no manufacturing employment in 1975 experiencing a surge in employment growth. This growth reflected large relative population shifts towards service-intensive areas, with little being due to differential growth in employment-population ratios.[13]

Employees spend at least a portion of their income locally. As such, these changes should imply slower growth in market size for new businesses in regions previously specialized in manufacturing. Lower growth should then cause lower rates of business formation, particularly in sectors that cater to local demand. Table 1 tests this prediction.

Using publicly available data from the QCEW, I regress changes in the number of estab-

_____

[13]See Appendix A.9 for details.

Figure 9: Employment Growth across Commuting Zones by Manufacturing Specialization
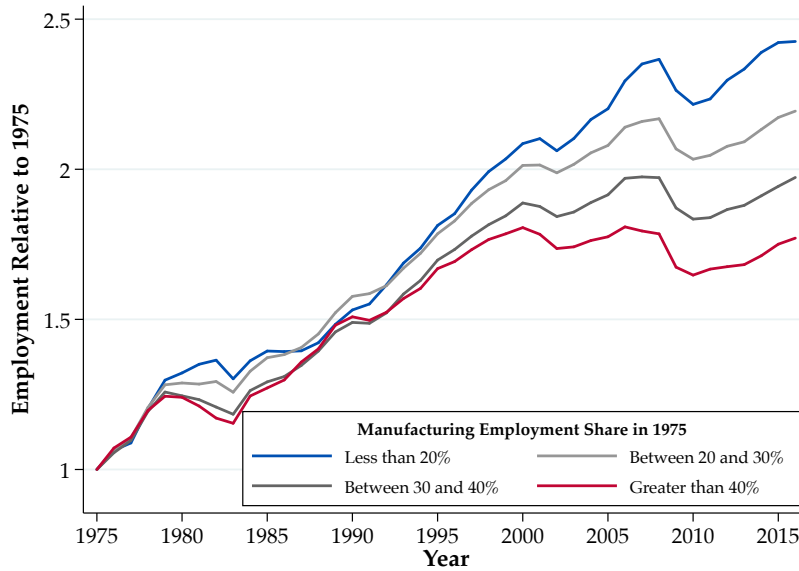
Note: This figure plots the average of annual employment relative to 1975 across four groups of U.S. commuting zones, where commuting zones are taken from the U.S. Census groupings of counties in 2000. The four groups are based on each commuting zones' manufacturing employment share in 1975. They also roughly correspond to quartiles of the manufacturing employment share distribution in 1975. Data comes from the Quarterly Census of Employment and Wages produced by the Bureau of Labor Statistics, and manufacturing employment uses the SIC code implementation for 1975.

Table 1: Firm Creation and Manufacturing Shares

| | | | Dependent variable: Log Change in Number of Establishments, 1990-2015 | | | | | |
| | | | Services | | | | | |
	All	Manufacturing	All	High-Skill	Local	Ed & Med	Construction	Trade
Manufacturing Share	−0.378***	−0.389***	−0.330***	−0.175	−0.319***	−0.718***	−0.668***	0.198
of Employment, 1975	(0.084)	(0.079)	(0.090)	(0.115)	(0.084)	(0.209)	(0.107)	(0.122)
Constant	1.359***	0.380	1.470***	1.906***	0.811**	0.656	0.622	1.058**
	(0.350)	(0.334)	(0.374)	(0.481)	(0.350)	(0.817)	(0.449)	(0.511)
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Size Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	510	496	510	504	510	377	488	501
Adjusted R^2	0.103	0.231	0.106	0.076	0.156	0.093	0.217	0.104

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

Note: This table shows regressions of the total log change in the number of firms by industry from 1990-2015 on the share of the local workforce employed in manufacturing in 1975. Geographic controls include the latitude of the centroid of the commuting zone and mean January temperature. Size controls are the commuting zone population in 1975 and the land area of the commuting zone in square kilometers. Demographic controls are the share of the population with less than a high-school degree, the average income of the commuting zone in 1975, the fraction of the population who was black in 1975, and the average age of the population in 1975. The changing number of observations by industry reflects censoring of the QCEW for some regions with low population.

lishments at the commuting zone level in the twenty-five years following 1990 on the share of employment in manufacturing in 1975. I report the results for four major sectors: Manufacturing, Services, Construction and Trade (which I classify as wholesale and transportation establishments). A consistent pattern emerges, with manufacturing-heavy areas seeing dampened business creation rates in subsequent decades across sectors.

I further separate Service industries into three categories: High-Skill (including finance, technology, consulting and management), Local (restaurants, retail and non-tradable services such as automotive repair), and Education and Medicine. The first category might be thought of as tradable services, and be less likely to respond to local demand shortfalls than Local Services and Education and Medicine, which is indeed what I find.

These results hold after including a rich set of controls for characteristics of each commuting zone in 1975, including size, geographic, and demographic characteristics.[14] As such, they support the idea that structural change and the shift of employment out of manufacturing had a direct effect on startup rates, depressing them in areas which were hit hardest by manufacturing's decline.

It is also interesting to note that while the number of manufacturing establishments contracted in most commuting zones during this period, they actually rose on average in commuting zones with less than 20% manufacturing employment in 1975. One interpretation is that the even in a period of national contraction, the large population and employment shifts presented in Figures 8 and 9 generated sufficiently increased demand for manufacturing products in these areas that they incentivized the creation of new firms to meet this demand.

4.2 Taking the Model to the Data

To take the model to the data, I now introduce two sectors: Manufacturing and Services. I suppose that final output in each location is composed of two final sectoral goods, given by a CES aggregator

$$
(19) \qquad Y_{j,t} = \left[(Y_{j,t}^M)^{\frac{\epsilon-1}{\epsilon}} + (Y_{j,t}^S)^{\frac{\epsilon-1}{\epsilon}} \right]^{\frac{\epsilon}{\epsilon-1}}
$$

where $Y_{j,t}^M$ and $Y_{j,t}^S$ are the final output of the manufacturing good and the service good. Both these final outputs are produced competitively by aggregating intermediate inputs within a location according to (5), but with sectoral specific elasticities σ_M and σ_S. I modify

---

[14]See the note in Table 1 for details.

the production function to accommodate sectoral differences, and assume that production of each intermediate variety is done according to

$$q_{j,t}^M(i) = B_{j,t} M_{j,t} \tilde{z}_t(i) l(i) \quad q_{j,t}^S(k) = B_{j,t} \tilde{z}_t(k) l(k)$$

for firms in the service and manufacturing sectors respectively. $B_{j,t}$ continues to define a sector-neutral location-specific TFP, while $M_{j,t}$ is a manufacturing-specific productivity shifter. This manufacturing shifter is the key cause of structural change in the two-sector model. It consists of two components: a location specific $m_{j,t}$, and an aggregate Π_t common to all locations, such that

$$M_{j,t} = m_{j,t} \Pi_t.$$

The comparative advantage of a location in manufacturing is captured by $m_{j,t}$, and could reflect a number of features that make manufacturing firms particularly productive in a location, including favorable transportation infrastructure, cheap power supply and historical local knowledge stocks. Without loss, I take these to be mean zero in logs across locations.

Π_t is an aggregate shifter which affects the output of all manufacturing firms within the economy. The effect it has on a particular location depends on local specialization in manufacturing. Such specialization can come from two sources; comparative advantage in productivity through $m_{j,t}$, and the number of firms $N_{j,t}^M$ producing manufacturing goods. To see this, note that at any time the share of the local labor force in services will be determined by

$$\frac{L_{j,t}^S}{L_{j,t}} = \frac{\left(\bar{Z}_{j,t}^S (N_{j,t}^S)^{\frac{1}{\sigma_S - 1}} \right)^{\epsilon - 1}}{\left(\bar{Z}_{j,t}^S (N_{j,t}^S)^{\frac{1}{\sigma_S - 1}} \right)^{\epsilon - 1} + \left(m_{j,t} \Pi_t \bar{Z}_{j,t}^M (N_{j,t}^M)^{\frac{1}{\sigma_M - 1}} \right)^{\epsilon - 1}}.$$

It is apparent that all else equal, falls in Π_t will raise service employment locally, as workers reallocate away from manufacturing locally. For a given change in Π_t, this effect will be more pronounced the higher manufacturing comparative advantage $m_{j,t}$, and the greater the number of manufacturing firms relative to service firms currently operating.

However, two other forces come in to play when aggregate manufacturing activity suffers. First, wages fall immediately, and fall more in places specialized in manufacturing. This causes some workers to leave the area until spatial equilibrium is restored. Second, as workers leave, investment in manufacturing-heavy areas becomes less attractive due to a lower overall level of spending in this area. This reduces the number of firms over time, and due to the economies of scale embedded in the production function in (19), wages

29

full even further, causing more workers to leave. All together, this can serve to *lower* total service employment in the long run.

4.3 Estimation Implementation

Estimation of the two sector model proceeds in two stages. First, I recover key firm-level parameters from direct estimation of the firm growth process on the LBD. Second, I use the structure of the model to generate moment conditions, based on counterfactual equilibrium changes in employment and firm creation, to estimate the fundamental local production parameters.

Firm-level Parameters. I first estimate the parameters describing the firm life cycle directly from the LBD. In a stationary equilibrium, growth in firm-level productivity can be directly inferred from growth in employment, conditional on survival. I estimate the firm level growth parameters ϕ_i and Δ_i for $i \in \{u, d\}$ using the first four moments of the employment growth distribution in the entire LBD. The notion of a firm in the model corresponds most closely to an establishment with a fixed address in the data, and as such I use the LBD to construct the distribution of establishment growth rates from 1980-2015, aggregating across all industries.

To estimate the parameters of the model, I simulate a panel of firms, and sample the data at a yearly time frequency to calculate the distribution of annual firm growth rates conditional on survival. The four growth parameters are chosen to match the first four moments of the empirical distribution of firm growth rates, and are reported in Table 6 in Appendix A.11. I estimate the death rate of establishments δ from the average annual exit rate from 1980 to 2015.

Structural Production Parameters. Next I estimate the local scale elasticities in the number of firms, σ_M and σ_S, and the elasticity of substitution between manufacturing and service production at the local level, ϵ. Market clearing will imply relative sectoral employments given by

$$
(20) \qquad \frac{L_{j,t}^S}{L_{j,t}^M} = \left(\frac{(N_{j,t}^S)^{\frac{1}{\sigma_S - 1}} \bar{Z}_{j,t}^S}{m_{j,t} \Pi_t (N_{j,t}^M)^{\frac{1}{\sigma_M - 1}} \bar{Z}_{j,t}^M} \right)^{\epsilon - 1}
$$

Local specialization depends on an external comparative advantage (captured through $m_{j,t}$), and the relative numbers of firms in both locations, which are equilibrium outcomes. We can use this ratio to derive an expression that the real wage must satisfy in a location,

given by

$$(21) \qquad w_{j,t} = (N_{j,t}^S)^{\frac{1}{\sigma_S - 1}} \left(1 + \frac{L_{j,t}^M}{L_{j,t}^S}\right)^{\frac{1}{\epsilon - 1}} \bar{Z}_{j,t}^S B_{j,t}$$

An analogous equation holds for the local wage in terms of the number of manufacturing firms and local manufacturing comparative advantage (recall that services productivity is normalized to 1 in every location).

The fundamental estimation problem is immediately clear from this equation. Wages will tend to be high where local TFP $B_{j,t}$ is high, and if there are scale economies, where the number of local firms $N_{j,t}$ is high. However, local TFP is unobserved, and the number of firms in a location will respond to TFP. Examining (21), higher productivity will draw more workers to an area, raising local profits and incentivizing entry. This is a problem that is not specific to this paper, and arises in a wide variety of contexts that attempt to estimate notions of agglomeration (see Ahlfeldt and Pietrostefani (2019) for a review).

Note also that estimating equation (21) in changes is not sufficient to address this issue. Write the estimating equation in log differences as

$$(22) \qquad \hat{w}_{j,t} = \frac{1}{\sigma_S - 1} \hat{N}_{j,t}^S + \frac{1}{\epsilon - 1}(1 + \widehat{L_{j,t}^M / L_{j,t}^S}) + \hat{\bar{Z}}_{j,t}^S + \hat{B}_{j,t}$$

where $\hat{x}_{j,t} \equiv log(\frac{x_{j,t}}{x_{j,t-1}})$ for any variable x and two points in time.[15] Changes in local TFP $B_{j,t}$ will also be correlated with changes in the local numbers of firms in both sectors, $N_{j,t}^S$ and $N_{j,t}^M$. Estimating this equation by OLS will give us inconsistent estimates. What is needed are instruments for the local number of firms in each sector and changes in sectoral employment which are themselves uncorrelated with disturbances to local productivity.

The structure of the model suggests a natural way to do this. [16] A decline in aggregate manufacturing productivity Π_t will imply different investments in the local number of firms, depending on initial manufacturing specialization. Using a guess for the fundamental parameters of the model, I calibrate the model to a steady state in 1975. In doing so I back out fundamental comparative advantages $m_{j,t}$ to exactly match each commuting zone's share of employment in manufacturing in 1975, given observed numbers of firms in each sector. I then solve for a perfect-foresight path of Π_t which exactly matches

---

[15]For the estimation I take these to be five-year changes in order to minimize the contributions of measurement error and Census sampling frame adjustments among small commuting zones.

[16]I defer to the appendix discussion of changes in local firm efficiency, $\hat{Z}_{j,t}^S$, which can be handled via a separate procedure.

the series for total U.S. manufacturing employment in the data. The model generates a time-series for $(\tilde{N}_{j,t}^S, \tilde{N}_{j,t}^M)$ as well as labor allocations $(\tilde{L}_{j,t}^S, \tilde{L}_{j,t}^M)$ in the absence of changes to manufacturing comparative advantage $m_{j,t}$ and local TFP $B_{j,t}$.

I use these counterfactual investment and employment patterns as instruments for the true changes, leaving us with the moment conditions

$$\mathbb{E}[\hat{B}_{j,t}\hat{Q}_{j,t}] = 0,$$

$$Instruments: \quad Q_{j,t} = \left[\begin{array}{c} \tilde{N}_{j,t}^S \\ \tilde{N}_{j,t}^M \end{array} \right].$$

Since the instruments are model-generated, with no additional local data past 1975, they are in one sense simply time-varying functions of the manufacturing shares of each commuting zone in 1975. Thus these moment conditions boil down to specifying that local changes in sector-neutral TFP (i.e. changes *not* involving structural change in manufacturing) are uncorrelated with initial manufacturing shares.

Before proceeding further, it is worth asking how well these instruments predict the actual observed changes in economic activity. The results from the first-stage, given in Appendix A.10, are quite strong, and in addition, the instruments do not systematically over- or under-predict their data counterparts. This might be taken as a partial validation of the model, as no information on entry patterns or firm investment is used to generate the counterfactual series for employment and establishment changes.

A second set of moment conditions pins down the scale parameters for manufacturing σ_M. This comes from the analog of (22) with manufacturing firms on the RHS, leaving us with the moment condition

$$\mathbb{E}[\hat{m}_{j,t}\hat{Q}_{j,t}] = 0$$

This condition also has a simple interpretation. It assumes that changes in *relative* comparative advantage in manufacturing are uncorrelated with manufacturing shares in 1975. It is worth specifying what this means; it does not say that manufacturing-heavy areas experienced no greater shifts in relative employment to services. Rather, it says that all of this effect is mediated through initial exposure to the investment decisions of firms. Places that were better at manufacturing in 1975 did not become systemically better or worse in the sector in later years.

Third is the strategy for ϵ, the local elasticity of substitution between manufacturing and services. Recall that for a given ϵ, to generate the instruments $(\tilde{N}_{j,t}^S, \tilde{N}_{j,t}^M)$ I solve for the

path for aggregate manufacturing productivity Π_t that exactly matches the U.S. manufacturing employment share. Changing ϵ influences the *regional* consequences of structural change. In particular, a high value of ϵ makes it easier for employment to move from manufacturing into services *locally*, without large effects on the local wage. In contrast, when ϵ is low, a fall in Π_t will depress wages in areas that are more specialized in manufacturing, and cause more workers to leave the area and move to service-intensive locations.

I thus choose ϵ to match the model-implied effects of structural change with what is observed in the data. In particular, I choose ϵ such that, given all other estimated parameters and a path for Π_t that matches the aggregate manufacturing employment share, the average difference in relative employment growth between manufacturing-intensives areas in 1975 (red commuting zones in Figure 9) and service-intensive areas (the blue commuting zones) by 2015 is exactly 36% (corresponding to the gap between the endpoints of the lines).

Lastly, we need to estimate a value for ζ, the parameter which determines how costly it is for the mass of entrants to adjust when local fundamentals change. Proposition 2 highlights that in the stationary equilibrium, this parameter also directly governs how average firm size increases with population size. As such, I use equation (17) directly, computing the average firm size elasticity and average wage elasticity with respect to population for the years 1980-2010 in the Quarterly Census of Employment and Wages, and use these to estimate ζ.

