The Role of Startups for Local Labor Markets

Gerald Carlino
Federal Reserve Bank of Philadelphia Research Department

Thorsten Drautzburg
Federal Reserve Bank of Philadelphia Research Department
The Role of Startups for Local Labor Markets

Gerald Carlino and Thorsten Drautzburg *

February 6, 2020

Abstract

There are substantial differences in startup activity across U.S. local labor markets. We study the causes and consequences of these differences. Startup productivity shocks are found to drive much of these cross-city differences in startup activity: They explain half of the forecast error variance of startup job creation, accounting for 40% of population growth and long-run changes in employment. Shocks to barriers to firm entry have economy-wide effects similar to those of startup productivity shocks but operate largely through the number of startups, rather than their size. We use a novel spatial panel VAR, identifying shocks using shift-share external instruments.

Keywords: Startups; entrepreneurship; local labor markets; proxy VAR; spatial panel VAR.

JEL classification: C33, C36, E24, L26, R11.

*Federal Reserve Bank of Philadelphia, 10 Independence Mall, Philadelphia, PA, 19106. Phone: (215) 574-6000. Email: jerry.carlino@phil.frb.org and tdrautzburg@gmail.com. An earlier version of this paper circulated as “All jobs are created equal? Examining the importance of startups for local labor demand”. This paper benefited from comments by Jonás Arias, Marco Del Negro (editor), Karel Mertens (discussant), Jonathan Wright, and audiences at the Federal Reserve Bank of Philadelphia, the 2016 AEA meetings, and the ERSA meetings. Catherine O’Donnell, Adam Scavette, Samuel Wascher, and Nick Zarra provided excellent research assistance. Any mistakes are our own. Thorsten Drautzburg gratefully acknowledges the hospitality of the Federal Reserve Bank of Chicago and the Becker-Friedman-Institute during the final stage of this project. The views expressed herein are our own views only. They do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia, the Federal Reserve System, or its Board of Governors.
1 Introduction

New business are important in the U.S. as a whole: In the average year, firms that were started within the last year create jobs for 1.11% of the population, more than the economy as a whole, which creates jobs for 0.68% of the population. Haltiwanger et al. (2013) also show that conditional on survival, young firms grow faster than incumbents, contributing to productivity and wage growth across the economy. However, there are substantial differences in the contribution of startup activity to overall employment at the sub-national level, defined as metropolitan statistical areas (MSAs). Startups in the median MSA create jobs for 1.04% of the population in the average year, compared to 0.68% for all firms. But startups in the top 10% of MSAs, such as Austin, TX and Los Angeles, CA, contribute jobs for at least 1.40% of the population, compared to just 0.75% or less in the bottom 10%, which are mostly smaller MSAs. In this paper, we study the causes and consequences of the differences in startup activity among MSAs over time.

Startup activity can be due to the direct effects of shocks to startups or by the indirect response of startups to other economic forces, such as shocks to overall productivity in the economy. We analyze to what extent startup activity within U.S. MSAs is driven by startup shocks and what the dynamic effects of startup shocks are for these local economies.\footnote{Past research has considered the role of demographics and labor supply for startup activity (Karahan et al., 2015), as well as differences in investment opportunities (Adelino et al., 2014). For countries other than the U.S. however, it has been documented that policy shocks to startup activity, which lowered the cost of entry, can stimulate startup activity (Hombert et al., 2014; Maggi and Felix, 2018; Kaplan et al., 2011).}

We define startup shocks as shocks to entrepreneurial productivity, which we define formally in Section 2 using the Lucas (1978) span of control model. Via their effect on labor demand, startup shocks can affect the entire local economy.\footnote{We also account for shocks to overall labor demand originating, for example, from shocks to common productivity.}

Our results indicate that much of the cross-city differences in startup activity are indeed driven by startup shocks. But not only are startup shocks an important determinant of startup activity directly, they also account for much of the cross-MSA variation in population growth and thus employment. Specifically, startup productivity shocks account for about half of the forecast error variance in startup activity, when measured as the change in the job creation of startups relative to the stock of employment in our baseline results. Our finding lines up with structural estimates of aggregate models. Sedlacek and Sterk (2017) and Drautzburg (2019) both estimate a large role of direct startup shocks that move both startup employment and average startup size in the macro data. Our findings suggest that similar forces are at play at the local level. But our estimates also leave room for other
shocks to propagate to startup activity. Demographic factors (Karahan et al., 2015; Engbom, 2019) are one prominent example. Structural models should thus consider shocks with both strong direct and indirect effects on startups.

A typical, i.e., a positive one-standard deviation, shock to startup labor demand increases startup employment initially by 0.6% of overall employment. This translates to 0.2 percent of the population, or about two thirds of difference in startup job creation between an MSA in the bottom 10% of the distribution and the median MSA. The initial effect on overall employment is ambiguous, as local economies largely reallocate jobs from incumbents to startups, as in Hombert et al. (2014). Not many new firms enter, but the ones that do are bigger. The formation of these new firms makes the local areas more attractive and leads to immigration and, in turn, an increase in population and overall employment. Eventually, overall employment rises by around 1% and then remains stable. The migration and population growth response we find also distinguishes startup labor demand shocks from overall labor demand shocks, which are shorter lived and, possibly because of their lower persistence, do not induce immigration and population growth.\(^3\) Our findings suggest that there may be scope for policies to directly promote entrepreneurship: policies that generate a positive startup productivity shock would lead to employment and population growth.

To arrive at our results, we estimate a spatial panel VAR for 354 MSAs, covering 1986 to 2013. The estimated model includes time fixed effects, so that the model speaks to the differences between MSAs, but not to national trends. Correspondingly, differences across the MSAs are the identifying variation in the model. The panel VAR is spatial, because it estimates a spatial error structure that accounts for the correlation of economic activity across space at any given time, similar to Mutl (2009). We employ a block bootstrap for inference. Our main results are based on impulse-response-functions and forecast error variance decompositions for a typical MSA.