The estimated parameters appear in Table 7 in Appendix A.11. The estimated scale elasticity for wages in the local number of manufacturing firms is significantly higher than that for services (captured in the relative differences between σ_M vs σ_s). This perhaps accords with intuition: it is commonly assumed by policymakers that new manufacturing activity generates greater local multipliers than the equivalent investment in service firms.

Housing and Labor Supply Elasticities. The discussion in Section 3 highlights the key role that land supply plays in propagating the feedback from firm creation and labor supply. Places with more restrictive physical or regulatory constraints will receive smaller population inflows, and reduced growth in wages as a result. To complete the estimation and obtain measures of these elasticities at the local level, I follow a similar model-instrumentation strategy to above. From equation (12), the rental rate of housing satisfies (in log changes)

$$(23) \qquad \hat{p}_{j,t} = v_j(\widehat{w_{j,t}L_{j,t}} - \hat{h}_j)$$

such that log rental changes are linear in log changes in the total wage bill in the area.

I lack detailed data on rental rates at the commuting zone level. Instead, I impute them by assuming that rental rates are proportional to house prices. I use house price indexes from the Federal Housing Finance Agency at the county level, from 1980 to 2017, and deflate these by the CPI in each year. I then construct measures of long percentage changes in house prices (over the 5 year horizon) to abstract from short-term volatility in these series. I aggregate these changes to the commuting zone level by taking population-weighted averages of these changes over each county within a given commuting zone.

The omitted variable problem in estimating (23) is clear. Unobserved changes in shifters of the housing supply function $\hat{h}_j$ (such as changes to local zoning laws) will tend to induce movements in $L_{j,t}$. Moreover, endogenous changes in zoning (which would be captured in $\hat{h}_j$) in response to increases in local income are well-documented (Saiz, 2010). This would suggest that OLS estimates of v_j in equation (23) will be biased, though the direction is unclear. To obtain consistent estimates, I again use the model to generate instruments for total payroll, denoted $\widetilde{w_{j,t} L_{j,t}}$. Identification requires now that changes in local housing supply shifters in any period are uncorrelated with manufacturing shares of employment in 1975, which again is the only local data used as inputs to solve the counterfactual model.

Due to the limited length of data for an individual commuting zone, I pool commuting zones and estimate a single elasticity. The results are presented in Table 7. The IV estimate is substantially higher than the OLS estimate.

5. Startups and Local Growth

I now use the estimated model to perform four exercises, which together quantify the contribution of firm creation to growth. First, I show the impulse responses of changes in local employment and wages to reductions in the cost of entry, incorporating the endogenous responses of employment and population movements. I decompose the total changes into a component due to the initial shock and a dynamic component that reflects feedback from labor mobility.

Second, I decompose changes in wages at both the local and aggregate level into contributions from firm creation and from incumbent efficiency improvements. Third, I study the outcomes of the model under persistent, anticipated changes to local fundamentals, and quantify the role of firm creation in amplifying local shocks.

Finally, I use the model to consider how local investment decisions amplified aggregate structural change in recent decades, and in particular the decline of the US Rustbelt. In

Figure 10: Impulse Responses of Employment and Wages to a Fall in Entry Costs τ
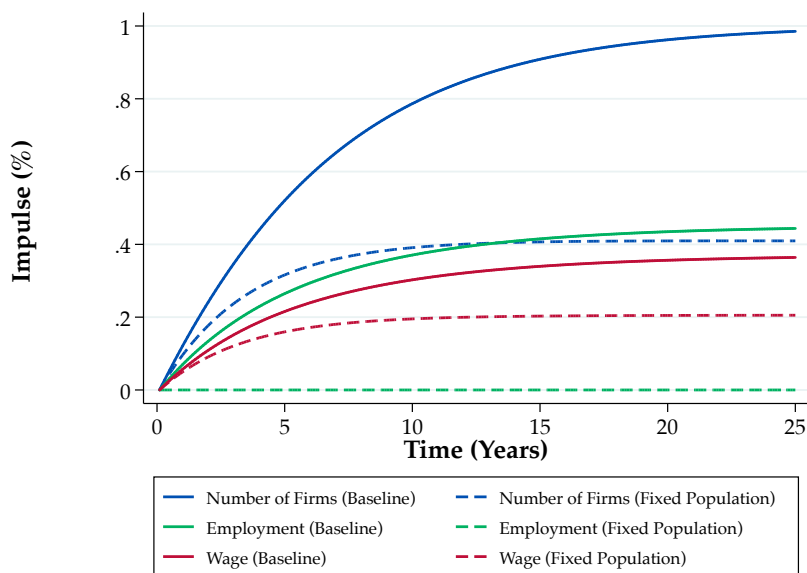
Note: This figure plots impulse response for employment, wages and the number of firms to a one-time, permanent shock to entry costs τ in that location, holding all other regions the same. The shock is calibrated to deliver a 1% increase in the long run number of firms. Dotted lines represent the counterfactual outcome when labor is not allowed to move.

particular, I study the contribution of firm creation dynamics to amplifying the differences in employment and wage growth between areas specialized in manufacturing versus those specialized in services in 1975, incorporating full general equilibrium dynamics.

5.1 Dynamic Contributions to Employment and Wage Growth

I first consider the dynamic response of local economic activity to a subsidy that lowers the cost of entry τ. Beginning from a stationary equilibrium, an unanticipated subsidy to the entry cost occurs in a single location, the size of which is chosen to increase the long-run number of firms by 1%. In Figure 10 I plot the dynamic response of local economic activity.

In the long-run, wages increase by 0.3% due to the estimated returns to scale in the production structure. Employment increases by 0.4% as workers move to the area to take advantage of the higher wages. Two central messages emerge from this exercise. First, the dynamics are slow; after 5 years, the baseline effects of the shocks are still barely halfway to their long run values. This reflects the difficulty of adjusting the mass of firms quickly in response to local shocks.

Second, there is substantial endogenous feedback. The dotted lines show the response

Figure 11: Impulse Responses of a Shock to Local Productivity

(a) Permanent Shock
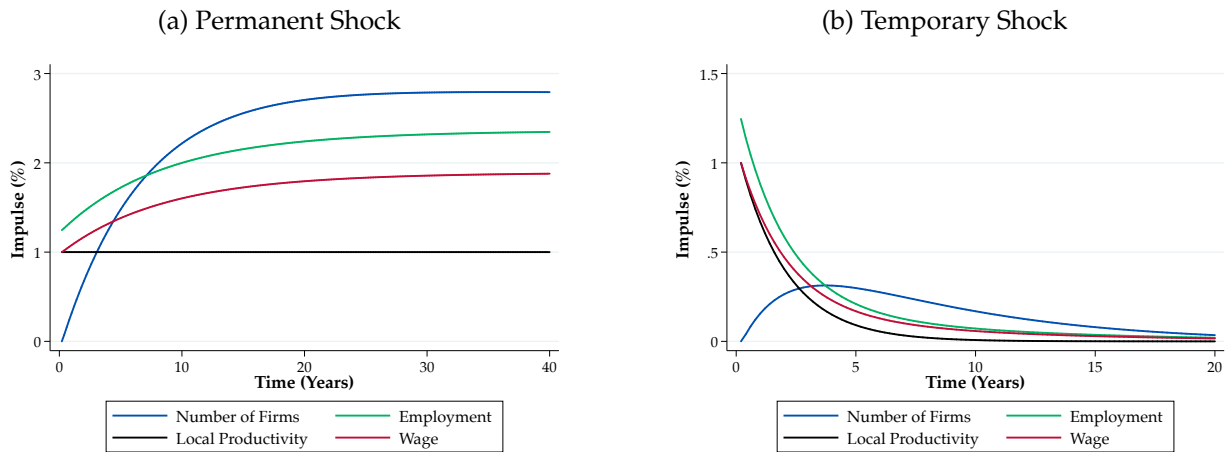
(b) Temporary Shock

Note: This figure plots impulse responses to a permanent (Panel (a)) and temporary (Panel(b)) shock to local productivity in the full estimated model.

of the mass of firms and wages if no labor is allowed to relocate to the affected location. The response here is 60% smaller, suggesting that new workers brought in by the initially higher wages themselves provide an important impetus to local demand.

To see this another way, in Figure 11 I show the dynamics out of steady state after an impulse to local TFP. Panel (a) considers a one time, unforeseen permanent increase in $B_{j,t}$ of 1% in one location only, which then persists for all time. The shock immediately causes the local wage to rise 1%, and since labor is free to move at all times, an immediate rise in employment. The dynamics of firm creation then respond to this increase in local demand by slowly rising to an almost 3% increase in the baseline level. The long run increase in wages is almost double the initial impulse. Panel (b) shows the propagation of a temporary shock, and reveals that the firm creation dynamics act to prolong and propagate the effects of the initial impulse over time.

I now turn to the role of firm creation in driving aggregate wage growth. The estimated model allows us to perform a simple decomposition of per-capita wage growth into a contribution for $N_{j,t}$, and a joint contribution of TFP $B_{j,t}$ and incumbent efficiency $Z_{j,t}$. For ease of interpretation, I abstract from sectoral differences for this exercise, and use equation (6) and the implied joint σ from the impulse response in 10 (which gives $\sigma = 3.9$) .

I apply this equation to data on changes in local wages at the commuting zone level. On the right hand side of the equation, I use rolling five-year windows of the changes in the number of local establishments. Two results are of interest, presented in Figure 12. First, in most years, a substantial fraction of wage growth at the commuting zone level is

attributable to firm creation. Panel (a) shows the inter-quartile range of this fraction for the past 25-years, as well as the median, which averages 26% during this period.

In Panel (b), I compute employment-weighted changes in real wages for each 5-year window at the commuting zone level. Summing these across all commuting zones yields a figure for average real wage growth for the entire United States. I do the same aggregation for changes in the numbers of establishments. We can see that U.S. real wage growth also has a substantial component due to firm creation. This contribution declined substantially in the Great Recession, along with real wage growth, reflecting a significant shortfall in the number of new firms. It has recovered somewhat in the years since. However, the recovery in firm creation is driven by a few large metropolitan areas, which have seen renewed business dynamism since 2008. It is not apparent in the median commuting zone, which, as Panel (a) in Figure 12 shows, has continued to see persistently low entry rates. This graph also shows that the surge in real wage growth in the 1990s was not due to a surge in business creation. Other work (e.g. Fernald (2015)) has documented that this surge likely reflected one-time TFP improvements from the ICT revolution, which have not been repeated since.

These figures for wage growth are substantially higher than the fraction of productivity growth due to entrants calculated in Garcia-Macia et al. (2019), despite the fact that the calibrated σ in that paper is 4, virtually identical to the 3.9 here. There are two reasons for this. First, by looking at 5-year horizons, the drag that the small size of entrants at entry reverses quickly, due to their fast initial growth. As such, at a longer time horizon the contribution from variety gains is much more apparent. Second, there is no role for creative destruction at entry in this model, whereas in Garcia-Macia et al. (2019) new entrants may steal existing products.

5.2 Persistent, Anticipated Shocks and Propagation

I now consider the role of firm creation in propagating and amplifying fundamental location shocks. To do so, I go beyond data description to postulate a simple process for location productivity $B_{j,t}$, and study how the model behaves under this process.
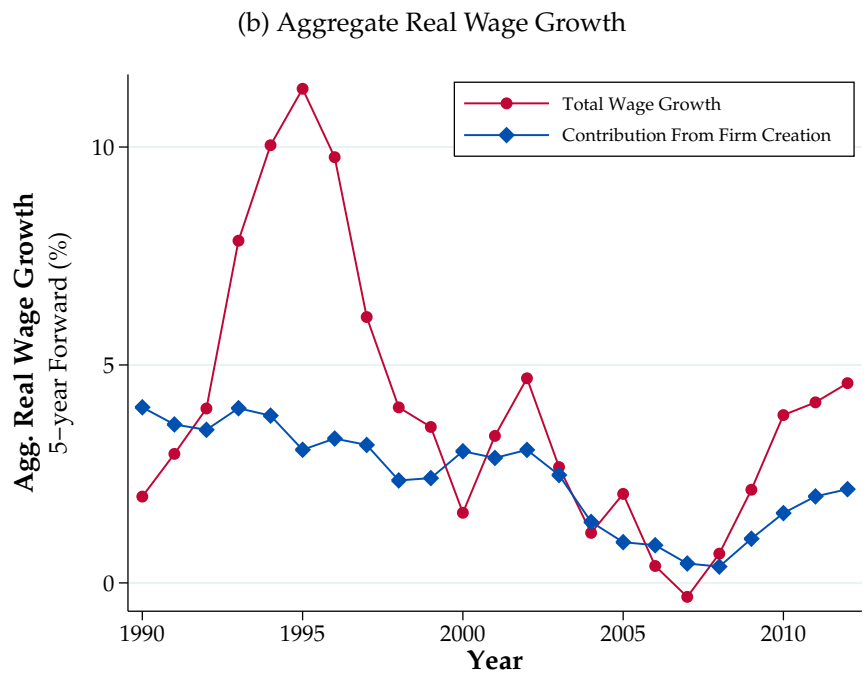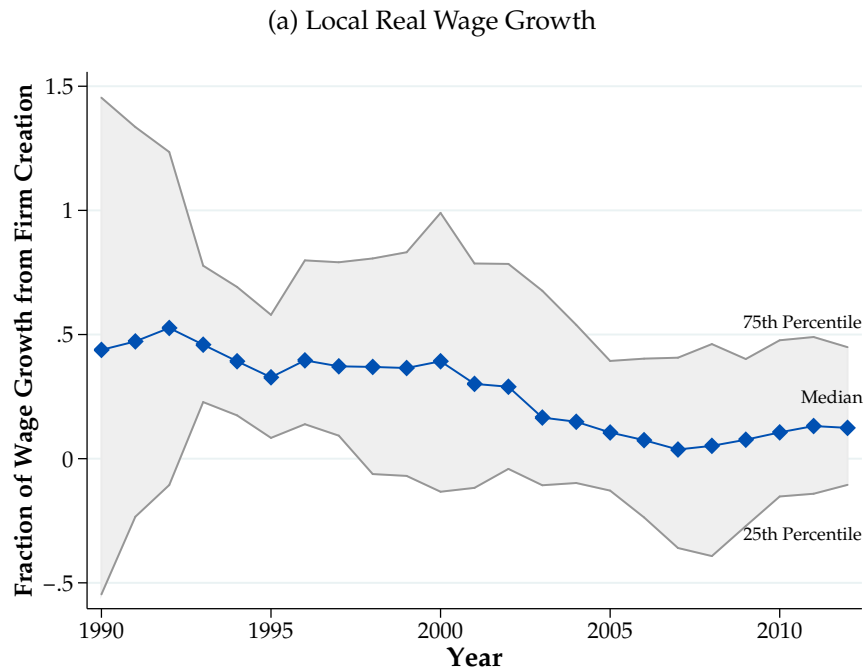
In particular, I suppose that the the growth rate of $B_{j,t}$, denoted $g_{j,t}^B$, follows an Ornstein-Uhlenbeck process, such that

$$dg_{j,t}^B = -\theta g_{j,t}^B dt + \varsigma dW_{j,t}$$

where $W_{j,t}$ is a standard Weiner process. This process is particularly convenient for cap-

Figure 12: Firm Creation and Real Wage Growth

(a) Local Real Wage Growth

(b) Aggregate Real Wage Growth

Note: Panel (a) plots the median fraction of commuting zone wage growth due to changes in the number of establishments in the estimated model, as well as the interquartile range. Wage growth is computed as the log change in the average real wage from year t to year $t + 5$. Real wages at the commuting zone level are computed by deflating nominal wages by the U.S. CPI. Panel (b) aggregates these changes for the whole U.S., weighting the observations for each commuting zone by employment in each year. Data is from the Quarterly Census of Employment and Wages.

Figure 13: Simulated Local Growth Rates for Full Model and Underlying Shocks
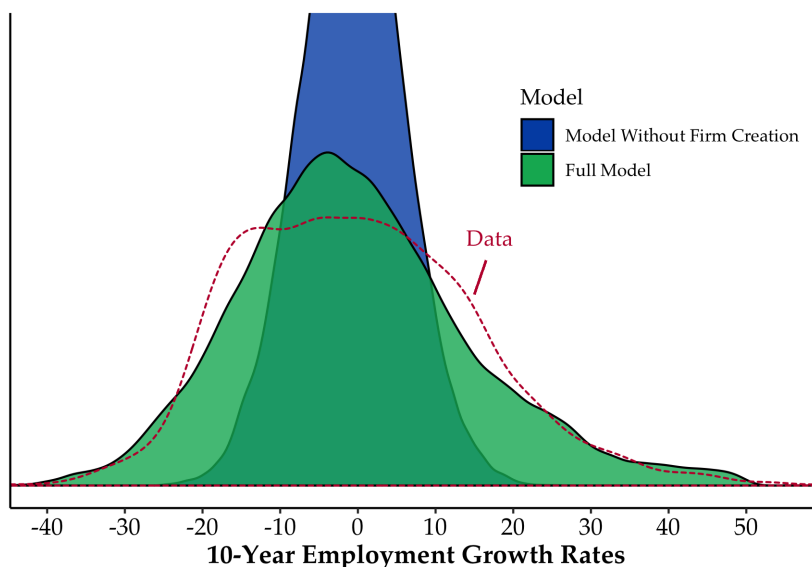
Note: This figure shows the distribution of simulated 10-year growth rates in employment under perfect foresight, and where the process for local productivity follows the process outlined in Section 5.2. 10-year employment changes are defined as the log change change in total employment at the commuting zone level between year t and year $t + 10$.

turing persistent differences in the growth of areas that is not driven by the model mechanism of the paper. In this formulation, θ indexes how quickly the growth rate of local productivity returns to its long-run (zero) mean, and controls the persistence of growth in location fundamentals. ς controls the volatility of the growth rate process. I assume that the realizations of these paths are drawn at time 0, and that firms continue to fully anticipate the paths for productivity and population in their location, making their entry decisions accordingly

I choose these parameters to exactly match two central moments from the data. Keeping all other model parameters as estimated in Section 4, I first target the variance in 10-year growth rates in employment at the commuting zone level, where I calculate this from the QCEW. Second, I target the autocorrelation of the growth rate in employment at a 10-year horizon. With calibrated values for θ and ς in hand, I then repeatedly simulate the model under this process for 100 year time horizons.