To identify the startup productivity shock in the VAR, we use an external instrument, as in Stock and Watson (2012) and Mertens and Ravn (2013). To our knowledge, this is the first paper to identify shocks with external instruments in a panel VAR framework. Unlike other identification methods, such as recursive (Cholesky) identification schemes or sign restrictions, VARs with external instruments do not directly impose restrictions on impulse-response functions. Instead, the identifying

\(^3\)Karahan et al. (2015) show that demographic shocks matter for startups via the cost of labor. Our result is complementary and implies that startup shocks also affect demographics. Our finding that startup dynamics are associated with net immigration is similar to Walsh (2019), who uses a fully structural model.
assumptions correspond to the relevance and validity requirements of the instrumental-variables literature. Panel VARs are particularly promising for working with external instruments, because panel VARs allow researchers to use cross-sectional differences to construct these instruments. For example, Beaudry et al. (2014) construct several instruments using cross-sectional regional variation.

The external instrument that identifies the startup productivity shock in the VAR is the sum of sectoral, nationwide employment growth of startups, weighted by the local employment share. As an intermediate step, the VAR also identifies shocks to overall labor demand, using an instrument based on the sectoral employment growth of all firms (Bartik, 1991). This procedure is valid under two assumptions: (1) Historical predictors of labor demand based on sectoral, nationwide employment growth are uncorrelated with local labor supply shocks. Blanchard and Katz (1992) argue that this assumption is reasonable if sectors are not overly geographically concentrated. (2) Controlling for overall local labor demand, labor demand by startups is driven by startup productivity. We discuss the economics in the context of our model, and formalize the procedure in Proposition 2 in Section 3 below.\(^4\)

Relevance of the instruments requires an MSA in which manufacturing, say, has historically been strong to also see higher startup employment in a year in which manufacturing startups nationwide had unusually high employment growth. Adelino et al. (2014) also use a Bartik (1991) instrument, but they study only the response of startups to overall labor demand changes.

We also show that our identified historical startup productivity shocks are sensible. To make that argument, we bring in data on the growth of venture-capital (VC) funded firms by MSA and year and compare our identified startup shocks to these data. Our motivating model of entrepreneurship implies a close connection between the startup productivity shock and financing costs for startups and we treat the growth of venture capital funded firms as a proxy for the cost of financing to startups. Specifically, in the model the implications of shocks to entrepreneurial productivity and to the startup-specific cost of inputs are observationally equivalent given the information that we have on startups. When available, the proxy correlates significantly with the identified startup productivity shocks.

To further validate the identified startup productivity shock, we contrast it with a shock to barriers to firm entry. To identify a startup shock to barriers to entry, we use an alternative instrument in the main VAR. This alternative instrument identifies shocks to the startup entry rate, rather than startup

---

\(^4\)With a single external instrument in a standard VAR, the instrument identifies the impact IRF as proportional to the covariance between the instrument and the VAR forecast errors. Here, the VAR forecast errors are purged of both temporal and spatial correlation. To identify both shocks, we assume that one instrument reflects only the labor demand shock, while the other instrument also reflects the startup productivity shock and the overall labor demand shock.
job creation, and uses national cross-industry variation in entry rate, again in the place of startup job creation. While the alternative barriers to entry shock operates through the extensive margin, the startup productivity shock operates through the intensive margin of startups, i.e., through their average size, in line with the motivating model. This may be of interest to local policy-makers, since barriers to entry may be readily affected by policy, unlike startup productivity.

This paper is structured as follows: Section 2 summarizes our data on startups, presents the model we use to motivate our shock identification, and discusses how we construct the external instruments. Section 3 presents our statistical framework. Section 4 contains the main results on startup shocks and summarizes robustness checks.

2 Startups: Data, shocks, and instruments

This section defines startups and describes the data we use to measure startup activity. It then defines our concept of startup shocks in context of the Lucas (1978) model of firms. Building on the data description and the model, we develop the instruments that we use in the empirical analysis.

2.1 Startup data and definition

We define startups as private firms that are less than one year old. Our data on startups and other firms come from the U.S. Census Bureau’s Business Dynamics Statistics (BDS) and its County Business Patterns (CBP). The BDS data give information on the number of employees on payroll, the number of firms, and, for older firms, the number of exits and net job creation. For startups, net job creation is virtually identical to the number of employees. The data are available by firm age for 354 U.S. metro areas, allowing us to compute local startup activity. When using non-BDS data, we construct MSAs from county level data to align with the 2009 MSA definition underlying the BDS. Since employment is the only measure of the intensive margin of startup activity, we focus on the employment of all startups in a certain location over the past year.

Startup employment is important in the U.S. (Figure 1(a)): For the U.S. as a whole, startups created, on average, jobs for 1.11% of the population from 1977 to 2013. This compares to an average

---

5One may think that all startup employment represents new jobs. However, some employees at some new firms, such as spin-offs, may have transferred from existing business, generating small discrepancies between startup employment and startup job creation. We use startup job creation when available, but substitute missing data on startup job creation with startup employment.

6At the national level, we also observe startup activity by broad industry.
net job creation by all firms in the U.S. of 0.68%. Averaging across MSAs gives similar numbers, with startups creating, on average, jobs for 1.06% of the population.

While startups are important for job creation, the averages mask substantial heterogeneity across MSAs and over time. The histogram in Figure 1(a) shows that the top 10% of MSAs in terms of their average job creation create jobs for at least 1.40% of the population, compared to just 0.75% or less in the bottom 10% of MSAs. Figure 1(b) shows the distribution of startup job creation (relative to population) over time. The dashed lines show that, for example, in 1980, in 5% (25%) of MSAs startups created jobs for 0.5% (0.8%) of the MSA population, compared with 1.0% for the median MSA (solid line). In the top 5% (25%), in contrast, startups create jobs for 2.0% (1.4%) of the population or more in 1980. While there is a common decline over time, which has been analyzed in the literature, year fixed effects account for only 24% of the variation in startup activity, leaving three-quarters of the variation across time and space unexplained.\(^7\) Below, we analyze to what extent variation of startup activity is driven by shocks to startups, and what the effects of these shocks are.