In Figure 13 I show the simulated distribution of 10-year employment growth rates.[17] I then assess the role of the fundamental shocks to local productivity in driving this dispersion by shutting off firm creation entirely, keeping the number of firms in each location fixed at their initial values. The distribution of growth rates just driven by local produc-

---

[17]I begin the simulation from a steady state calibrated to data in 1975, though the results are robust to this choice.

tivity is show in blue, and the distribution in the full model is shown in green.

The standard deviation in 10-year employment growth rates is almost three times as high in the full model as in the model that only experiences local productivity shocks and no firm dynamics. Places that see local employment booms find these booms amplified significantly by the process of firm creation, which adds to the demand for new workers and significantly propagates surges in employment.

5.3 The Startup Multiplier and the Decline of the US Rustbelt

I lastly consider the role that firm creation has played in facilitating the changing patterns of employment growth documented in Section 4. To do so, I take structural change in the model to be driven by a fall in aggregate manufacturing output Π_t which affects all manufacturing firms in the economy, in the same way as I did when generating the instruments for estimation. The most straightforward interpretation of this shock is a trade shock driven by developing country competition. Ordinarily, we might think of this as affecting the equilibrium relative prices of manufacturing goods. Without an explicit model of trade, such a notion can be mapped onto movements in Π_t, which directly affects the relative prices of goods and services within a location.

I again solve the full model for the path for Π_t which exactly matches the series for the aggregate share of U.S. employment in manufacturing, and recover implied commuting-zone level growth paths. I retain the assumption of perfect foresight on the part of individual firms, such that they anticipate the fall in manufacturing productivity through Π_t from 1975 onwards.[18] I also take the path for aggregate employment L_t as exogenous and fully anticipated on the part of the firms. Lastly, I assume no changes to location fundamentals ($B_{j,t}$, $A_{j,t}$ and m_{jt}) from their values in 1975, which are calibrated to rationalize observed data on wages, employment by sector and number of firms by sector within each commuting zone for that year.

Panel (a) in Figure 14 plots the implied paths for employment relative to 1975 in every commuting zone in the full estimated model. The colors correspond to the same groupings of commuting zones used in Figure. In Figure 31 in Appendix A.12 I show average employment growth within each of the four groupings. Recall that ϵ was estimated such that the gap between the least and most manufacturing-intensive group (blue and red) in 2015 matches the data.

In Panel (b) I plot the outcomes of the model when shutting down endogenous investment dynamics. I equalize the firm entry rates in every region to be the exogenous exit

---

[18]Firms still face idiosyncratic uncertainty with respect to their efficiency before before and after entry.

Figure 14: Contribution of Startup Deficit to Spatial Structural Change
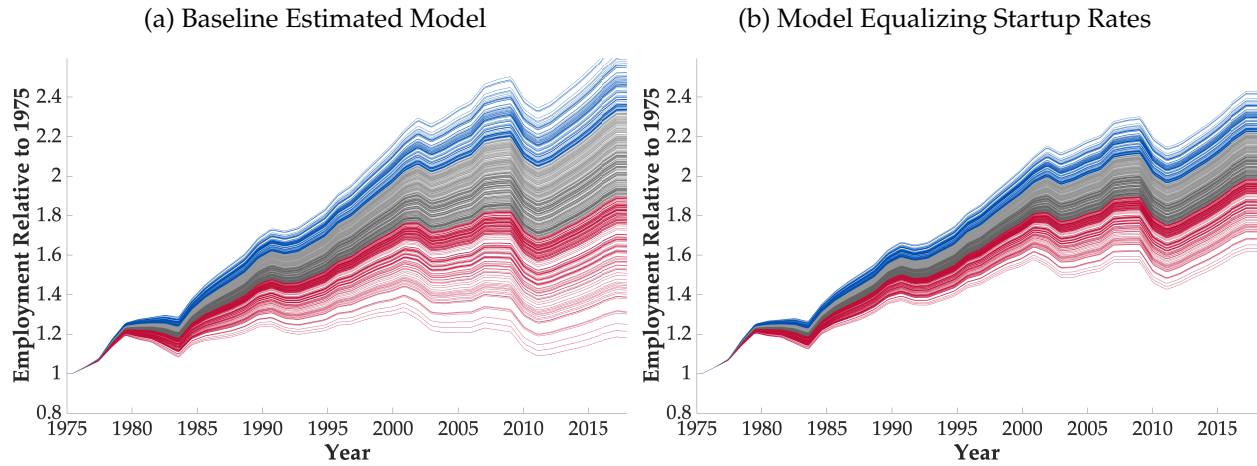
(a) Baseline Estimated Model   (b) Model Equalizing Startup Rates

Note: This figure plots the evolution of employment across commuting zones using the baseline estimated model under two scenarios. Panel (a) shows the model-implied growth in employment under the path for manufacturing productivity which matches the estimated decline in manufacturing employment. Panel (b) shows growth when fixing the number of firms in both manufacturing and services to their baseline values in 1975. The four color groups correspond to the four groups of 1975 manufacturing shares employed in Figure 9: blue (less than 20%), dark grey (20-30%), light grey (30-40%) and red (over 40%).

Figure 15: Model-Implied Startup Deficits Due to Structural Change

Startup Deficit (%)
0.6
0.8
1.0
1.2
1.4
Missing
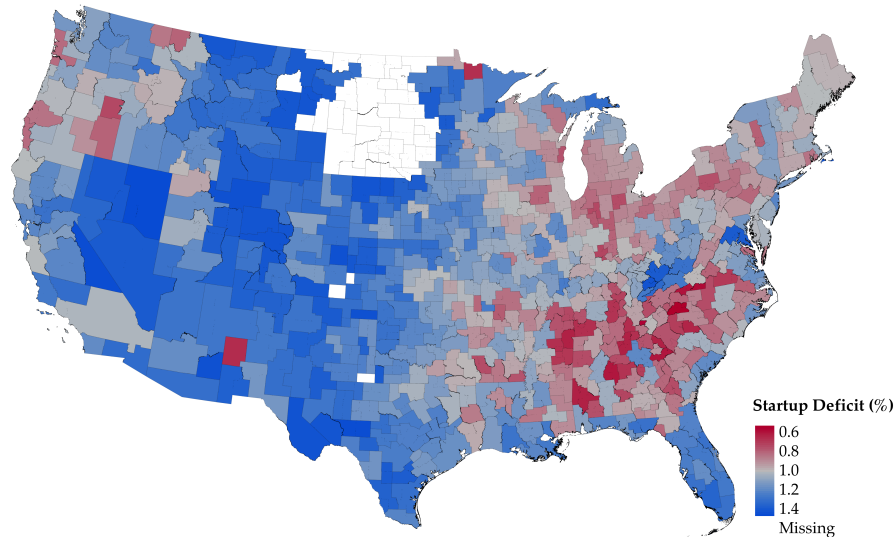
Note: This figure shows the model-implied startup deficit in each U.S. commuting zone when the productivity process for aggregate manufacturing productivity $\Pi_{j,t}$ is chosen to exactly match aggregate U.S. manufacturing employment. The deficit is calculated as the ratio of the model-implied mass of firms at the local level in 2015 against the counterfactual outcome in the same year where $\Pi_{j,t}$ is held constant from 1975-2015. Cells are missing where manufacturing employment in the 1975 QCEW is suppressed.

rate δ for the entire period of 1975-2017. As such, the entire effect of structural change is captured by the movement of workers *across* firms. Under this scenario, dispersion is greatly reduced; the difference in employment growth between the service heavy and manufacturing heavy commuting zones is on average only 37%. Put another way, had entry rates been equalized across regions during this period, the manufacturing-heavy commuting zones in the U.S. Northeast and Midwest would have almost 10% more employment relative to what we see today. While the decline of manufacturing would have still driven marked regional divergence, this divergence was amplified and propagated by endogenous firm creation decisions.

To see the regional implications directly, in Figure 15 I map estimates of the "startup deficit" implied by the model in each commuting zone. I define this deficit as the total number of local firms in the full model in 2015 relative to that in the scenario where entry rates are equalized across space between 1975 and 2015. Commuting zones in red are those for which the deficit is positive, reflecting an initial specialization in manufacturing, and consequently slower firm creation dynamics over the last forty years. Two regions stand out as "missing" a significant number of firms. The first are the classic US Rustbelt states, stretching from New York to Wisconsin along the shores of the Great Lakes. A second zone is apparent moving through Tennessee, Kentucky and parts of Virginia and North Carolina. While these regions have also experienced a range of other changes that are not captured in the model, this analysis suggests that to understand their recent growth experiences, a deficit of new firms driven by the flight of manufacturing and local workers is an essential part of their story.

6. CONCLUSION

New firm creation is central to economic growth. This paper has studied how firm creation affects employment and wage growth at the local level, and thereby shapes the fortunes of cities and local economic areas. There are three key takeaways.

First, the local startup rate is essential to understanding the dynamics of local growth. Its persistent contribution to new jobs, combined with the uniform contraction of incumbents across space, places startups as the dominant driver of local growth experiences in a statistical sense. The absence of differential firm lifecycles across space only further serves to highlight the role of firm creation. At the same time, this absence provides a cautionary bound on the importance of density-based externalities and models of selection to explaining differences in firm output across space.

Second, one way to understand the persistence of the startup rate and its contribution to growth is through the interaction of increasing returns and labor mobility. At the local level, startups drive significant wage and employment growth, attracting new workers, whose demand then spurs further firm creation. Though this is an old idea in economic geography, estimation on new data here suggests that the quantitative feedback is strong.

Third, considering firm creation dynamics sheds new light on fundamental questions facing the U.S. economy. The decline of manufacturing has devastated the fortunes of many cities and towns throughout the U.S. This paper has emphasized that it is not only the direct effects of structural change which matter. Manufacturing's flight has spilled over onto firm creation decisions in other local sectors, compounding the pain significantly and trapping areas in a spiral of decline. Understanding these dynamics is the first step towards developing solutions.

REFERENCES

ACEMOGLU, D. AND P. RESTREPO (2019): "Robots and Jobs: Evidence from US Labor Markets," *Journal of Political Economy*, Forthcoming. [25]

ADAO, R., C. ARKOLAKIS, AND F. ESPOSITO (2019): "Spatial Linkages, Global Shocks, and Local Labor Markets: Theory and Evidence," Tech. rep., National Bureau of Economic Research. [3, 25]

AHLFELDT, G. M. AND E. PIETROSTEFANI (2019): "The Economic Effects of Density: A Synthesis," *Journal of Urban Economics*, 111, 93–107. [31]

ALLEN, T. AND C. ARKOLAKIS (2014): "Trade and the Topography of the Spatial Economy," *The Quarterly Journal of Economics*, 1085, 1139. [20]

ALLEN, T., C. ARKOLAKIS, AND Y. TAKAHASHI (2019): "Universal Gravity," *Journal of Political Economy*, Forthcoming. [3, 25]

AUTOR, D. AND D. DORN (2013): "The Growth of Low-skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, 103, 1553–97. [5]

AUTOR, D., D. DORN, AND G. H. HANSON (2013): "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, 103, 2121–68. [25]

BLACK, D. AND V. HENDERSON (1999): "A Theory of Urban Growth," *Journal of Political Economy*, 107, 252–284. [4]

CARLINO, G. A. AND T. DRAUTZBURG (2017): "The Role of Startups for Local Labor Markets," . [4]

COMBES, P.-P., G. DURANTON, L. GOBILLON, D. PUGA, AND S. ROUX (2012): "The Productivity Advantages of Large Cities: Distinguishing Agglomeration from Firm Selection," *Econometrica*, 80, 2543–2594. [58]

DAVIS, D. R. AND J. I. DINGEL (2019): "A Spatial Knowledge Economy," *American Economic Review*, 109, 153–70. [9]

DENT, R. C., B. W. PUGSLEY, H. WHEELER, ET AL. (2018): "Longitudinal Linking of Enterprises in the LBD and SSL," Tech. rep., Center for Economic Studies, US Census Bureau. [58]

DESMET, K., D. K. NAGY, AND E. ROSSI-HANSBERG (2018): "The Geography of Development," *Journal of Political Economy*, 126, 903–983. [4]

DURANTON, G. AND D. PUGA (2001): "Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products," *American Economic Review*, 91, 1454–1477. [9]

——— (2014): "The Growth of Cities," in *Handbook of economic growth*, Elsevier, vol. 2, 781–853. [4]

ECKERT, F., S. GANAPATI, AND C. WALSH (2019): "Skilled Tradable Services: The Transformation of US High-Skill Labor Markets," *Available at SSRN 3439118*. [61]

FABER, B. AND C. GAUBERT (2019): "Tourism and Economic Development: Evidence from Mexico's Coastline," *American Economic Review*, 109, 2245–93. [3]

FABERMAN, R. J. (2011): "The Relationship Between the Establishment Age Distribution and Urban Growth," *Journal of Regional Science*, 51, 450–470. [7]

FABERMAN, R. J. AND M. FREEDMAN (2016): "The Urban Density Premium Across Establishments," *Journal of Urban Economics*, 93, 71–84. [59]

FERNALD, J. G. (2015): "Productivity and Potential Output Before, During, and After the Great Recession," *NBER Macroeconomics Annual*, 29, 1–51. [37]

FORT, T. C. AND S. D. KLIMEK (2018): "The Effects of Industry Classification Changes on US Employment Composition," *Center for Economic Studies, U.S. Census Bureau.* [10, 49, 61]

FORT, T. C., J. R. PIERCE, AND P. K. SCHOTT (2018): "New Perspectives on the Decline of U.S. Manufacturing Employment," *Journal of Economic Perspectives*, 32, 47–72. [5, 25]

FUJITA, M. AND J.-F. THISSE (2003): "Does Geographical Agglomeration Foster Economic Growth? And Who Gains and Loses From It?" *The Japanese Economic Review*, 54, 121–145. [4]

GARCIA-MACIA, D., C.-T. HSIEH, AND P. J. KLENOW (2019): "How Destructive is Innovation?" *Econometrica*, 87, 1507–1541. [37]

GLAESER, E. L., H. D. KALLAL, J. A. SCHEINKMAN, AND A. SHLEIFER (1992): "Growth in Cities," *Journal of Political Economy*, 100, 1126–1152. [9]

GLAESER, E. L., S. P. KERR, AND W. R. KERR (2015): "Entrepreneurship and Urban Growth: An Empirical Assessment with Historical Mines," *Review of Economics and Statistics*, 97, 498–520. [4]

GLAESER, E. L., W. R. KERR, AND G. A. PONZETTO (2010): "Clusters of Entrepreneurship," *Journal of Urban Economics*, 67, 150–168. [4]

GOURIO, F., T. MESSER, AND M. SIEMER (2016): "Firm entry and macroeconomic dynamics: a state-level analysis," *American Economic Review*, 106, 214–18. [4]

GREANEY, B. (2019): "Distributional Effects of Uneven Regional Growth," *Working Paper*. [4]

HALTIWANGER, J., R. JARMIN, R. KULICK, AND J. MIRANDA (2016): "High Growth Young Firms: Contribution to Job, Output and Productivity Growth," *CES 16-49.* [5]

HALTIWANGER, J., R. S. JARMIN, AND J. MIRANDA (2013): "Who Creates Jobs? Small Versus Large Versus Young," *Review of Economics and Statistics*, 95, 347–361. [8, 12]

HUBMER, J. (2018): "The Race Between Preferences and Technology," *Working Paper*. [25]

JACOBS, J. (1970): *The Economy of Cities*, vol. 584, Vintage. [9]

KRUGMAN, P. (1991): "Increasing Returns and Economic Geography," *The Journal of Political Economy*, 99, 483–499. [4, 15]

MOREIRA, S. (2015): "Firm Dynamics, Persistent Effects of Entry Conditions, and Business Cycles," . [5, 60]