### 2.2 Startup shocks in a model of entrepreneurship

To fix ideas about what shocks to entrepreneurial productivity are, we now introduce entrepreneurial productivity in the Lucas (1978) model, a canonical model of firms. In our adaptation of the model, firm \(i\) located in MSA \(m\) produces output \(Q_{i,m,t}\) at time \(t\) using labor input \(N_{i,m,t}\):

\[
Q_{i,m,t} = \left(\phi_{i,m,t} \times \chi_{m,t}\right)^{\alpha} \frac{N_{i,m,t}^{1-\alpha}}{1-\alpha}. \tag{1}
\]

Here, \(\phi_{i,m,t}\) represents the entrepreneur’s span of control – in equilibrium, higher skilled entrepreneurs manage a larger work force, hence the label. \(\phi_{i,m,t}\) varies by entrepreneur, but has a common cohort-MSA component. \(\chi_{m,t}\) represents productivity that is common to all firms in location \(m\) at time \(t\). The labor share in production is given by \(1 - \alpha\). Let \(\pi_{i,m,t}\) represent profits, which are just output minus the cost of the labor input:

\[
\pi_{i,m,t} = \max_{N_{i,m,t}} \left(\phi_{i,m,t} \times \chi_{m,t}\right)^{\alpha} \frac{N_{i,m,t}^{1-\alpha}}{1-\alpha} - (w_{m,t} \times \xi_{i,m,t}^{\alpha}) N_{i,m,t}. \tag{2}
\]

\(^7\)I.e., the R-squared of year fixed effects is 24%. MSA-specific fixed effects account for 34% of the variation.
Figure 1: The distribution of startup job creation across MSAs and over time

\( w_{m,t} \) is the local wage rate and \( \xi_{i,m,t}^{\alpha} \) (scaled for convenience) denotes the premium that entrepreneur \( i \) pays over the local wage rate – for example, this premium could reflect the cost that firm \( i \) might have to borrow to pay wages in advance.

If entrepreneurs hire workers to maximize profits, the following equation holds:

\[
N_{i,m,t}^* = \frac{w_{m,t}}{\phi_{i,m,t}} \chi_{m,t} \times \frac{\phi_{i,m,t}}{\xi_{i,m,t}}.
\]

Equation (3) indicates that entrepreneurs hire more workers the lower the local wage rate \( w_{m,t} \), the higher the local productivity \( \chi_{m,t} \), and the higher their span of control \( \phi_{i,m,t} \). Entrepreneurs also hire more workers if they face proportionally lower input prices \( \xi_{i,m,t} \). Since we can only observe employment for firms, we cannot distinguish between the span of control and lower input prices, and we summarize them by their ratio \( \frac{\phi_{i,m,t}}{\xi_{i,m,t}} \), which we refer to as entrepreneurial productivity. Firms’ profits move in the same direction as employment in response to changes in either of these variables.\(^8\)

Our empirical model uses data by MSA and firm age, but not for individual firms (since data at the firm level are proprietary). We therefore sum over all new firms to get the total employment of

---

\(^8\)Profits are \( \pi_{i,m,t}^* = \frac{\alpha}{1-\alpha} (\phi_{i,m,t} \times \chi_{m,t})^{\alpha} N_{i,m,t}^{1-\alpha} = \frac{\alpha}{1-\alpha} \phi_{i,m,t}^{\frac{1}{1-\alpha}} \chi_{m,t}^{\frac{1}{1-\alpha}} w_{m,t}^{\frac{1}{1-\alpha}} \). In more detailed models, profits and size both depend on demand (“taste”) for startups and their productivity, which one could tell apart with more data. Using data on multi-product firms (of any age), Hottman et al. (2016) find that such firm-level demand accounts for at least half of the variation in firm size.
This equation shows that startup employment is the product of total productivity of new entrepreneurs in location $m$ at time $t$, i.e., $\sum_{i|\text{age}_{i,t}=0} \phi_{i,m,t} \xi_{i,m,t}$, as well as market-wide wages and TFP.

This simple model has clear predictions for the direct effects of a positive shock to entrepreneurial productivity, which affects labor demand in area $m$. Consider a positive shock to local entrepreneurial ability: For given wages, local employment by startups (firms of age 0) increases. In equilibrium, wages rise – with the magnitude governed by the local labor supply elasticity. Higher wages lower labor demand by incumbents and startups. Unless wages overshoot, however, startup employment in a given location should still be higher than absent the shock. Moreover, in the startups that always enter, the average size (measured as employment) increases. The overall startup responses depend on firm entry, however.

For startup productivity shocks, we expect entry to affect our results only qualitatively. Consider a simple model of entry. Potential entrepreneurs enter until the present value of profits, possibly net of costs of operating, equals the cost of entry:

$$E_{i,m,t}[\text{NPV}\{\pi_{i,m,s}(\phi_{i,m,t}, \xi_{i,m,t})\}_{s=t}^\infty] = \Phi,$$

where $E_{i,m,t}$ denotes the expectations entrepreneur $i$ forms about the net present value $\text{NPV}$ of future profits, and $\Phi$ is the fixed cost of entry. Recall that $\phi_{i,m,t}$ has a common and an idiosyncratic component and let $\xi$ be the same for all entrepreneurs. Then there exists an entry threshold $\bar{\phi}$, and following a increase of the common component, more entrepreneurs (with a lower idiosyncratic component) enter. For given wages, $\phi_{i,m,t}$ for the marginal entrepreneur is then unchanged following a startup productivity shock. For the marginal entrepreneur, average size is then constant. For the cohort as a whole, average size therefore rises when wages are given. Similarly, when $\xi_{i,m,t}$ falls, average cohort size rises. When wages increase in equilibrium, this can lower the average size of the marginal entrant,

\footnote{Unlike Pugsley et al. (2017), the idiosyncratic component here is a sufficient statistic for startup productivity and, implicitly, productivity dynamics. In Pugsley et al. (2017), productivity has both permanent and transitory components, so that productivity at startup is not sufficient to determine entry.}
but unless labor supply is very inelastic, we expect this effect to be second order.\footnote{Changes in wages affect size more than profits, while changes in $\phi_{i,m,t}$ affect profits and size the same. When wages rise in response to an increased common component of $\phi$, the size of the marginal entrant thus falls with the wage. But the wage effect is likely small, since startups account only for a fraction of overall labor demand. This effect is also dampened by high labor supply elasticities, e.g., when migration is responsive. Note that when the common $\xi$ falls, the size of the marginal startup increases relatively more, since profits are less elastic with respect to changes in $\xi$ than size.}

Different types of startup shocks differ in their predictions. To see this, compare a startup productivity shock with a shock to barriers to entry. While the shocks to startup productivity, $\phi_{i,m,t}$, and financing cost, $\xi_{i,m,t}$, have the same prediction for our observables, shocks to barriers to entry predict lower average size. To see this, let the fixed cost in (5) vary by MSA and time, replacing $\Phi$ by $\Phi_{m,t}$. If the fixed cost falls, less productive entrepreneurs still make profits high enough to warrant entry. As a result, the average productivity of startups falls, lowering the average size – with the magnitude governed by the mass of entrants near the entry threshold. The prediction for average size is thus opposite the prediction following a startup productivity shock.

The model predicts that shocks to local labor supply or economy-wide local productivity shocks have qualitatively similar effects as shocks to entrepreneurial productivity. Going forward, we construct instruments that jointly account for entrepreneurial productivity and overall productivity, and exclude the effects of local labor supply shocks. The remaining challenge is how to disentangle shocks to entrepreneurial productivity from shifts in overall productivity. We do so by assuming that one instrument only loads on overall labor demand in our baseline specification, as discussed in Section 3.

\subsection{Constructing shock proxies in the data}

Our strategy for constructing relevant shock proxies posits that the past industrial structure in a location facilitates how much (new) firms in an area typically benefit from sectoral changes in (startup) productivity. Urban economics suggests that this should, indeed, be the case. For example, areas with a past presence of heavy industry, such as steel in Pittsburgh and autos in Detroit, may make it less likely for an area to benefit from higher startup productivity in other sectors, because the dominance of a few large firms in the heavy industries is likely to drive out potential entrepreneurs (Chinitz, 1961; Chatterji et al., 2013). On the contrary, a presence of firms in innovative sectors may facilitate entry or expansion by startups through input sharing with incumbents (Helsley and Strange, 2002).

Formally, we construct an instrument that exploits the covariance between the past industrial structure with the current growth of startups across sectors. Since Bartik (1991), past industrial
structure has been used as an instrument to isolate sources of exogenous variation in overall local labor demand. We use a Bartik (1991)-type instrument for startup employment $N$ for firms aged zero ($N_0$) in MSA $m$ at time $t$, $Z^{N_0}_{m,t}$, computed as the industry $j$ share weighted average, $\omega$, of national startup (age 0 firms) employment growth at time $t$, $\Delta \ln N^{nat}_{0,j,t}$:

$$Z^{N_0}_{m,t} = \sum_j \omega_{m,j,t} \times \Delta \ln N^{nat}_{0,j,t}. \quad (6)$$

In words, the instrument is constructed by holding fixed an MSA’s industrial composition of employment for a given base year, and computing the startup job growth that would have occurred if the industrial composition in an MSA did not change but startup employment in each industry had only grown at the national rate for startups in that industry.

In the special case of constant industry shares, $\omega$, subtracting year and MSA fixed effects, $\hat{\gamma}_t$ and $\hat{\mu}_m$, as we do in our analysis, highlights the residual variation we use to identify shocks:

$$Z^{N_0}_{m,t} - \hat{\mu}_m - \hat{\gamma}_t = \sum_j (\omega_{m,j} - \bar{\omega}_{m,j}) \times (\Delta \ln N^{nat}_{0,j,t} - \Delta \ln N^{nat}_{0,\bar{j},t}) = \text{Cov}[^{\omega_{0,m}}_m, \Delta \ln N^{nat}_{0,\bar{j},t}]. \quad (7)$$

The residual variation predicts higher startup employment in MSA $m$ at time $t$ when $m$ has historically been stronger in industries that had above average startup employment growth in time $t$.

As this discussion makes clear, this instrument does not reflect all local determinants of startup growth. For example, local policy changes are not picked up, even if they affect some industries more than others (such as easing non-compete clauses). We capture other factors, such as a younger workforce, only if these factors are correlated with industry structure. Moreover, employment at young firms also reflects overall labor demand in the economy because employment of young firms reflects both their productivity and the average labor demand in their labor market. To address this, we introduce a second instrument that controls for overall labor demand.

To identify shocks to overall labor demand, we follow the original Bartik (1991)-type instrument for startup employment, $Z^N_{m,t}$, defined as the industry share weighted average of national startup employment growth of all firms, $\Delta \ln N^{nat}_{all,j,t}$:

$$Z^N_{m,t} = \sum_j \omega_{m,j,t} \times \Delta \ln N^{nat}_{all,j,t}. \quad (8)$$
Note that $Z_{m,t}^N$ in (8) is similar to $Z_{m,t}^{N_0}$ in (6), except that it is for employment at firms of all ages.

As Blanchard and Katz (1992, p. 49) discuss in the context of our instrument for overall labor demand, these instruments are valid for labor demand if national industry-level employment growth, overall or of startups, is uncorrelated with local supply shocks in areas $m$. A special case of an episode captured by this standard Bartik instrument would be the decline of MSAs exposed to industries that suffered from Chinese import competition as in the work of Autor et al. (2013).

![Figure 2: Static reduced form first stage relationship](image)

The figure shows binned scatter plots of the change in the startup job creation rate on the startup Bartik variable $Z_{m,t}^{N_0}$ (panel (a)) and of the employment to population ratio on the overall labor demand Bartik variable $Z_{m,t}^N$ (panel (b)). We winsorize all four variables at the 0.5% and 99.5% levels and use deviations from time and MSA fixed effects. The observations are binned with the radius of the circles indicating the number of observations. The plot compares the (solid) regression line with the (dashed) 45-degree line. In the raw data, Bartik (1991)-type instruments for startup activity and overall labor demand have good predictive power.