MYRDAL, G. (1957): *Economic Theory and Under-Developed Regions*, G. Duckworth London. [4]

NAGY, D. K. (2017): "City Location and Economic Development," *Working Paper*. [4]

PETERS, M. (2016): "Refugees and Local Agglomeration - Evidence from Germany's Post-War Population Expulsions," *Working Paper*. [4]

PIERCE, J. R. AND P. K. SCHOTT (2016): "The Surprisingly Swift Decline of US Manufacturing Employment," *American Economic Review*, 106, 1632–62. [25]

PUGSLEY, B. W., P. SEDLACEK, AND V. STERK (2018): "The Nature of Firm Growth," . [58]

SAIZ, A. (2010): "The Geographic Determinants of Housing Supply," *Quarterly Journal of Economics*, 125. [34]

TARR, D. G. (1988): "The Steel Crisis in the United States and the European Community: Causes and Adjustments," in *Issues in US-EC trade relations*, University of Chicago Press, 173–200. [3]

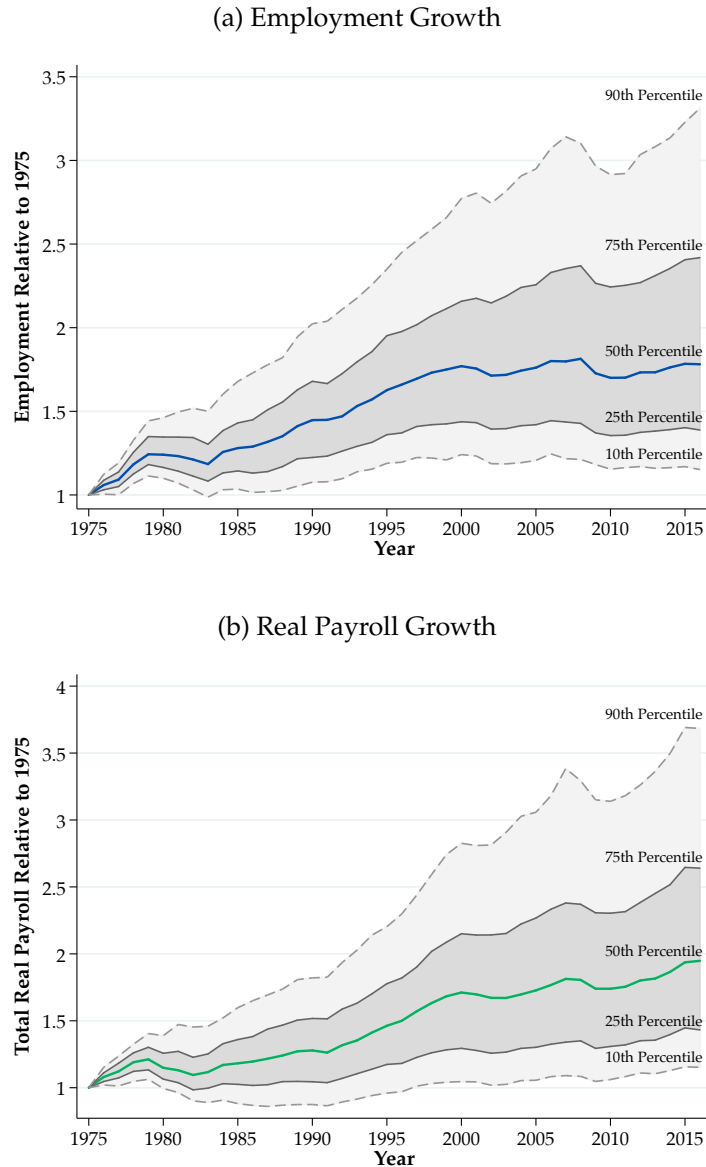TOLBERT, C. M. AND M. SIZER (1996): "US Commuting Zones and Labor Market Areas: A 1990 Update," *ERS Staff Paper*. [5]

Online Appendix for "The Startup Multiplier"

A. Empirical Appendix

A.1 Employment and Wage Growth at the Local Level

Figure 16: Distribution of Commuting Zone Level Employment and Payroll Growth

(a) Employment Growth

(b) Real Payroll Growth

Note: This figure shows an estimate of the distribution of cumulative growth in employment (Panel (a)) and payroll (Panel (b)) at the commuting zone level. The source data is the public-use County Business Patterns. Commuting zones are defined using the boundaries from the U.S. Census for the year 2000. Aggregation from county files to commuting zone level is done by the author. Percentiles of cumulative growth are calculated for each year, and averages are taken within percentiles. Real payroll for each commuting zone is calculated via deflating nominal payroll by the CPI in each year.

In this section of the Appendix I document the variability of growth at the local level in the U.S. over the last 40 years. I use the public-use County Business Patterns to construct a panel of 703 commuting zones which together form a complete partition of the U.S. I then examine how total employment and total payroll have evolved within each of these commuting zones over the past 40 years. Figure 16 shows how variable growth at the local level has been by representing the distribution of cumulative growth at the local level during this period. For employment, mean cumulative growth between 1975 and 2016 was 105%, with a standard deviation of 108%. The 90-10 ratio is substantial: at the 90th percentile, cumulative growth in employment was 231%, while at the 10th percentile it was 15%. Total real payroll growth shows even more dispersion across commuting zones. In Figure 17 I plot illustrative sample paths for some well-known areas.

Figure 17: Sample Commuting Zone Employment Growth Paths
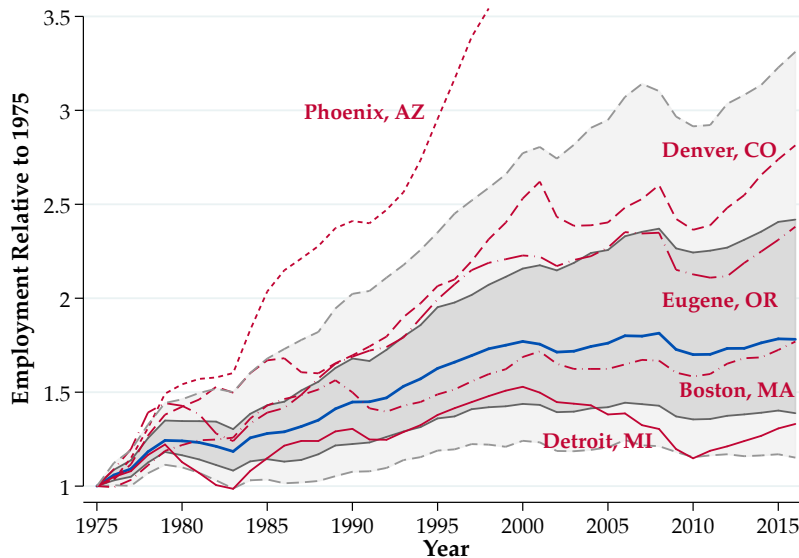
Note: This figure overlays sample employment growth paths for five commuting zones onto Panel (a) of Figure 16.
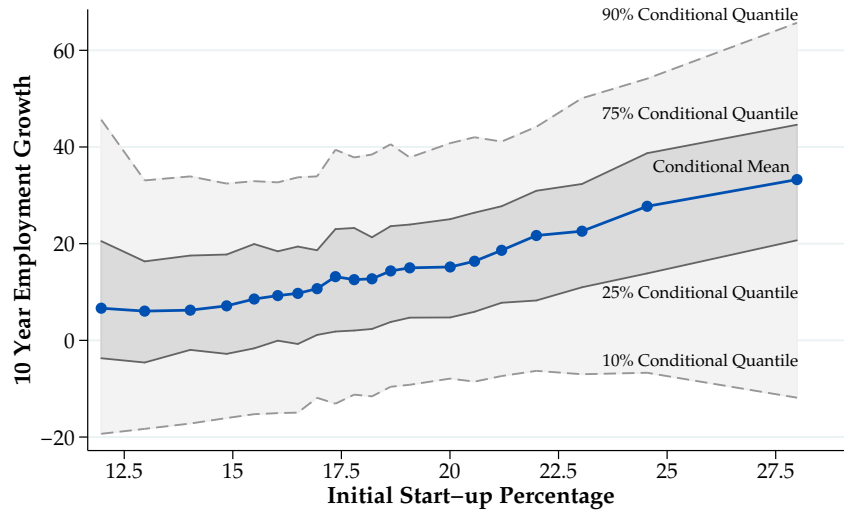
A.2   Employment Growth and The Startup Rate

This section of the Appendix explores in more detail the relationship between employment growth and the startup share presented in Figure 1.

In Figure 18 I show an estimate of the conditional distribution of 10-year employment growth at the local level against the startup share in that commuting zone. Census disclosure rules prevent the release of scatter plots, but much of the same information can be obtained from examining a representation of the conditional distribution. I summarize this by four quantile regressions (at the the 90%, 75%, 25% and 10% level) of 10-year forward employment growth on 20 quantiles of the

startup share, as well as reproducing the conditional mean function from Figure 1. While there is variation around the mean, the current start-up share robustly predicts growth in employment.

Figure 18: The Conditional Distribution of 10-Year Employment Growth at the CZ Level

Note: This figure shows an estimate of the conditional distribution of 10-year forward employment growth against the start-up percentage in a given year at the commuting zone. Conditional quantiles are computed by conducting quantile regressions of 10-year employment growth on 20 dummy variables for the startup percentage during. Employment growth is defined as the log change in total employment at the commuting zone level between year t and $t + 10$, and the startup percentage is the fraction of establishments who are either aged 0 or 1 year in year year t.

In Table 2, I estimate the relationship between local employment growth and the start-up percentage with controls. Panel A regresses employment growth in log changes at the local level from year t to year $t + 10$ on the fraction of establishments that are startups (establishments of age 0 or 1) in year t, and using the years 1980, 1990 and 2000. Panel B conducts the same analysis using employment growth from year t to year $t + 1$. The local unit of analysis is either commuting zone or county.

Columns I.A and II.A. regresses the log employment change on the startup fraction with no controls. I.B and II.B controls for industry employment shares in the location in year t, where these industries are defined as employment at the 1-digit NAICS level. NAICS codes are mapped to all establishments using the work of Fort and Klimek (2018). Columns I.C and II.C controls for the log of employment, and I.D and II.D includes local area fixed effects. In all cases, these start-up fraction is strongly associated with growth in both the next year, and the subsequent 10 years.

I now consider more detail in the conclusion of Fact 2. While Figure 2 decomposes growth across areas ranked by their initial startup percentage, it is instructive to ask how this picture changes when we look at areas ranked by their growth performance. In Figure 19 I repeat the estimation

Table 2: Estimates for Local Employment Growth By Start-Up Percentage

	Dependent Variable: Log Employment Change × 100							
	Commuting Zone Level				County Level			
Panel A: 10-Year Growth	I.A	I.B	I.C	I.D	II.A	II.B	II.C	II.D
Startup Fraction	156.2***	196.7***	195.0***	99.51***	180.6***	188.2***	189.8***	126.2***
	(14.12)	(15.44)	(15.74)	(10.38)	(12.98)	(14.94)	(15.19)	(10.38)
Log Employment			0.798*	-1.458***			-0.581	-2.412***
			(0.436)	(0.333)			(0.514)	(0.209)
1-digit Industry Emp. Shares	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Local FE (CZ or County)	No	No	No	Yes	No	No	No	Yes
R-squared	0.107	0.107	0.169	0.354	0.097	0.111	0.112	0.246
Observations	2,100	2,100	2,100	2,100	9,200	9,200	9,200	9,200
Panel B: 1-Year Growth								
Startup Fraction	27.53***	35.07***	35.45***	28.02***	25.72***	28.02***	28.23***	19.70***
	(0.9789)	(1.202)	(1.250)	(1.627)	(0.529)	(0.609)	(0.622)	(0.705)
Log Employment			-0.079**	-9.904***			-0.065***	-7.211***
			(0.040)	(0.271)			(0.026)	(0.114)
1-digit Industry Emp. Shares	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Local FE (CZ or County)	No	No	No	Yes	No	No	No	Yes
R-squared	0.045	0.055	0.055	0.233	0.031	0.034	0.034	0.170
Observations	24,000	24,000	24,000	24,000	104,000	104,000	104,000	104,000

Note: This table reports the estimated coefficients from regressing employment growth at the local level on the start-up percentage of the location, defined as the fraction of establishments of age 0 and 1. The sample period is from 1980-2005 for 10-year growth (for example, the observation for 2005 includes 10-year growth from 2005 to 2015). Observation counts have been rounded to accord with U.S. Census Disclosure rules. *** denotes significant at the 1% level, * denotes significant at the 5% level, * denotes significant at the 10% level.

of equation (1), with one key difference. In each year, I assign each commuting zone to one of 10 deciles according to their (forward) 10 year employment growth from that year. I then take the means of each term in equation (1) across each one of the ten deciles.

Figure 19: Entrant Contributions to 10-Year Employment Growth
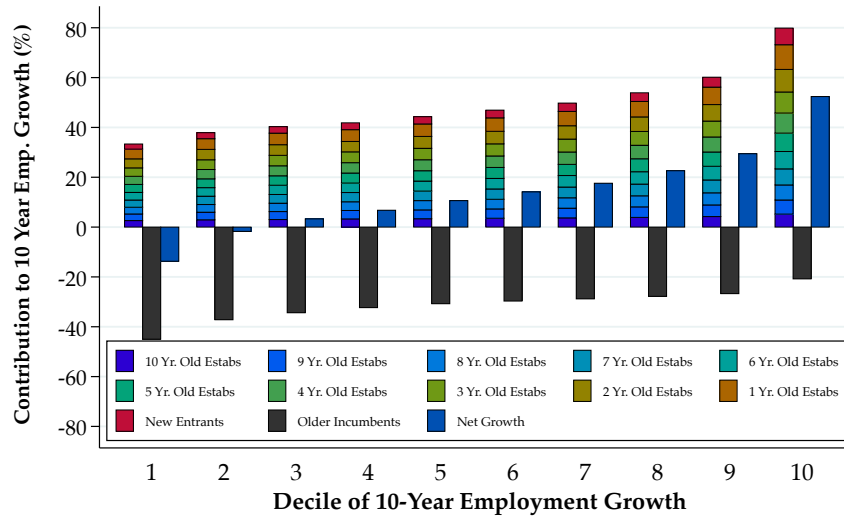
Note: This figure plots the contributions to future 10-year employment growth at the commuting zone level by successive cohorts of entrants, as well as current incumbents. Areas are assigned in each year to 10 deciles of 10-year employment growth. Decomposition is as in equation (1). Years are from 1990-2015 (for example, 10-year growth for 1990 is taken for the period 1980-1990). $N = \{24,000\}$, where this count has been rounded to accord with U.S. Census disclosure rules.

There is of course a larger degree of variation in total 10 year employment growth across these deciles than is apparent when splitting commuting zones by their startup percentage. However, the two conclusions I drew in Section 2 are sustained in this analysis. First, there is a strong contribution to total employment growth from each cohort of entrants that starts up during the 10 year period, with this contribution much greater in areas with high rates of employment growth. The persistence of the startup contribution is very salient; the higher deciles have high entrant contributions across all cohorts.
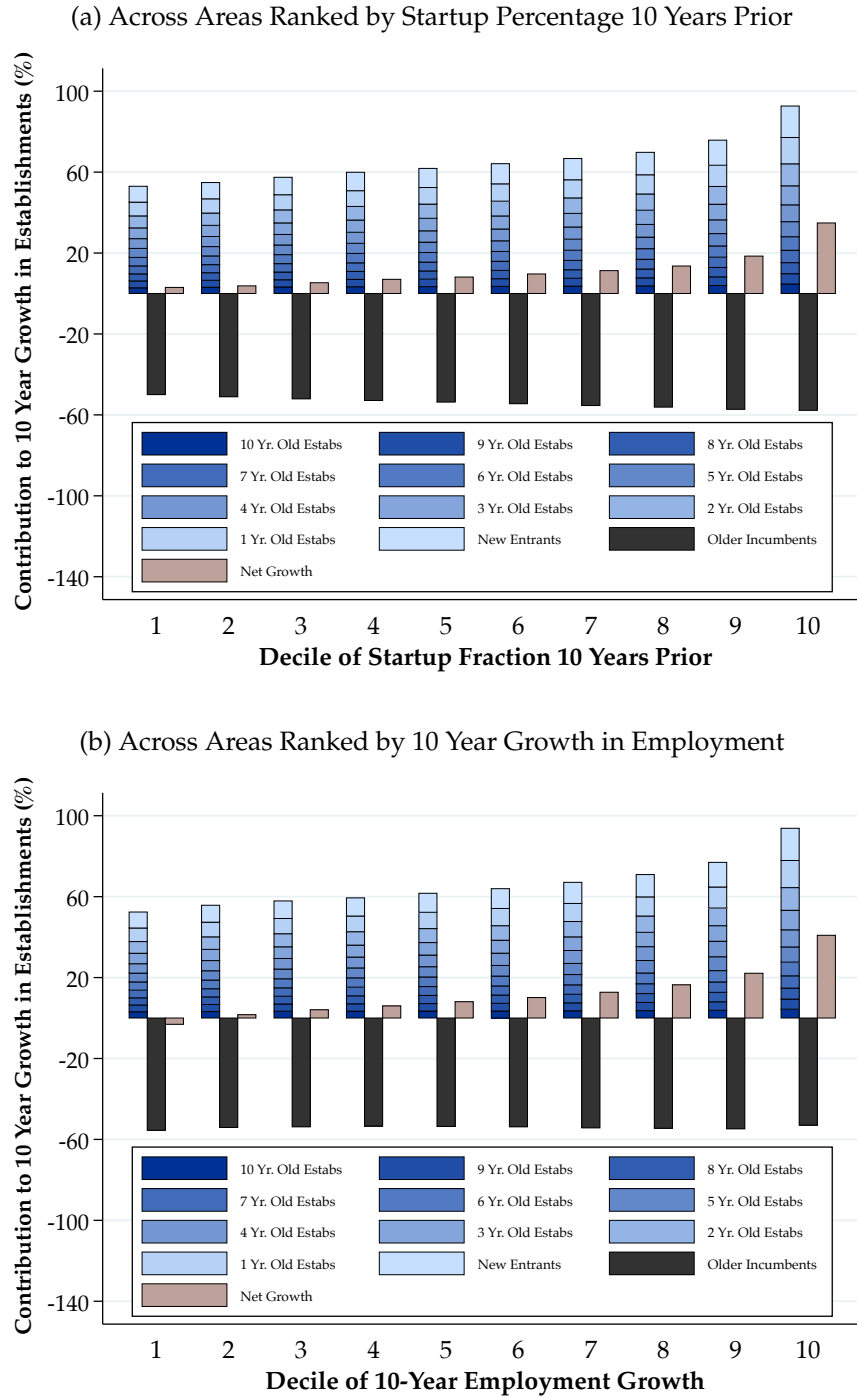
Second, it remains the case that incumbent establishments subtract from employment growth over the 10-year time horizon. However, split by growth deciles, there is more variation in the contribution of incumbents than when looking by the startup percentage in Figure 2.