The Bartik (1991) logic indeed has predictive power for startup employment and overall employment, as Figure 2 shows. Panel (a) of the figure shows a plot of the change in startup job creation (expressed as a rate in terms of overall employment) in an MSA against the corresponding prediction of the job creation from the startup Bartik instrument, pooling all periods and taking out MSA and year fixed effects. This startup instrument predicts the local change in startup job creation reasonably well with an $F$-statistic of 31.1. Panel (b) of the same figure plots the employment to population ratio against the (standard) overall labor demand Bartik instrument, also net of fixed effects. Here,

\[^{11}\text{Data permitting, we compute } \Delta \ln N_{j,t}^{nat} \text{ excluding the MSA whose outcome we want to predict to further guard against a mechanical relationship between national } \Delta \ln N_{j,t}^{nat} \text{ and local } \Delta \ln N_{m,t}. \text{ Data limitations prevent us from doing this for startup labor demand. But we use data at the sector level that are less geographically concentrated. We also check that the identification is robust to using only MSAs with less than 1% of the population in 1977. The 18 MSAs with at least 1% of the population account for 44% of the population and 58% of employment of all MSAs in 1977.}\]
the predictive power is even higher with an $F$-statistic of 67.8, possibly due to the finer industry
definition. These statistics are, however, only suggestive: They do not account for spatial correlation
or predictable variation in the outcomes. To do this, we now turn to a spatial panel VAR. In this
VAR, the Bartik (1991)-type predictors serve as external instruments to identify orthogonal shocks.

Going forward, we identify shocks by attributing the variation in the well-understood Bartik in-
strument for overall labor demand entirely to the labor demand shock. This makes sure the overall
labor demand shock inherits the standard interpretation in the literature, starting with Bartik (1991)
and with Blanchard and Katz (1992) for VAR models. As a robustness check, we also construct an
analog to the standard Bartik instrument based on TFP growth of public companies. The residual
variation in the startup Bartik identifies the startup shock.

3 Empirical Methodology

We now describe how to identify shocks in a spatial panel VAR with external instruments. Our
approach is to specify a reduced form model for the dynamics and spatial correlation in the variables
of interest and the instruments, and to identify shocks from the correlations between the residuals of the
variables of interest and the instruments. Absent the spatial dimension, our approach would collapse to
the VARs with external instruments introduced by Stock and Watson (2012) and Mertens and Ravn
(2013). We also discuss how to construct a bootstrap algorithm for inference.

3.1 Reduced form VAR

Our VAR is meant to model local labor markets and ultimately identify shocks based on how well na-
tional industry-specific trends predict outcomes in these local markets. To rule out that the estimation
and identification is biased by common aggregate factors or excluded MSA specific characteristics, we
allow for both location fixed effects $\mu_m$ and year fixed effects $\eta_t$. This leads us to specify the following
equation for the $n \times 1$ vector $y_{m,t}$ of observables in location $m$ at time $t$:

$$y_{m,t} = k \sum_{\ell=1}^k A_{\ell} y_{m,t-\ell} + \mu_m + \eta_t + u_{m,t}. \quad (9)$$

Bold lower case letters denote vectors, and bold upper case letters denote matrices. The $A_{\ell}$ matrices
capture within-MSA dynamics and $u_{m,t}$ is a vector of forecast errors.
The spatial econometrics literature often models spatial dependence explicitly. We also do so, using distance weights \( D = [d_{mn}]_{m,n=1}^M \). \( d_{mn} \) measures the geographic proximity between locations \( m \) and \( n \). \( D \) is row-standardized such that non-zero rows sum to one and the matrix has a maximal eigenvalue of unity.\(^{12}\) The resulting model for location \( m \) at time \( t \) links the forecast errors \( u_{m,t} \) to the underlying structural shocks \( \varepsilon_{m,t} \) in its own location and, indirectly, in all interconnected locations:

\[
\begin{align*}
  u_{m,t} &= \sum_{n=1}^{M} d_{mn} R u_{n,t} + B \varepsilon_{m,t}, \\
  \varepsilon_{m,t} &\sim iid \mathcal{N}(0, I_n).
\end{align*}
\]

\( B \) maps the structural shocks to the forecast error, and is only partially identified via instruments and restrictions below. \( R = \text{diag}([\rho_1, \ldots, \rho_n]) \) with \(-1 < \rho_1, \ldots, \rho_n < 1\) parametrizes the degree of spatial dependence. In contrast to the forecast errors, which are typically spatially dependent, the structural shocks \( \varepsilon_{m,t} \) are \( iid \) across time and space. For robustness, however, our inference allows for possible residual temporal and spatial dependence in \( \varepsilon \).

An intuitive expression arises with a simplifying assumption and after stacking variables across MSAs. Given \( E_t \equiv [\varepsilon_{m,t}]_{m=1}^M \), and in the special case of \( R = I_n \rho, |\rho| < 1 \), considered in Mutl (2009), we can plug in recursively to arrive at the following expression:

\[
\begin{align*}
  u_{m,t} &= B E_t e'_m + \rho B E_t D' e'_m + \rho^2 B E_t (D')^2 e'_m + \ldots = B E_t (I - \rho D')^{-1} e'_m,
\end{align*}
\]

where \( e_m \) is an \( M \times 1 \) vector of zeros with a one in its \( m \)th location, i.e., a vector that selects row \( m \). This expression shows how the forecast error in location \( m \) depends on the structural shocks in its own location \( \varepsilon_{m,t} = E_t e'_m \) as well as the (spatially discounted) shocks to all interconnected locations. Clearly, when \( \rho = 0 \), this expression collapses to the standard case in the VAR literature of \( u_{m,t} = B \varepsilon_{m,t} \), e.g., Uhlig (2005).

In the general case, we define \( Y_t \equiv [y_m]_{m=1}^M \) and \( U_t \equiv [u_m]_{m=1}^M \) and show in Appendix C.2 that we can write the VAR as follows:

\[
\begin{align*}
  \text{vec}(Y_t) &= (I_M \otimes A) \text{vec}(X_{t-1}) + \text{vec}(u_t), \\
  \text{vec}(U_t) &= (I_{Nq} - (D \otimes R))^{-1} ((I_M \otimes B) \text{vec}(E_t)),
\end{align*}
\]

\(^{12}\)\( d_{mm} = 0 \) so that a region is not a neighbor to itself. In the special case of discrete distance measures, \( D \) would be the adjacency matrix representing the graph of the network of MSAs. We discuss our specific proximity measures below.
where \( \text{vec}(\mathcal{E}_t) \sim \mathcal{N}(0, I_{MN}) \); \( X_{t-1} \equiv [Y_{t-1}, \ldots, Y_{t-k}] \). \( \otimes \) represents the Kronecker product and \( \text{vec} \) the vectorization operator: \( \text{vec}(A) = [A'_1, A'_2, \ldots]' \). This form of the spatial VAR is tractable as it expresses the forecast error in terms of the \text{iid} standard normal residuals \( \text{vec}(\mathcal{E}_t) \) and therefore allows us to write down the likelihood function or to derive the form of the impulse-responses.

The standard identification challenge in VARs is how to go from the forecast errors to the structural shocks – i.e., how to identify columns of \( B \). Following the recent literature, we use external instruments to identify the shocks of interest, as we now discuss.

### 3.2 Identification

To set up our model for identification via external instruments, we now extend the baseline spatial panel VAR to include the instruments. Let \( z_{m,t} \) denote the vector of instruments and define:

\[
\tilde{u}_{m,t}^\ast \equiv z_{m,t} - (\mu_m^\ast + \eta_t^\ast).
\]

Then we can stack this residual of the instrument, net of year and location fixed effects, with the VAR forecast error \( u_{m,t} \) to get the generalization of equation (10):

\[
\tilde{u}_{m,t} \equiv \begin{bmatrix} u_{m,t} \\ \tilde{u}_{m,t} \end{bmatrix} = \sum_{n=1}^{M} d_{mn} \tilde{R} \tilde{u}_{n,t} + \begin{bmatrix} B & 0 \\ B_{z\varepsilon} & B_{zz} \end{bmatrix} \begin{bmatrix} \varepsilon_{m,t} \\ \varepsilon_{m,t}^z \end{bmatrix}, \quad \varepsilon_{m,t} \sim \mathcal{N}(0, I_{n+n_z}). \tag{14}
\]

\( u_{m,t}^\ast \) is a vector of \( n_z \) noisy measures of the shocks of interest, or a linear combination thereof. \( \varepsilon_{m,t}^z \) is the measurement error. The zero restriction on the upper right corner of \( \tilde{b}B \) simply states that the VAR is not affected by the measurement or prediction error in the instruments. The standard VAR assumption that there are as many shocks as VAR observables implies this restriction.

Next, we build on Stock and Watson (2012) and Mertens and Ravn (2013) to show how having \( u_{m,t}^\ast \) available allows us to identify the first \( n_z \) columns of \( B \) and thus the causal effects of the first \( n_z \) shocks in the VAR, where the ordering is without loss of generality.

For identification, we assume that we can write the matrix \( B_{z\varepsilon} \) as follows, where \( G \) is invertible:

\[
B_{z\varepsilon} = \begin{bmatrix} G, 0_{n_z, n-n_z} \end{bmatrix} B.
\]
The zero restriction in \( [G, 0_{n_z, n-n_z}]B \) corresponds to the exclusion restriction in the instrumental variables literature: The instruments are uncorrelated with the other structural shocks. Invertibility of \( G \) implies instrument relevance. Together these assumptions imply that the \( n_z \) external instruments identify the \( n_z \) structural shocks that are ordered first.

External instruments narrow the identification problem. In our application, having instruments for labor demand and the associated exclusion restrictions requires us to only differentiate between \( n_z = 2 \) labor demand shocks – instead of dealing with all \( n \) shocks, which include labor supply shocks. Mertens and Ravn (2013) show that the instruments reduce the identification problem from imposing the standard \( \frac{n(n-1)}{2} \) identifying restrictions in the VAR to the lower dimensional problem of imposing the \( \frac{n_z(n_z-1)}{2} \) restrictions to factor the \( n_z \times n_z \) covariance matrix of the forecast errors spanned jointly by the two identified shocks. Calling this covariance matrix \( S_1S_1' \), yields the following equation:

\[
\beta_{[1]} = \left[ (I - \eta\kappa)^{-1} \right] \left( S_1S_1' \right)^{1/2}.
\]

(15)

See Drautzburg (forthcoming, Appendix A.4) for closed-form expressions of \( \eta, \kappa, \) and \( S_1S_1' \) in terms of the reduced-form VAR covariance matrix \( V \equiv \tilde{B}\tilde{B}' \).

Our identifying assumption is that the instrument for overall labor demand is independent of shocks to startup productivity. Since we use a standard Bartik (1991)-style to identify overall labor demand shock, this means that we identify this shock as the literature (e.g., Blanchard and Katz, 1992) does. Formally, our identification works by identifying the first shock as proportional to the covariance between the first instrument, here the Bartik (1991) instrument, and the within-MSA forecast errors. We then choose the other shocks to be the orthogonal shocks that explain the residual variation that is jointly explained by all instruments. In our case, there are only \( n_z = 2 \) instruments and this pins down both shocks exactly. Appendix C proves the following formal statement.

Proposition 1 (Identifying shocks.). Let \( V = \tilde{B}\tilde{B}' \) and \( \Gamma = V_{n+1:n+n_z,1:n} \). Partition \( \Gamma = [\Gamma_1', \Gamma_2']' \), where \( \Gamma_1 \) is \( n_z \times n_z \). Assume \( \Gamma_1 \) is invertible, so that \( \kappa = \Gamma_2\Gamma_1^{-1} \) is well defined.
(a) If the first instrument is correlated only with the first shock, we have $\beta_{[1]} = [1, (\Gamma_2 e_1)' \Gamma_1^{-1} e_1]' \times \bar{c} \propto \Gamma e_1$ for $\bar{c} > 0$ (defined in the proof).