To help understand why, I conduct a related exercise, and decompose total growth in the *number of establishments* at the commuting zone level over 10 year horizons. From this we can see what role differential exit of establishments is playing in Figures 19 and 2. In particular, for each commuting zone I write the total growth in the number of establishments as

$$(24) \quad \frac{Nestabs_{i,t} - Nestabs_{i,t-10}}{Nestabs_{i,t-10}} = \sum_{a \in \{0,1,\ldots,10\}} \frac{Nentrants_{i,t}^a}{Nestabs_{i,t-10}} + \left(\frac{Nincumb_{i,t}}{Nestabs_{i,t-10}} - 1 \right)$$

where $Nestabs_{i,t}$ is the number of establishments in commuting zone i in year t, $Nentrants_{i,t}^a$ is the number of surviving entrants of age a who entered during the 10 year period, and $Nincumb_{i,t}$ is the number of incumbents who existed 10 years prior and are still surviving by period t.

Figure 20: Decomposition of Growth in the Number of Establishments Across Cohorts

(a) Across Areas Ranked by Startup Percentage 10 Years Prior

(b) Across Areas Ranked by 10 Year Growth in Employment

Note: This figure shows a decomposition of 10-year growth in the number of establishments at the commuting zone level into contributions from incumbents and successive cohorts of entrants. Panel (a) splits the decomposition by the deciles of the startup percentage 10 years prior, while Panel(b) splits by deciles of 10-year employment growth. Decomposition is as in equation (1). Years are from 1990-2015 (for example, 10-year growth for 1990 is taken for the period 1980-1990). $N = \{24,000\}$, where this count has been rounded to accord with U.S. Census disclosure rules.

I then assign each commuting zone to the same two groupings as above: 10 deciles of the startup percentage 10 years prior at the commuting zone level, and 10 deciles of 10-year employment growth at the commuting zone level, for the period 1990-2015. I average each of the terms of the right hand side of equation 19 within these deciles. The results are plotted in Figure 20.

The black bars are effectively the 10-year exit rate of incumbent establishments. This is strikingly uniform across areas in both decompositions. Effectively, incumbent establishments exit at the same rate over a 10-year horizon regardless of the employment or entry dynamics.

A.3 Wage Growth and The Startup Rate

In Table 3 I repeat the regressions in Table 2, but now with 10-year growth in average wages at the local level as the dependent variable. The central right hand side variable continues to be the startup percentage in year t, and growth is taken to be the log change in average wages from year t to year $t + 10$.

Table 3: Estimates for Local Wage Growth By Start-Up Percentage

| | *Dependent Variable*: Log Wage Change $\times$ 100 | | | | | | | |
| | Commuting Zone Level | | | | County Level | | | |
10-Year Growth	I.A	I.B	I.C	I.D	II.A	II.B	II.C	II.D
Startup Fraction	91.41***	104.3***	34.23***	46.31***	55.13***	57.9***	30.39***	32.68***
	(4.22)	(4.32)	(4.12)	(6.60)	(2.60)	(2.70)	(2.95)	(2.76)
Log Employment		-1.40***	-0.82***	-1.16		-1.25***	-1.22***	-1.17*
		(0.079)	(0.081)	(1.25)		(0.06)	(0.06)	(0.58)
1-digit Industry Emp. Shares	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Local FE (CZ or County)	No	No	No	Yes	No	No	No	Yes
R-squared	0.075	0.114	0.377	0.647	0.029	0.0496	0.203	0.522
Observations	11,000	11,000	11,000	11,000	50,000	50,000	50,000	50,000

Note: This table reports the estimated coefficients from regressing average wage growth at the local level on the start-up percentage of the location, defined as the fraction of establishments of age 0 and 1. The sample period is from 1980-2005 for 10-year growth (for example, the observation for 2005 includes 10-year growth from 2005 to 2015). Average wage at the local level is computed by aggregating total local payroll and dividing by total employment in the local area. Observation counts have been rounded to accord with U.S. Census Disclosure rules. *** denotes significant at the 1% level, * denotes significant at the 5% level, * denotes significant at the 10% level.

A.4 Autocorrelation of the Startup Rate

Figure 21: Autocorrelation of Local Startup Rate
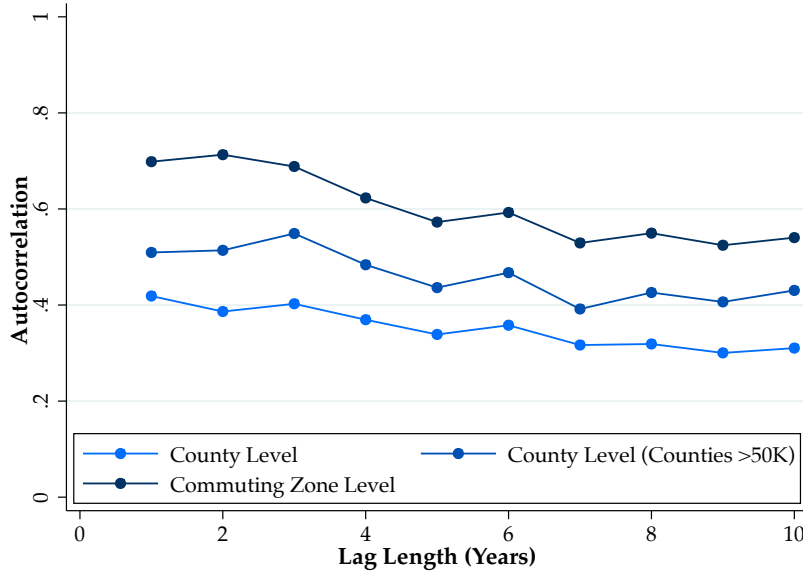
Note: This figure plots estimates of the autocorrelation of the startup rate at the local level at different time horizons. The startup rate is defined as the fraction of establishments who are of age 0at time t. Autocorrelation is computed across the panel of commuting zones, using the correlation of all observations of the startup-rate in time t with the same variable at time $t - x$, for $x = \{1, ..., 10\}$. Data comes from the LBD. $N = \{17,500; 76,000; 113,000\}$, where these counts has been rounded to accord with U.S. Census disclosure rules.
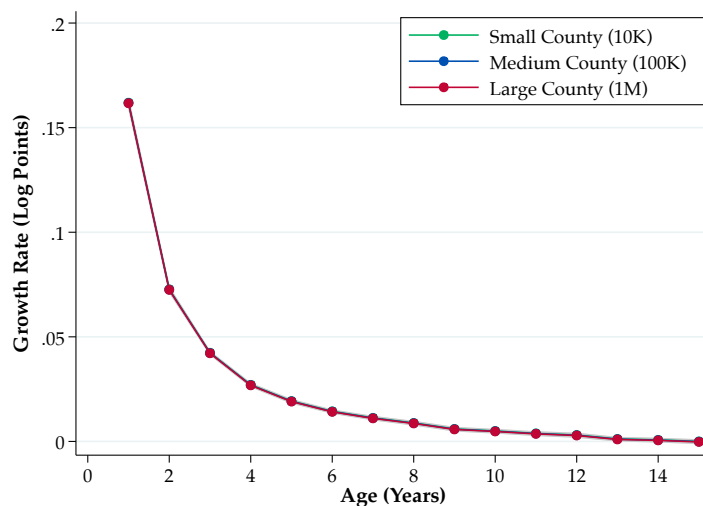
A.5 Establishment and Firm Growth Rates

Figure 22 reproduces Figure 4 with the population of the county in the year of the establishment's birth as the local population. The coefficient β_a is restricted to be the same for all ages. The results are unchanged- establishments do not grow faster in denser areas.

Figure 23 plots the results from reestimating equation (2) allowing β_a to vary by age. The confidence intervals include the decreased precision for the estimated growth rates by age (not reported in Table 4). Overall, it appears that establishments grow slightly faster in denser areas for the first four years of life, and then slightly slower for the rest of their life. However, for each size classification considered (small, medium or large counties), the estimated confidence intervals overlap. When considered over the first 15 years of life, establishments do not grow systematically faster in denser areas.

In Figure 24 I estimate the same equation with firms as the unit of analysis. Firms can grow their employment in two ways: intensively within the establishments they own, or extensively by adding new establishments. I consider only total employment on the LHS of equation (2). Population on the RHS is now population at the time of birth, and in the place of birth, for firms who

Figure 22: Establishment Growth Rates Across Space: County Estimates

Note: This figure plots the estimated growth rates of establishments by age in three different categories of county population size using the results from Table 4 . 95% confidence intervals in grey. $N = 104,000,000$, where this count has been rounded to accord with U.S. Census disclosure rules.

Figure 23: Establishment Growth Rates Across Space

(a) CZ-level Estimates
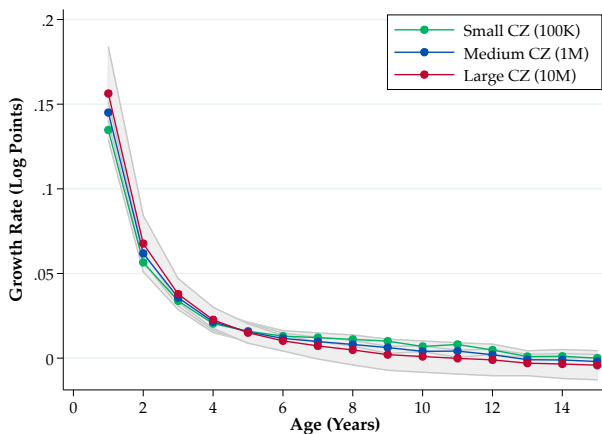
(b) County-level Estimates
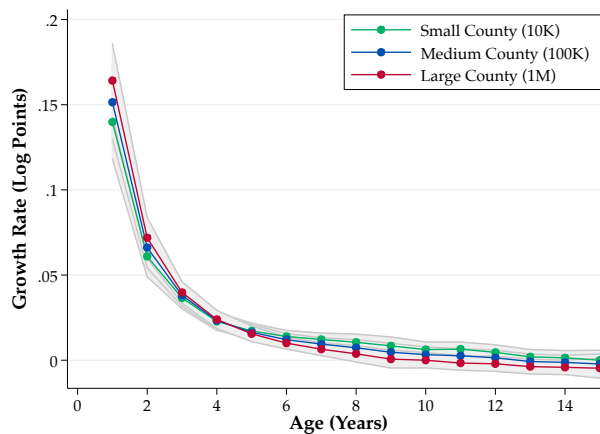
Note: This figure plots the estimates of growth by age at the establishment level from estimating equation (2). Panel (a) contains estimates for commuting zones, and Panel (b) for counties. Growth profiles are given for three different area sizes in each case. $N = 104,000,00$, where this count has been rounded to accord with U.S. Census disclosure rules.

Table 4: Estimates for Establishment Growth Rate Differences By Density

	Commuting Zone Population			County Population		
	I.A	I.B	I.C	II.A	II.B	II.C
Log Population	0.00931 (0.02042)			-0.00659 (0.02046)		
Log Population × Age = 1		0.48*** (0.068)	0.4884*** (0.062)		0.561*** (0.101)	0.578*** (0.095)
Log Population × Age = 2		0.262*** (0.043)	0.271*** (0.036)		0.244*** (0.055)	0.262*** (0.048)
Log Population × Age = 3		0.103*** (0.029)	0.112*** (0.022)		0.075*** (0.030)	0.093*** 0.024)
Log Population × Age = 4		0.057** (0.026)	0.067*** (0.019)		0.023 (0.025)	0.041** (0.019)
Log Population × Age = 5		-0.018 (0.024)	-0.008 (0.018)		-0.044*** (0.022)	-0.026 (0.017)
Log Population × Age = 6		-0.064*** (0.018)	-0.054*** (0.016)		-0.095*** (0.015)	-0.077*** (0.016)
Log Population × Age = 7		-0.115*** (0.021)	-0.105*** (0.016)		-0.140*** (0.016)	-0.122*** (0.016)
Log Population × Age = 8		-0.147*** (0.022)	-0.137*** (0.019)		-0.164*** (0.020)	-0.146*** (0.020)
Log Population × Age = 9		-0.188*** (0.019)	-0.178*** (0.021)		-0.186*** (0.020)	-0.168*** (0.025)
Log Population × Age = 10		-0.142*** (0.023)	-0.132*** (0.019)		-0.154*** (0.019)	-0.135*** (0.017)
Log Population × Age = 11		-0.194*** (0.019)	-0.184*** (0.018)		-0.196*** (0.017)	-0.178*** (0.019)
Log Population × Age = 12		-0.139*** (0.024)	-0.129*** (0.019)		-0.164*** (0.019)	-0.146*** (0.017)
Log Population × Age = 13		-0.095*** (0.020)	-0.085*** (0.020)		-0.140*** (0.017)	-0.140*** (0.022)
Log Population × Age = 14		-0.110*** (0.024)	-0.100*** (0.021)		-0.138*** (0.017)	-0.120*** (0.019)
Log Population × Age = 15		-0.100*** (0.024)	-0.090*** (0.026)		-0.115*** (0.023)	-0.098*** (0.028)
4-digit NAICS Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	No	Yes	Yes	No	Yes
R-squared	0.009	0.009	0.009	0.009	0.009	0.009
Observations	104,000,000	104,000,000	104,000,000	104,000,000	104,000,000	104,000,000

Dependent Variable: Log Employment Change × 100

Note: This table reports the results from estimating equation (2) at the establishment level on data from the LBD, for both county and commuting zone populations. Column *A* the density coefficients to be equal across establishment age. $N = \{104,000,000\}$, where these counts have been rounded to accord with U.S. Census disclosure rules.

begin as single-establishment firms of age zero. Age of the firm is defined as the age of the oldest establishment the firm owns.

Defining the birthplace of the firms that begin with a single establishment is complicated the process by which firm ID's are assigned in the the LBD. As discussed in Ding et al. (2019), firm ID's break by construction when firms grow or shrink in a particular way. In particular, when a single-establishment (SU) firm becomes/joins a multi-establishment (MU) firm, it is assigned a new firm ID. Likewise, if an existing MU firm loses all but one establishment, its firm ID will also change.

To account for this, I modify the LBD firm ID's according to the following rules:

a) If two or more SU establishments join an MU firm at t, the oldest establishment receives the MU firm ID since its birth. All other establishments keep their SU IDs until the year they join the MU firm.

b) In rule (a), if two or more establishments have the same maximum age at t, then the largest establishment receives the MU firm ID since its birth. All other establishments keep their SU IDs until the year they join the MU firm.

c) In rule (b), if two or more establishments have the same maximum size at t, then the firm is flagged as not possessing a birthplace, and dropped from the sample

d) In rule (a), if a younger establishment is 5 times larger in terms of total employees than the oldest establishment, then this largest establishment receives the MU ID since birth. All other establishments keep their SU IDs until the year they join the MU firm.

e) If a MU firm changes its MU firm ID at t, but possesses the same establishments in year t and $t-1$, then that new firm ID is replaced everywhere it appears by the original firm ID.