(b) We can factor $(S_1 S_1')^{1/2} = [\nu_1, \nu_2]$ in (15), where $\nu_1 = \bar{c}_1(I - \eta \kappa)\Gamma_1 e_1$ and $\nu_2 = F \text{chol}(\Lambda)$. Here, $F$ are the $n_z - 1$ eigenvectors and $\Lambda$ the diagonal matrix of strictly positive eigenvalues of $S_1 S_1' - \nu_1 \nu_1'$. Then the first identified shock is identified only from the first instrument, i.e., $\beta_{[1]} e_1 \propto \Gamma e_1$.

While we focus on shocks identified according to Proposition 2 below, we also report shocks based on the conditional Cholesky factorization. We find that a lower Cholesky factorization of $S_1 S_1'$ yields results similar to attributing the entire variation in the first instrument to the first shock. This indicates that the identification is robust to small violations of the assumption that the standard Bartik instrument is uncorrelated to startup shocks.

3.3 Inferring additional responses

Using data rather than zero-restrictions to identify shocks makes shock identification more uncertain. This is particularly relevant in higher dimensions, because the number of parameters of the covariance matrix that drive the identification grows in the square of the size of the VAR. We therefore find it useful to proceed in a block-recursive manner, following Zha (1999) and Uhlig (2003): (1) We identify shocks in a relatively small-scale VAR that contains $n$ core variables, as specified in (9); and (2) we estimate the dynamics of a second block of $n_p$ variables. These peripheral variables respond to all shocks and variables in the system, but do not influence the core variables themselves.

This approach allows us to infer the response of a number of additional variables without impeding inference. Of course, if the information set of the smaller VAR were insufficient, this approach would be invalid. But we show that our findings in the core-periphery VAR are consistent with estimates of a larger VAR. The larger VAR has qualitatively similar implications but wider confidence intervals for some responses.

Formally, we add two additional sets of equations to the model, analogous to (9) and (10):

$$y_{m,t}^p = \sum_{\ell=0}^{k} A_{\ell}^{p,c} y_{m,t-\ell} + \sum_{\ell=1}^{k} A_{\ell}^{p,p} y_{m,t-\ell} + \mu_m^p + \eta_t^p + u_{m,t}^p$$

(16)
\[ u_{mt}^p = \sum_{n=1}^{M} d_{mn} R^p u_{nt}^p + B^{pp} \varepsilon_{mt}, \quad \varepsilon_{mt} \sim \mathcal{N}(0, I_{n_p}). \]  

(17)

We only use the own lags for the peripheral variables, so that \( A_{t,lp}^{p,p} \) is diagonal. But importantly, we include all the contemporaneous core variables \( y_{m,t-\ell} \) in the peripheral VAR without restrictions on \( A_{t,lp}^{p,c} \). This allows us to infer impact responses of the peripheral variables as \( A_{0}^{p,c} \beta_{[1]} \). More generally, the responses of \( y_{m,t}^p \) at longer horizons depend on the interplay of the core and the peripheral variables. The combined system has a VAR(k) representation with zero restrictions; see Uhlig (2003).\(^{14}\)

We can estimate the error structure, however, without any complications. Below we define \( \tilde{R} = \text{diag}(\rho_1, \ldots, \rho_n, \rho_1^p, \ldots, \rho_n^p) \). Also we define \( A^p = [A_0^{p,c,}, \ldots, A_k^{p,c,}, A_1^{p,p,}, \ldots, A_k^{p,p,}] \).

### 3.4 Estimation

We adapt the two-step procedure in Mutl (2009) for our setup with varying degrees of spatial autocorrelation. First, we difference the data and use an IV procedure with lags of the VAR variables as instruments to identify the dynamic relationship between the variables of interest, i.e., \( A^p \) and \( A^p \).

Second, we compute the concentrated likelihood for given \( D \) and conditional on \( \hat{A}, A^p \) and choose \( \rho_1, \ldots, \rho_n, \rho_{n+1}, \ldots, \rho_{n+n_z} \) to maximize the concentrated (log) likelihood of \( \{u_t, z_t\} \), which is given by:

\[
\log L^c = -\frac{1}{2} \left( MT \log(2\pi) + 2T \log(|I_{M(n+n_p+n_z)} - D \otimes \tilde{R}|) - MT \log(|\hat{V}|) \right) + c, \tag{18}
\]

where \( \hat{V} \) is the sample covariance matrix of \( \{\tilde{\varepsilon}_{m,t}\} \), i.e., of the residuals \( [u'_{mt}, z'_{mt}, u'_{mt}] \) after spatial filtering. \( c \) is a constant independent of parameters. The concentrated likelihood thus depends directly on \( \{\rho_i\} \) through the determinant \( |I_{Mn} - D \otimes \tilde{R}| \) and indirectly through the determinant of \( \hat{V} \).

Intuitively, more spatial autocorrelation tends to lower the remaining variance of the spatially filtered residuals \( \{\tilde{\varepsilon}_{m,t}\} \) and thereby increases the likelihood, but simultaneously lowers the likelihood by lowering the determinant of the Jacobian of the spatial transform, i.e., \( |I_{M(n+n_z+n_p)} - D \otimes \tilde{R}|. \)

\(^{14}\)To trace out the dynamic effects, we can write the VAR without constants and spillovers as \( [y_{m,t}; y_{m,t}^p] = \sum_{\ell=1}^{k} [I; 0; A_0^{p,c,}; I][A_0; 0; A_1^{p,c,}; A_0^{p,p}][y_{m,t-\ell}; y_{m,t-\ell}^p] + [B; 0; A_0^{p,c} B, B^{pp}][\varepsilon_{m,t}]. \)
3.5 Inference

Since the distribution of the implied estimator is non-standard given our two-step estimation procedure, we use a bootstrap procedure.\textsuperscript{15} The main bootstrap procedure exploits that the residuals are spatially uncorrelated once run through an appropriate spatial filter. We therefore draw \textit{iid} with replacement from the spatially filtered residuals vectors, pooled across MSAs and blocks of time of length \(T\).\textsuperscript{16} We also consider a version of our bootstrap that draws blocks of spatial “width” and spatial “width” \(\mathfrak{M}\) and length \(\mathfrak{T}\). See Algorithm 1 for details. In our baseline, we set \(B = 500\), \(\mathfrak{T} = 3\), and \(\mathfrak{M} = 1\). \(\mathfrak{T} = 3\) equals \(T^{1/3}\) up to rounding and we report in the appendix that using \(\mathfrak{M} = 7 \approx M^{1/3}\) does not affect our results, so that we focus on the results from the simpler procedure in the main text.