These rules allow the MU firm to be connected to a physical birthplace. Alternative algorithms for modifying the firm ID's are also available, developed in particular by Dent et al. (2018), and used for the first time by Pugsley et al. (2018). These rules focus on more comprehensive procedures for accounting for mergers and acquisitions at the firm level. In future research, I plan to explore the robustness of the conclusions of this section to using these alternative firm ID's.

As can be see in Figure 24, the conclusion for establishments holds true for firms as well. Firms born in denser areas do not grow their employment faster conditional on survival in a systematic way over the lifecycle.

A.6   Scale Elasticities

It is well known that establishments in denser areas are larger, pay higher wages and have higher sales even controlling for detailed industry differences (see e.g. Combes et al. (2012)). Fact 3c

Figure 24: Firm Growth Rates By Place of Birth

(a) CZ-level Estimates

(b) County-level Estimates
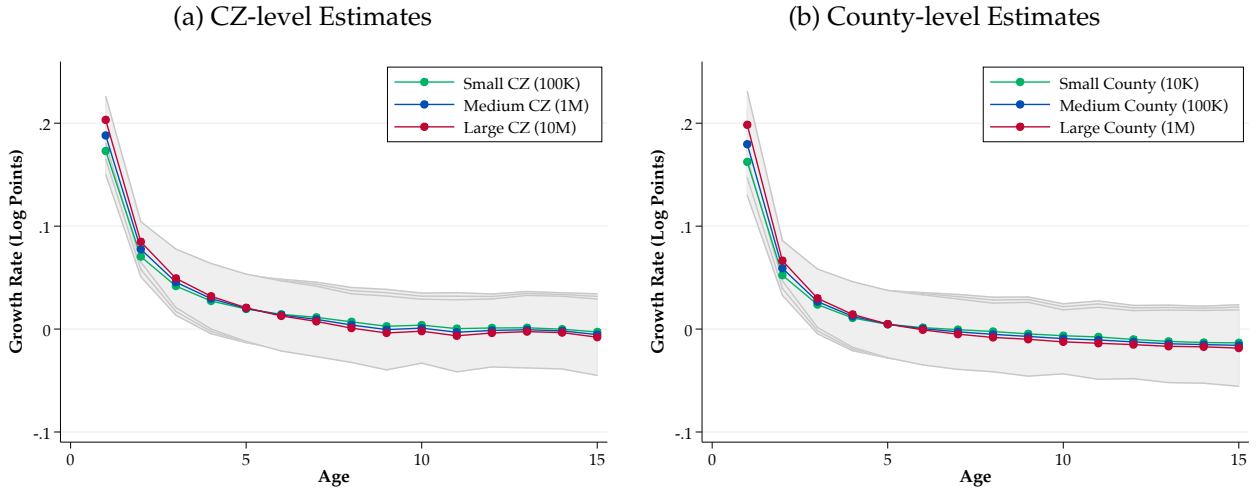
Note: This figure plots the estimates of growth by age at the firm level from estimating equation (2). Panel (a) contains estimates for commuting zones, and Panel (b) for counties. Growth profiles are given for three different area sizes in each case. $N = 76,000,00$, where this count has been rounded to accord with U.S. Census disclosure rules.

shows that these differences are constant with age. To establish this fact, I obtain the density elasticity of establishment scale by age through a full set of non-parametric dummies by age and detailed industry controls. In particular, I estimate the following series of equations:

$$(25) \qquad y_{it} = \sum_{a=0}^{A} \gamma_a \mathbb{I}_{Age_{it}=a} + \sum_{a=0}^{A} \beta_a \mathbb{I}_{Age_{it}=a} \times log(Pop_i) + Ind_i + \mu_t + X'_{it}\delta + \epsilon_{it}$$

Here y_{it} is log employment of the establishment, log average wages, and log sales per worker. The right hand side variables are as in equation (2). While equation (2) measures growth, equation (25) measures how scale evolves with age.

Our interest is again in the coefficients β_a, which describes how the elasticities of establishment scale variables with respect to local population evolve as the establishment ages. These are plotted for the full sample in Figure 25 for the three outcome variables, at both the commuting zone and county level. Confirming earlier studies, in the LBD establishments tend to be larger, pay higher wages and have higher sales per worker in more populous areas. However, in all three cases, these elasticities are approximately constant over the lifetime of an establishment. As such, not only are establishments in more populous areas not growing their employment faster (as outlined in Fact 3a), they are not seeing faster growth in payroll or sales.

The result for average wages is already known in the urban literature, and was first documented by Faberman and Freedman (2016). I extend that result to a holistic analysis of the evolution of establishment scale by density. Together, these facts suggest that the post-entry life-cycle of an establishment does not differ in densely populated areas in an appreciable way.

Figure 25: Scale Elasticities by Age

(a) Commuting Zone Level
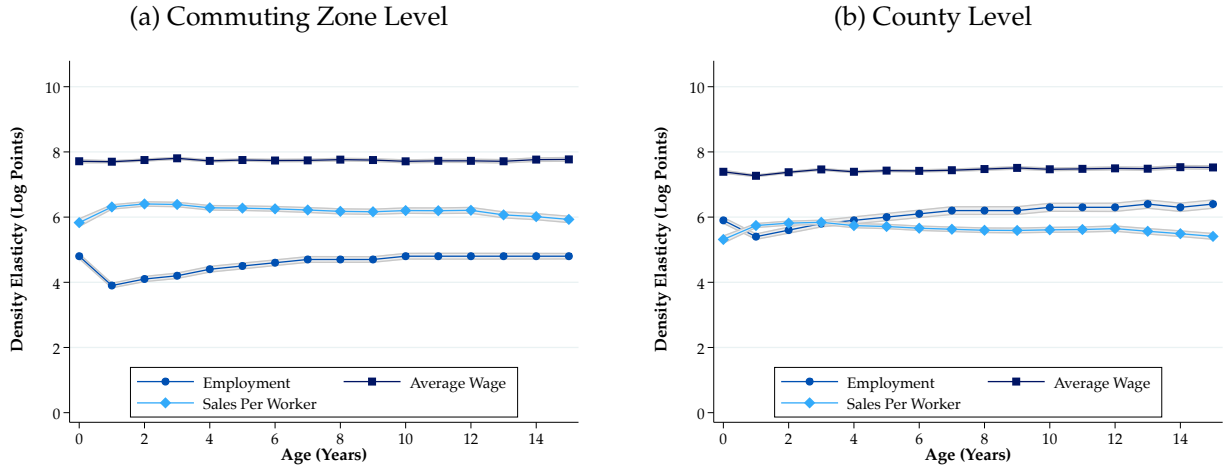
(b) County Level

Note: This figure plots the density coefficients by age from estimating Equation (25), at the commuting zone level (Panel (a)) and the county level (Panel (b)). Sales Per-Worker is reported for single-establishment firms; results using all establishments are quantitatively very similar. 95% confidence intervals in grey. $N = \{125,000,000; 125,000,000; 47,000,000\}$ for Employment, Average Wage and Sales Per Worker respectively, where these counts have been rounded to accord with U.S. Census disclosure rules.

While both equation (2) and equation (25) control for mean differences in industry growth rates and industry scale, they do not do so by age. As a final check, I estimate scale elasticities for individual industry groups, rerunning equation (25) separately for each group. The output is given in Figure 26. The headline conclusion continues to hold: establishment scale evolves in parallel across areas of different density over the lifecycle.

Constructing the scale elasticity for sales per-worker involves some work to determine establishment level sales, which is not a variable which appears in the current release (though there are plans to make this variable available to approved researchers). Instead, a measure of revenue in the Business Register is available for most firms that appear in the LBD from the year 1992 onwards. I follow the matching process of Moreira (2015) to match these revenues to the Longitudinal Business Database.

For single-establishment firms, the match is straightforward. Revenue can be matched on the basis of Employer Identification Number. Following this procedure results in a revenue variable for 86% of firms. Considering sales-per worker for multi-unit establishments is complicated by the fact sales are only observed at the firm level, and based on administrative tax records. The vast majority of firms in the Business Register consist of only a single establishment, and I report the results for only a single establishment firm first.

However, multi-establishment firms account for large fractions of output and employment, and so an attempt is also made to include these in the analysis. I consider two primary methods of apportioning sales across multi-establishment firms. First, I assign sales to each establishment in proportion to the employment of each establishment. Second, I assign them in proportion to the

Figure 26: Employment Scale Elasticities by Age: Industries
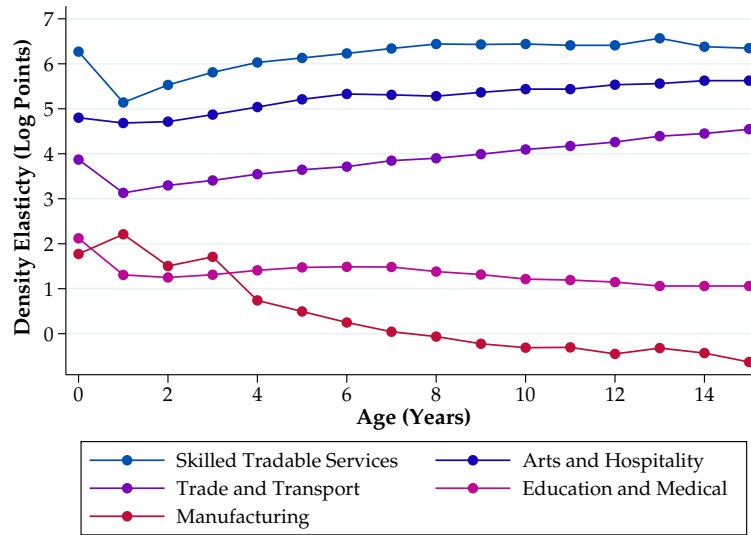
Note: This figure plots the density coefficients by age from estimating Equation (25) separately for major groupings of two digit NAICS codes, using the Fort and Klimek (2018) concordance to map NAICS codes to SIC codes. Skilled Tradable Services are NAICS 51-55, and include Finance, Information and Management of Companies, as defined in Eckert et al. (2019). Trade and Transportation is composed of NAICS 42-29, and includes Retail, Wholesale and Transportation and Warehousing. Arts and Hospitality is composed of NAICS 71 and 72, while Education and Medical is Sectors 61 and 62. Manufacturing is composed of NAICS 31-33.Estimation is done at the commuting zone level.

wage bill of the establishment. Results are virtually identical under both specifications.

In Figure 26 I repeat the estimation of scale elasticities at the establishment level for five major sectoral groupings: Skilled Tradable Services, Arts and Hospitality, Trade and Transport, Education and Medical, and Manufacturing. The estimation uses the population at the commuting zone level. An enumeration of the NAICS codes corresponding to each one of these groupings is found in the footnotes of Figure 26.

No systematic pattern of how the scale elasticities evolve with age is observable. A slight increase with age is seen for the Arts and Hospitality and Trade and Transport, while decreases are seen for Education and Medical and Manufacturing. These more detailed estimates give scant reason to change the overall conclusions of Fact 3. While establishments start out larger in more populous areas across industries, they do not appear to grow systematically faster, or see their scales increase more rapidly in these areas.

A.7   Commuting Zone and County Partition Summary Statistics

In Table 5 I report summary statistics for the partition of the set of commuting zones and counties employed throughout the paper. This partition splits the geographic units into 10 groups, with approximately 10% of the U.S. population in each group. The split is done using populations in

Table 5: Commuting Zone and County Partitions Summary Statistics

	Commuting Zone			County		
Decile	Count	Mean Population	Example	Count	Mean Population	Example
1	435	64,519	Jackson, WY	1,898	14,818	Peach County, GA
2	115	242,717	Burlington, VT	565	49,750	Columbia County, PA
3	59	475,226	Savannah, GA	281	100,238	Napa County, CA
4	36	768,921	Ann Arbor, MI	159	176,522	Arlington County, DC
5	23	1,254,128	Salt Lake City, UT	92	303,893	Pulaski County, AR
6	16	1,755,426	Columbus, OH	57	491,407	Seminole County, FL
7	10	2,600,746	Denver, CO	39	720,857	Baltimore County, MD
8	7	3,770,746	Phoenix, AZ	27	1,040,510	Palm Beach County, FL
9	5	4,905,792	Houston, TX	16	1,705,231	Riverside County, CA
10	3	11,953,495	New York, NY	7	4,213,464	Cook County, IL

Note: This table reports summary statistics of the commuting zone and county partitions into ten deciles employed throughout the paper. These deciles are employed in Figure 6 for establishment size by age, and Figure 5 for establishment exit rates. County population data is taken from the U.S Census in 2000. Aggregation from county to commuting zone level is done by the author using the commuting zones defined by the Census Bureau for the year 2000.

the year 2000, which is approximately the midpoint of the data from the LBD employed in this study.

A.8 Establishment Survival Function Estimation

In this Section I report the survival function estimates corresponding to the exit rates in Figure 5. These employ the non-parametric Kaplan-Meier estimator. Establishments in the LBD are split into ten groups according to the size of the commuting zone or county in which they were born. I use the same ten groups as in Figure 6 and Table 5. I estimate the survival functions for all establishments born in the period from 1980 to 1995, and use data from 1980 to 2015 on their lifetime outcomes. Exit occurs when the establishment is flagged as an establishment death in the LBD. There is some truncation given that not all such establishments born between 1980 and 1995 have exited by 2015. The Kaplan-Meier estimator handles this truncation naturally. The results are shown in Figure 27.

Differences in exit rates between densely and sparsely populated areas are quite negligible. We can see in Panel (a) that exit rates are marginally higher at the start of life in large places. This is somewhat more pronounced than the results for counties in Panel (b). These differences lead to a

Figure 27: Survival Function Estimates

(a) Establishment Exit Rates by CZ Size

(b) Establishment Survival Function by CZ Size
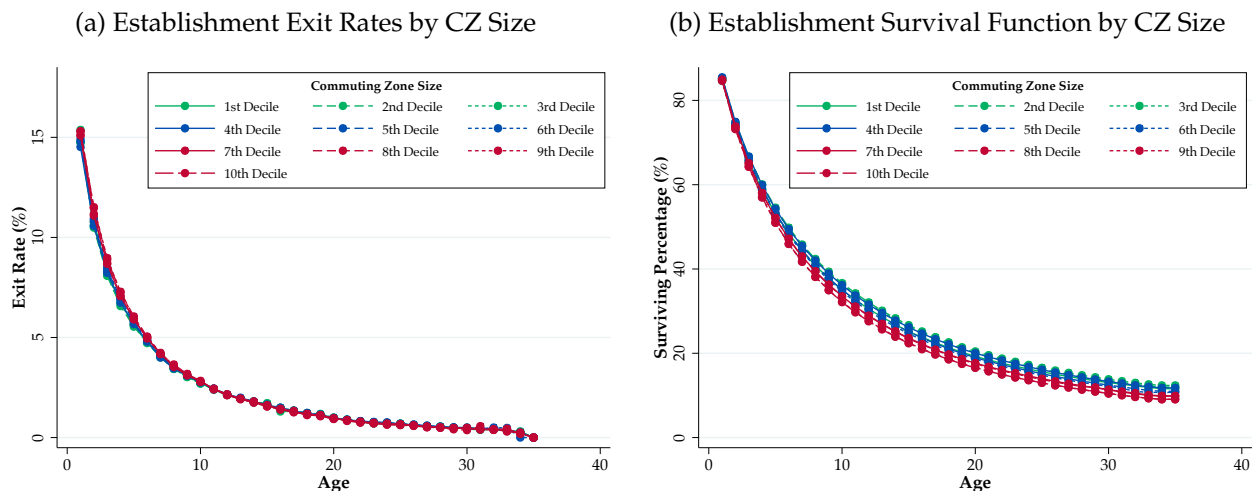
Note: This figure plots the estimated Kaplan-Meier survival functions for all establishments born between 1980 and 1995, across ten deciles of commuting zone or county size in which the establishment was born. $N = 125,000,000$, where this count has been rounded to accord with U.S. Census disclosure rules. Panels (a) plots the increments (exit rates) and Panel (b) presents the survival functions. Deciles of commuting zone and county sizes are labelled in ascending order (1st being the smallest), and correspond to the groupings summarized in Table 5.

small dispersion in estimated exit probabilities over the first 35 years of life: an establishment in the least dense group of commuting zones has around a 2% higher chance of survival than one more in the most populous group.

In the main text in Figure 5, the three groupings used correspond to Decile 2 (100k population), Decile 5 (1M population) and Decile 10 (10M population).
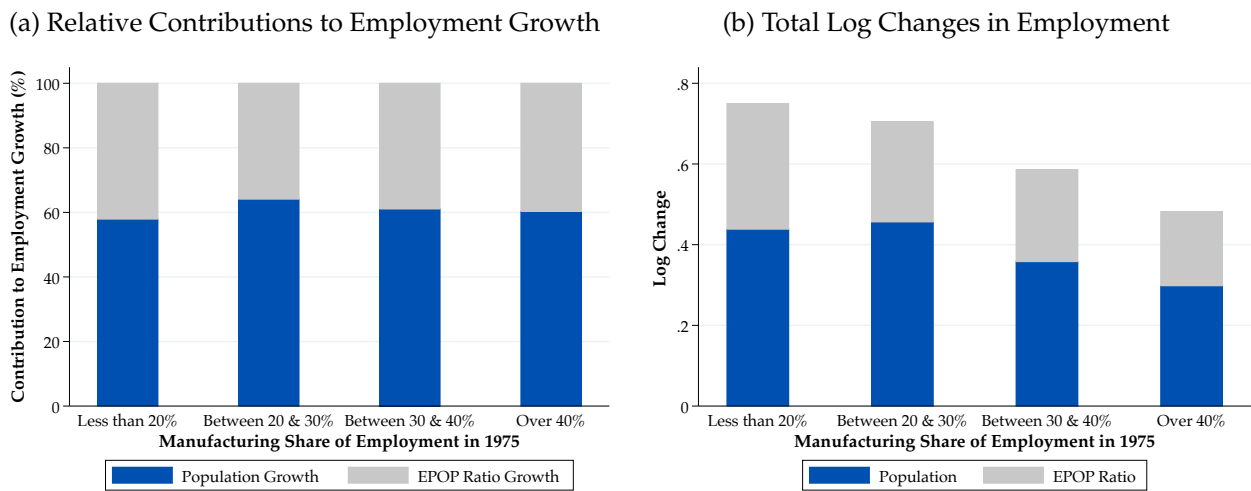
A.9 Local Employment Growth Contributions

In this section I decompose employment growth at the local level into two separate parts: a contribution from population growth, and another contribution from changing employment-population ratios. Between any two periods, we can write the log of employment growth as

$$
(26) \qquad log(\frac{E_t}{E_{t-1}}) = \underbrace{log(\frac{E_t}{P_t}) - log(\frac{E_{t-1}}{P_{t-1}})}_{EPOP \ Ratio \ Growth} + \underbrace{log(\frac{P_t}{P_{t-1}})}_{Population \ Growth}
$$

where E_t is the total private employment at the commuting zone level, and P_t is total adult population at the commuting zone level. Employment comes from the QCEW at the county level, and total adult population at the county level comes from the U.S. Census. I aggregate these to the commuting zone level.

Figure 28: Decomposition of Commuting Zone Employment Growth 1975-2015

(a) Relative Contributions to Employment Growth

(b) Total Log Changes in Employment

Note: Figure (a) shows the relative contributions to employment growth at the commuting zone level since 1975 of population growth and changes in the EPOP ratio, averaged across four groups of commuting zones depending on their manufacturing share of employment in 1975. These contributions are as defined in equation (26). Relative contributions are computed by dividing each term on the right-hand side of (26) by the total log change in employment since 1975. Figure (b) shows the total contribution both components of the right-hand side of (26). Data on population uses the U.S. Census population estimates at the county-level, which I aggregate to the commuting zone level using commuting zone boundaries in the year 2000. Data on employment and manufacturing specialization comes from the Quarterly Census of Employment and Wages produced by the Bureau of Labor Statistics, and employs the SIC implementation of industry classification.

I then average these contributions at the commuting zone level into four categories depending on the manufacturing share of employment in 1975. These are the same categories presented in Figure 9. Panel (a) of Figure 28 shows the mean contributions of Population Growth and EPOP Ratio growth as a fraction of total employment growth, and Panel (b) gives the mean log change from equation 26 across the four categories of manufacturing specialization.

During this period, the employment-population ratio grew markedly, due in part to the continued entry of women into the labor force. However, population growth was the largest contributor to employment growth across all regions, and this did not differ by manufacturing specialization in 1975.

A.10 Structural Estimation First Stage

In this section I present the first stage estimates for the model generated instruments for the numbers of manufacturing firms ($N_{j,t}^M$) and service firms ($N_{j,t}^S$). Recall that the moment conditions exploit variation in changes in these instruments. Figure 29 plots the log-changes in the actual numbers of firms at the commuting zone level over 5-year periods against the simulated numbers from the model.

The fit is quite strong, with an R^2 of around 0.2 in both cases. Intriguingly, the model-generated predictions neither under- or over-predict the true changes in numbers of firms on the micro level: the intercept is close to zero, and the slope is close to 1.

A.11 Structural Model Estimates

In this Section I report the parameter estimates for the structural model. In Table 6 I report the parameters describing the dynamics of the firm lifecycle, with the estimation as discussed in Section 4.3. I am able to match the first four moments of the empirical growth rate distribution in the Longitudinal Business Database exactly.

Table 7 records the structural production estimates: the local scale elasticities and the house price elasticities. OLS estimates are additionally given for the house price elasticities. Parameters are rather tightly estimated, where standard errors are computed using the delta method (which is necessary given the moment conditions in fact provide estimates of $(\sigma_S - 1)^{-1}$ etc.).

In Figure 30 I show the data which pins down the choice of the congestion parameter ζ. I again aggregate the Quarterly Census of Employment and Wages to the commuting zone level, and use it to compute average real wage and average establishment size in each commuting zone, for the years 1980-2015.

I then regress the log of both of these variables on the log of commuting zone employment, obtaining the coefficients $\hat{\beta}_w$ and $\hat{\beta}_{size}$. A representative plot from the year 2000 is shown in Figure 30. ζ

Figure 29: First-Stage Fit of Predicted Changes in Number of Firms Vs. Actual
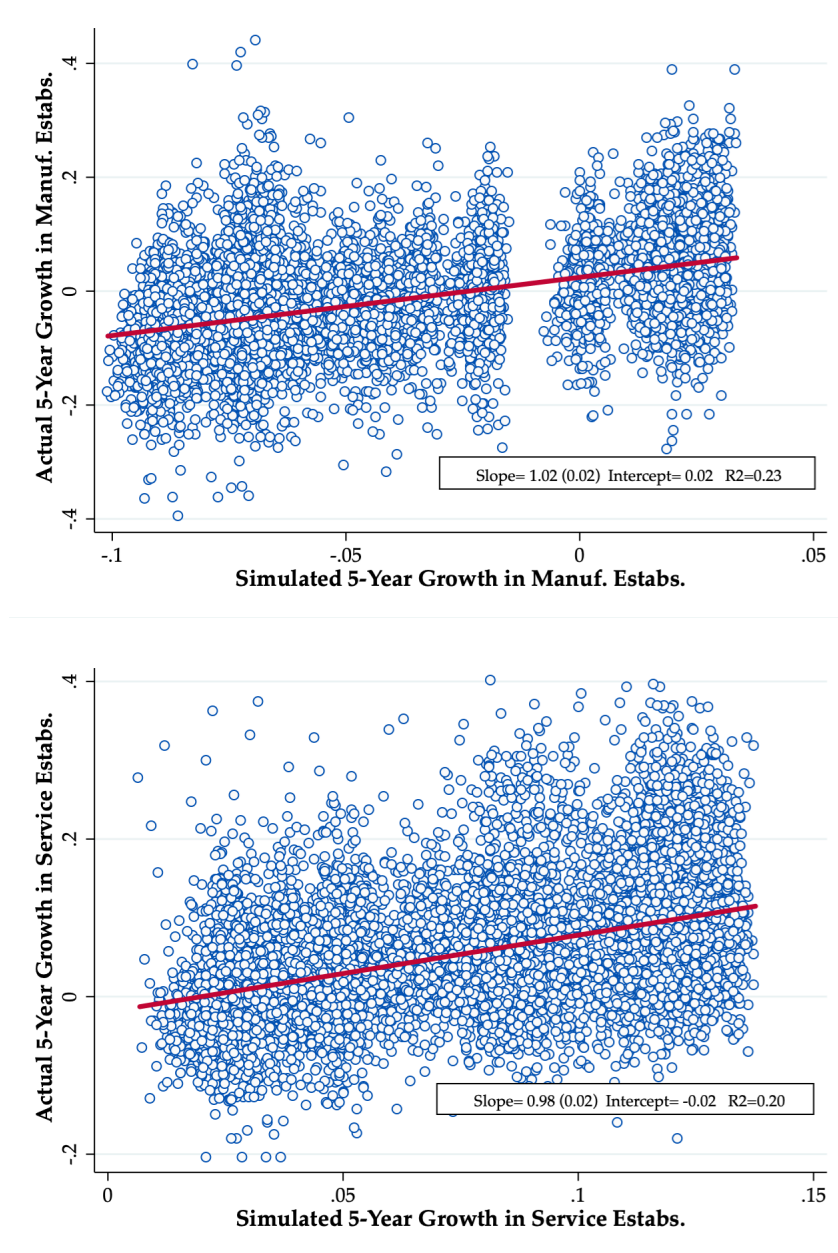
Note: This figure compares the model-based simulations of changes in manufacturing firms (above) and services firms (below) at five year time horizons to what actually occurred in the data from 1975 to 2015. The unit of observation is the commuting zone. Simulated values are generated assuming the economy is in a steady state in 1975, and then solving for the value of Π_t which exactly matches the series for the aggregate labor share in manufacturing.

Table 6: Structural Parameter Estimates for the Firm Lifecycle

| Parameter	Description	Baseline Value	Source/Moment
ϕ_u	Arrival rate of productivity improvement	0.482	Mean employment growth of 0.75%
Δ_u	Improvement step size	0.054	S.D in employment growth of 25%
ϕ_d	Arrival rate of productivity deterioration	0.498	Skewness in employment growth -0.05 %
Δ_d	Deterioration step size	0.036	Excess kurtosis in employment growth of 9 %
δ	Exogenous death rate of firms	0.1	Average Exit rate of 10%
γ	Curvature of utility for local capitalists	1	Set exogenously
ρ	Discount rate of local capitalists	0.05	Set exogenously

Note: This table reports the estimated baseline model parameters. Estimation is done via. Simulated Method of Moments, matching the first four moments of employment growth rates in the LBD, and the average death rate of establishments.

Table 7: Structural Production Parameter Estimates

Parameter	Description	Estimate	Standard Error	First-Stage F Statistic
	Production Parameters			
σ_M	Scale elasticity of wages for Manufacturing	2.36	(0.4)	642.3
σ_S	Scale elasticity of wages for Services	4.01	(0.95)	789.9
ϵ	E.O.S. between Manufacturing and Services	2.43	-	-
v	House price elasticity	0.92 (OLS=0.52)	(0.18)	348.2

Note: This table reports the estimated structural production parameter estimates. Estimation is done via. exploiting moment conditions based on aggregate structural change shifters. For the house price elasticities, OLS estimates are reported in brackets. Standard errors for $\{\sigma_S, \sigma_M, \epsilon\}$ use the delta method, given that the moment conditions provide estimates and asymptotic standard errors for $\{(\sigma_S - 1)^{-1}, (\sigma_M - 1)^{-1}, (\epsilon - 1)^{-1}\}$. Estimation is conducted for commuting zones with at least 50,000 people.

Figure 30: Cross Section of Average Wages and Average Establishment Size

(a) Average Wages By Area Size

(b) Establishment Size By Area Size
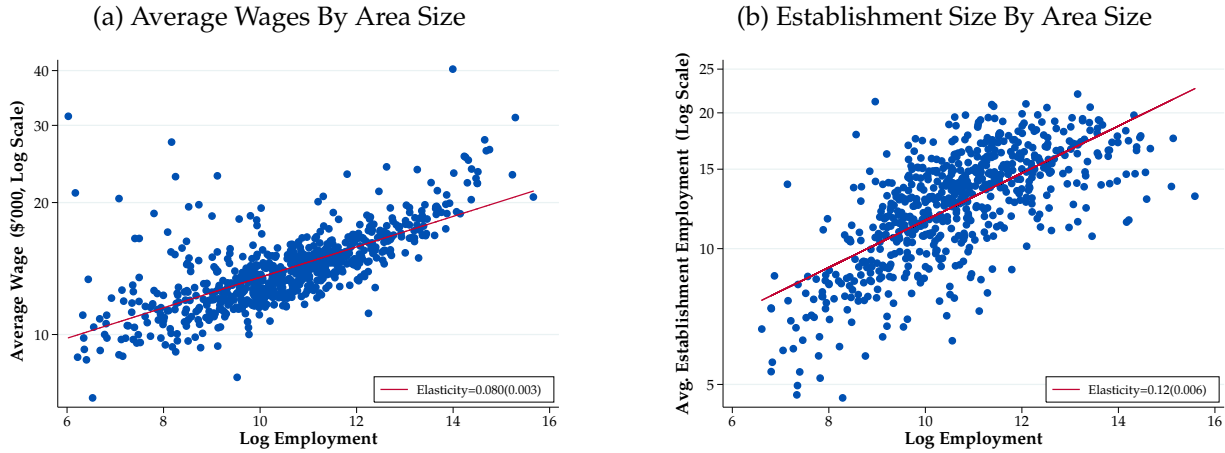
Note: Figure (a) shows average wage at the commuting zone level for the year 2000. Average wage is computed by dividing total payroll deflated by the U.S. CPI in the commuting zone by total employment. Figure (b) shows average establishment size at the commuting zone level for the year 2000. Average establishment size is computed by dividing total employment by the total number of establishments. Data for average wages is from the County Business Patters, and average establishment size is from the Quarterly Census of Employment and Wages produced by the Bureau of Labor Statistics.
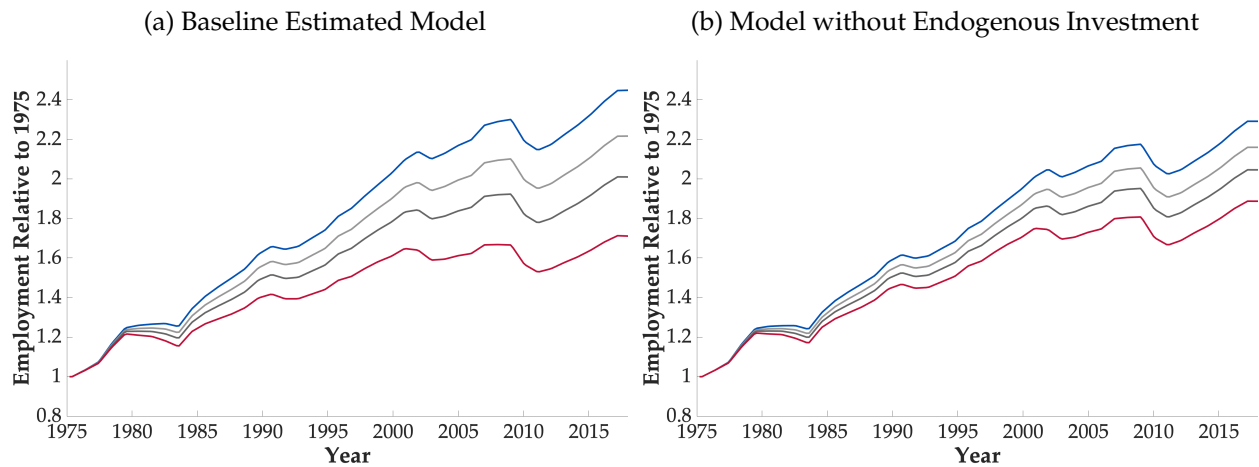
is solved from equation (17) as the root of

$$\hat{\beta}_{size} = \frac{1}{1+\zeta} \left(\zeta - \hat{\beta}_w \right)$$

A.12   Contribution of Firm Creation to Spatial Structural Change

Figure 31 provides the means of cumulative employment growth in each year in the four groupings of commuting zones considered in Figure 14.

Figure 31: Contribution of Startups to Spatial Structural Change

(a) Baseline Estimated Model
(b) Model without Endogenous Investment

Note: This figure plots the evolution of employment across commuting zones using the baseline estimated model under two scenarios. Panel (a) shows the model-implied growth in employment under the path for manufacturing productivity which matches the estimated decline in manufacturing employment. Panel (b) shows growth when fixing the number of firms in both manufacturing and services to their baseline values in 1975. The four color groups correspond to the four groups of 1975 manufacturing shares employed in Figure 9: blue (less than 20%), light grey (20-30%), dark grey (30-40%) and red (over 40%).

B. THEORY APPENDIX

B.1 Proof of Lemma 1

Write the discounted value of profits in sequence form as

$$V_{j,t}(z) = \mathbb{E}_t \int_t^\infty \delta e^{-\delta(T-t)} \left[\int_t^T e^{-\int_t^s r_v dv} \frac{1}{\sigma} \left(\frac{B_{j,s}}{w_{j,s}} \right)^{\sigma-1} Y_{j,s} z_s ds \right] dT$$

where the expectation is taken over the stochastic random variable z_s (all other variables are known with perfect foresight). At time s, z_s is given by

$$z_s = z_t D_s^u D_s^d$$

where (D_s^u, D_s^d) are independent counting processes taking on values on the countable set $Y^u \times Y^d$, where

$$Y^u \equiv \{1, \Delta_u, \Delta_u^2, \ldots\} \quad Y^d \equiv \{1, \Delta_d, \Delta_d^2, \ldots\}$$

with joint probability mass function

$$P(D_s^u = Y_n^u, D_s^d = Y_m^d) = \frac{(\phi_u s)^n}{n!} e^{-\phi_u s} \cdot \frac{(\phi_d s)^m}{m!} e^{-\phi_d s}$$

and similarly for D_s^d, where Y_n^u and Y_m^d represent the nth and m^{th} elements of Y^u and Y^d respectively. Hence $D_s^u D_s^d$ is independent of z_t. Furthermore, each firm is infinitesimal, and the paths for aggregate variables are taken as given from the point of view of the firm. So we can write

$$V_{j,t}(z) = z \mathbb{E}_t \int_t^\infty \delta e^{-\delta(T-t)} \left[\int_t^T e^{-\int_t^s r_v dv} \frac{1}{\sigma} \left(\frac{B_{j,s}}{w_{j,s}} \right)^{\sigma-1} Y_{j,s} D_s^u D_s^d ds \right] dT$$

$$= \nu_{j,t} z$$

Sufficient conditions for finite values. For simplicity we will assume aggregate labor L_t is constant; growing aggregate labor supply can be easily incorporated. First note that $Y_{j,s}$ is bounded above by $\frac{\sigma-1}{\sigma} w_{j,s} L$, where L is aggregate labor supply, so we can write

$$V_{j,t}(z) \leq Lz \mathbb{E}_t \int_t^\infty \delta e^{-\delta(T-t)} \left[\int_t^T e^{-\int_t^s r_v dv} \frac{1}{\sigma} (B_{j,s})^{\sigma-1} w_{j,s}^\sigma D_s^u D_s^d ds \right] dT$$

Now supposing the value of the firm is finite, by Fubini we can interchange the integrals to get

$$= Lz \int_t^\infty \delta e^{-\delta(T-t)} \left[\int_t^T e^{-\int_t^s r_v dv} \frac{1}{\sigma} (B_{j,s})^{\sigma-1} w_{j,s}^\sigma L_s \mathbb{E}_t[D_s^u D_s^d] ds \right] dT$$

$$= Lz \int_t^\infty \delta e^{-\delta(T-t)} \left[\int_t^T e^{-\int_t^s r_v dv} \frac{1}{\sigma} (B_{j,s})^{\sigma-1} w_{j,s}^\sigma L_s e^{(\phi_u(\Delta_u-1)+\phi_d(\Delta_d-1))s} ds \right] dT$$

$$= Lz \int_t^\infty \delta e^{-\delta(T-t)} \left[\int_t^T e^{-\gamma \int_t^s g_v dv} e^{-\rho s} \frac{1}{\sigma} (B_{j,s})^{\sigma-1} w_{j,s}^\sigma e^{(\phi_u(\Delta_u-1)+\phi_d(\Delta_d-1))s} ds \right] dT$$

using the Euler equation of the local capitalist. Now we impose some simple assumptions on location fundamentals. Suppose that:

a) There exist numbers K_1 and k such that

$$\max_{j}\{B_{j,s}\} \le K_1 e^{ks}$$

for all s.

b) There exists some K_2 such that

$$\max_{j,k}\{\frac{B_{j,s}}{B_{k,s}}\} < K_2$$

for all s.

c) $B_{j,s} > F_s$ for some weakly increasing surjection $F_s : \mathbb{R}_+ \to \mathbb{R}_+$

Together, (b) and (c) can be shown to imply $\lim_{s\to\infty} e^{-\gamma \int_t^s g_v dv} = 1$. Moreover, using (a) we have

$$V_{j,t}(z) \le Lz \int_t^\infty \delta e^{-\delta(T-t)} \left[\int_t^T e^{-\gamma \int_t^s g_v dv} e^{-\rho s} \frac{1}{\sigma} (K_1 e^{ks})^{\sigma-1} (N_{j,t} \bar{Z}_{j,t})^{\frac{\sigma}{\sigma-1}} e^{(\phi_u(\Delta_u-1)+\phi_d(\Delta_d-1))s}] ds \right] dT$$

If $N_{j,t}$ is not vanishing in the limit (which can be shown to be implied by conditions (a)-(c)), then $\bar{Z}_{j,t}$ is bounded from above. Moreover, $N_{j,t}$ cannot grow faster than $B_{j,s}$ in the limit without violating feasibility. As such, there is some K_3 such that

$$V_{j,t}(z) \le K_3 Lz \int_t^\infty \delta e^{-\delta(T-t)} \left[\int_t^T e^{-\gamma \int_t^s g_v dv} e^{-\rho s} \frac{1}{\sigma} (e^{ks})^{\sigma-1+\frac{\sigma}{\sigma-1}} e^{(\phi_u(\Delta_u-1)+\phi_d(\Delta_d-1))s}] ds \right] dT$$

Then it can be directly verified this integral is finite if

$$\delta + \rho > \sum_{i\in u,d} \phi_i(\Delta_i - 1) + k(\sigma - 1 + \frac{\sigma}{\sigma-1})$$

In words, the value of the firm will be finite if the sum of the discount and death rates is larger than the average growth rate of firm efficiency and a multiple of the maximum growth in local productivity.

B.2 Proof of Proposition 2

In a stationary equilibrium, the mass of firms is constant, which requires $N^e_{j,t}/N_{j,t} = \delta$ for all locations. Given that both the mass of firms and the entry rate are constant, the consumption of

local capitalists is also constant, which requires $r_{j,t} = \rho$. Lastly, average efficiency $\bar{Z}_{j,t}$ is constant in each location. Using the law of motion in (14), this implies that $\bar{Z}_{j,t}$ is the same across areas, and given by

$$\bar{Z}_{j,t} = \frac{\delta}{\Phi} \bar{z}^E$$

Combining this with the results of Proposition 1, we get an expression for payroll across areas, which satisfies

$$\frac{w_{j,t} L_{j,t}}{N_{j,t}^{1+\zeta}} = \delta^\zeta (\rho - \Phi)(\sigma - 1)\tau \bar{z}^{E-1}$$

Substituting both of these expression into equation (27), we can write

(27)
$$w_{j,t} = k_1 B_{j,t}^{\frac{(1+\zeta)(\sigma-1)}{(1+\zeta)(\sigma-1)-1}} \left(L_{j,t} \right)^{\frac{1}{(1+\zeta)(\sigma-1)-1}}$$

where k_1 is combination of model constants. Utility maximization among workers must imply that

$$L_{j,t}^{\bar{v}_j} = w_{j,t}^{1-\bar{v}_j} A_{j,t} \bar{h}_j^{\bar{v}_j} \bar{U}^{-1}$$

for all locations, where $\bar{U}$ is the equilibrium level of utility in the stationary economy. Combining this with (27) and dropping time subscripts, we find that L_j is pinned down as a function of the realizations of location fundamentals, in

(28)
$$(L_j)^{\frac{\bar{v}_j}{1-\bar{v}_j} - \frac{1}{(1+\zeta)(\sigma-1)-1}} = A_j^{\frac{1}{\bar{v}-1}} h_j^{\frac{\bar{v}}{\bar{v}-1}} \frac{\bar{U}^{\frac{1}{\bar{v}-1}}}{k_1} B_j^{\frac{(1+\zeta)(\sigma-1)}{(1+\zeta)(\sigma-1)-1}}$$

Suppose that $\frac{\bar{v}_j}{1-\bar{v}_j} > \frac{1}{(1+\zeta)(\sigma-1)-1}$. Then we can use this expression to write

$$\frac{L_j^{\bar{v}_j - \frac{1-\bar{v}_j}{(1+\zeta)(\sigma-1)-1}}}{L_k^{\bar{v}_k - \frac{1-\bar{v}_k}{(1+\zeta)(\sigma-1)-1}}} = \frac{(k_1 B_j^{\frac{(1+\zeta)(\sigma-1)}{(1+\zeta)(\sigma-1)-1}})^{1-\bar{v}_j} A_j \bar{h}_j^{\bar{v}_j}}{(k_1 B_j^{\frac{(1+\zeta)(\sigma-1)}{(1+\zeta)(\sigma-1)-1}})^{1-\bar{v}_k} A_k \bar{h}_k^{\bar{v}_k}}$$

and together with the labor market clearing conditioning, this equation pins down unique values for L_j as a function of all location fundamentals. This is the unique stationary equilibrium. In this equilibrium, using (27) the conditional log wage function at time 0 in the steady state (before long run fundamentals are realized) is

$$\mathbb{E}[log(w_j)|log(L_j)] = log(k_1) + \frac{1}{(1+\zeta)(\sigma-1)-1} log(L_j) + \frac{(1+\zeta)(\sigma-1)}{(1+\zeta)(\sigma-1)-1} \mathbb{E}[log(B_j)|log(L_j)]$$

For average firm size, recall we have

$$\frac{w_{j,t} L_{j,t}}{N_{j,t}^{1+\zeta}} = \lambda$$

using equation ([15](#)). As such we can can write

$$(1+\zeta)log(L_{j,t}/N_{j,t}) \propto \zeta log(L_{j,t}) - log(w_{j,t})$$

from which the result follows immediately.

If $\frac{\bar{v}_j}{1-\bar{v}_j} \leq \frac{1}{(1+\zeta)(\sigma-1)-1}$, it can be shown that utility does not equalize across areas, such that equation ([28](#)) does not hold, and in equilibrium only a single location will be inhabited.

Proof of Lemma 2. To derive the conditional expectation of $log(B_{j,t})$ given employment in a location, under the assumptions in Lemma 2, we make the following observations. Note that the vector

$$\begin{bmatrix} log(L_j) \\ log(B_j) \end{bmatrix} = c \underbrace{\begin{bmatrix} log(\frac{\bar{U}^{\frac{1}{\bar{v}-1}}}{k_1}) \\ 0 \end{bmatrix}}_{\equiv b} + c \underbrace{\begin{bmatrix} \frac{1}{\bar{v}-1} & \frac{\bar{v}}{\bar{v}-1} & \tilde{\sigma} \\ 0 & 0 & 1 \end{bmatrix}}_{\equiv \Gamma} \begin{bmatrix} log(A_j) \\ log(h_j) \\ log(B_j) \end{bmatrix}$$

where $c \equiv (\frac{\bar{v}_j}{1-\bar{v}_j} - \frac{1}{(1+\zeta)(\sigma-1)-1})^{-1}$ and $\tilde{\sigma} \equiv \frac{(1+\zeta)(\sigma-1)}{(1+\zeta)(\sigma-1)-1}$, is an affine transformation of a multivariate normal, and is thus itself multivariate normal, such that

$$\begin{bmatrix} log(L_j) \\ log(B_j) \end{bmatrix} \sim \mathcal{N}(b + \Gamma \mu, \Gamma \Sigma \Gamma').$$

Then the conditional expectation of $log(B_j)$ takes a simple form, and is given by

$$\mathbb{E}[log(B_{j,t})|log(L_{j,t})] = \mu_B + \Gamma \Sigma \Gamma'_{21} \Gamma \Sigma \Gamma'^{-1}_{11} (log(L_{j,t}) - c \, log(\frac{\bar{U}^{\frac{1}{\bar{v}-1}}}{k_1})$$
$$- c(\frac{1}{\bar{v}-1}\mu_A + \frac{\bar{v}}{\bar{v}-1}\mu_h + \tilde{\sigma}\mu_B))$$

Some algebra reveals that, in the case that h_j is constant across space,

$$\frac{d}{log(L_{j,t})}\mathbb{E}[log(B_{j,t})|log(L_{j,t})] = \frac{\frac{1}{1-\bar{v}}\Sigma_{BA} + \tilde{\sigma}^2\Sigma_B^2}{(\frac{1}{\bar{v}-1})^2\Sigma_A^2 + \frac{\tilde{\sigma}}{\bar{v}-1}\Sigma_{BA} + \tilde{\sigma}^2\Sigma_B^2}$$

Linearized Dynamics. Before examining the full case use to prove local stability, I derive the expressions presented in Section ([3.2](#)). First, assume that there are no post-entry dynamics, so that $\bar{Z}_{j,t}$ is always constant at the mean entry productivity $\bar{z}^E$. Also, under risk neutrality of the local capitalist, we have $r_{j,t} = \rho$. To examine the local dynamics around the unique steady state, we combine equation ([10](#)) and ([13](#)) to get

$$(N^E_{j,t})^\zeta = \frac{\bar{z}^E A^{\frac{1}{\bar{v}_j}}_j B^{\frac{1-\bar{v}_j}{\bar{v}_j}}_j h}{\sigma \tau} N^{\frac{1}{\bar{v}_j(\sigma-1)}-1}_{j,t} \bar{Z}^{\frac{1}{\bar{v}_j(\sigma-1)}-1}_{j,t} \frac{1}{\rho - \zeta g^{n^e}_{j,t} - \Phi} U_t$$

where U_t is the equilibrium level of utility in the economy which ensures labor market clearing across space. In the text we suppose only one location is out of steady state, and that this location is small, such that we U_t is close to its steady state value of $\bar{U}$. Linearizing this expression around the unique steady state yields

$$\zeta(N_{j,t}^E - \bar{N}_j^E) = \delta(\frac{1}{\bar{v}_j}\frac{1}{\sigma - 1} - 1)(N_{j,t} - \bar{N}_j) + \frac{\zeta}{\rho - \Phi}(\dot{N}_{j,t}^E - \dot{\bar{N}}_j^E)$$

where $\bar{N}_j$, for example, denotes the steady state number of firms in that location. Noting that $\dot{N}_{j,t} = N_{j,t}^E - \delta N_{j,t}$, we can write this as a second order, homogenous linear differential equation in $N_{j,t}$, as

$$\frac{\zeta}{(\rho + \delta)}(\ddot{N}_{j,t} - \ddot{\bar{N}}_j) - \frac{\zeta\rho}{\rho + \delta}(\dot{N}_{j,t} - \dot{\bar{N}}_j) + \delta(\frac{1}{\bar{v}_j(\sigma - 1)} - 1 - \zeta)(N_{j,t} - \bar{N}_j) = 0$$

The solution is given by

$$N_{j,t} = (N_{j,0} - \bar{N}_j)e^{xt} + \bar{N}_j$$

$$x = 0.5(\rho \pm \sqrt{\rho^2 + 4\delta\frac{(\rho + \delta)}{\zeta}(1 + \zeta - \frac{1}{\bar{v}_j(\sigma - 1)})})$$

Given that $\left(\frac{1}{\bar{v}_j(\sigma-1)} - 1 - \zeta\right) < 1$, the solution for $N_{j,t}$ has one stable and one unstable root. Standards arguments imply that the explosive solution will violate the transversality condition of the local capitalist. As such, the solution for this special case has

$$N_{j,t} = (N_{j,0} - \bar{N}_j)e^{xt} + \bar{N}_j$$

$$x = 0.5(\rho - \sqrt{\rho^2 + 4\delta\frac{(\rho + \delta)}{\zeta}(1 + \zeta - \frac{1}{\bar{v}_j(\sigma - 1)})})$$

Now we consider the full case with CRRA utility of the capitalist. From market clearing in the final good, the consumption of the capitalist in location j must satisfy

$$C_{j,t} = \frac{1}{\sigma}B_t(N_t\bar{Z}_t)^{\frac{1}{\sigma-1}}L_t - \tau(N_{j,t}^E)^\zeta$$

where the right hand side is dividends net of entry costs. The linearized Euler equation of the capitalist requires

$$\dot{C}_t = \bar{C}\frac{1}{\gamma}(r_{j,t} - \bar{\rho})$$

Inserting this into the free entry condition gives us

$$(N_{j,t}^E)^\zeta = \frac{\bar{z}^E A^{\frac{1}{\bar{v}_j}}B_j^{\frac{1-\bar{v}_j}{\bar{v}_j}}h}{\sigma\tau}N_{j,t}^{\frac{1}{\bar{v}_j(\sigma-1)}-1}Z_{j,t}^{\frac{1}{\bar{v}_j(\sigma-1)}-1}\frac{1}{\gamma\frac{\dot{C}_{j,t}}{\bar{C}_j} + \rho - \zeta g_{j,t}^{ne} - \Phi}U_t$$

I linearize this equation ignoring deviations of $\bar{Z}_{j,t}$ and $\bar{U}_t$ from their steady state values, and then show that doing so is valid for economies sufficiently close to the stationary equilibrium in all locations. Doing so yields

$$
\zeta(N_{j,t}^E - \bar{N}_j^E) = \delta(\frac{1}{\bar{v}_j}\frac{1}{\sigma-1} - 1)(N_{j,t} - \bar{N}_j) + \frac{\zeta}{(\rho-\Phi)}(\dot{N}_{j,t}^E - \dot{\bar{N}}_j^E)
$$

(29)
$$
- (\frac{\bar{N}_j^E}{\bar{C}_j})\frac{\bar{\gamma}}{(\rho-\Phi)}\left(\frac{1}{\bar{v}_j}\frac{1}{\sigma-1}\bar{N}_j^\zeta(\dot{N}_{j,t} - \bar{N}_j) - \tau(N_j^E)^{\zeta-1}(\dot{N}_{j,t}^E - N_{j,t})\right)
$$

Unlike in the simple case above, the solution to this equation now depends on location, and quantities such as the steady state entry flow $\bar{N}_j^\zeta$. However, it can be verified that in all locations, the solution to equation (29) $N_{j,t}$ has one stable and one unstable root.