Blocking of residuals over time accounts for potential residual correlation in the instrument equations and VAR residuals. We find that it matters little in practice. This is intuitive for the VAR, because its lags should capture the autocorrelation. And because we use growth rates to construct the instruments, these should be close to uncorrelated over time absent strong mean reversion.

When applying this procedure we notice small biases in the estimated autocorrelation. We therefore adapt the “bootstrap-after-bootstrap” procedure from Kilian (1998) and run the bootstrap twice. We first estimate the bias in the VAR coefficients and the spatial autocorrelation with half the number of repetitions. We then re-run the algorithm with the bias-corrected coefficients. See Algorithm 2.

In Appendix D we also provide an extension of our VAR model that allows for all model coefficients to differ by (ex ante) MSA characteristics, while taking the spatial correlation among all MSAs into account. We use this extension in several robustness checks.

4 Results

We are now ready to analyze the effect of startup shocks on local labor markets using the spatial panel VAR. The variables in the core VAR capture local labor demand and supply factors. They include the change in the job creation rate by startups, the (log) of the employment-to-population ratio, population growth, and the growth of average wages. We back out the response of the employment level by accumulating the (log) population response and adding this cumulative change to the response of the

\textsuperscript{15}The wild bootstrap in Mertens and Ravn (2013) does not take uncertainty about the covariance between instruments and forecast errors into account. Jentsch and Lunsford (2016) show that, in a pure time-series setup, a moving block bootstrap provides asymptotically valid inference.

\textsuperscript{16}For the terminal period, the terminal block has length \(\mathfrak{T} + \mod (T; \mathfrak{T})\). We sample \([T/\mathfrak{T}]\) blocks with replacement and use only the first \(T\) observations.
Algorithm 1 Bootstrap
For $b = 1, \ldots, B$:

1. Set the initial observations $Y_{i,0}^{(b)}, \ldots, Y_{i,-(k-1)}^{(b)}$ equal to the actual values.

2. Initialize $\mathbb{L} = \emptyset$ as the list of “drawn” MSAs. For $n = 1, \ldots, \lfloor M/2N \rfloor$:
   (a) (Draw target block of MSAs.) Draw $m(b, n)$ from $\{1, \ldots, M\} / \mathbb{L}$. Let $M(b, n)^-$ be the set of
       min\{$\mathfrak{M} - 1, |\mathbb{L}|-1$\} closest remaining neighbors such that $d_{m, n}^- \geq d_{m, n}^- \forall n \in M(b, n)^-, \hat{n} \notin M(b, n)$. Let $M(b, n) \equiv \{m(b, n)\} \cup M(b, n)^-$, ordered as a row vector. Order $M(b, n)^-$ in decreasing proximity from $m(b, n)$.
   (b) For $r = 1, \ldots, \lfloor T/2 \rfloor$:

        • (Draw origin time-space block.) Draw $\hat{m}(b, n, r)$ iid with replacement from $\{1, \ldots, M\}$ and draw $s(b, n, r)$ iid with replacement from $\{1, \ldots, \lfloor T/2 \rfloor\}$. Determine the set of origin MSAs $\mathfrak{M}(b, n, r)^-$ as the set of $|M(b, n)^-|$, closest neighbors of $\hat{m}(b, n, r)$ as in the previous step, where $|M(b, n)^-|$ is the cardinality of $M(b, n)^-$.

        • For $\ell = 1, \ldots, |M(b, n)|$: Set $\{(\hat{\varepsilon}_{e,b,M(b,n)},_{1})_{1=t+1}^{\tau+1} s(b,n,r)\} = \hat{\varepsilon}_{e,b,M(b,n),s(b,n,r)}$.

        • Retain only first $t$ observations to account for the (weakly) longer length of the terminal block.
   (c) Set $\mathbb{L} = \mathbb{L} \cup M(b, n)$.

3. For $t = 1, \ldots, T$ and $n = 1, \ldots, M$:
   (a) Generate spatially correlated error terms from spatially filtered residuals: $\text{vec}(\text{res}_{\ell}^{(b)}) = (I_{Mn} \otimes (D \otimes \hat{R}))^{-1} \text{vec}(\hat{e}^{(b)})$.

   (b) Generate variables of interest:

        • $y_{m,t}^{(b)} = \mu_{m}^{y} + \eta_{t}^{y} + \sum_{\ell=1}^{k} \lambda_{\ell}^{y} y_{m,t-\ell}^{(b)} + \text{res}_{m,t,1:p}^{(b)}$, where $\text{res}_{m,t,1:p}^{(b)}$ denotes elements $1, \ldots, n$ of $\text{res}_{m,t}^{(b)}$.

        • $y_{m,t}^{(b)} = \mu_{m}^{p} + \eta_{t}^{p} + \sum_{\ell=1}^{k} \lambda_{\ell}^{p} y_{m,t-\ell}^{(b)} + \sum_{\ell=1}^{k} \lambda_{\ell}^{p} y_{m,t-\ell}^{(b)} + \text{res}_{m,t,n+n_{p}+1:n+n_{p}+n_{p}}^{(b)}$, where $\text{res}_{m,t,n+n_{p}+1:n+n_{p}+n_{p}}^{(b)}$ denotes the last $n_{p}$ elements of $\text{res}_{m,t}^{(b)}$.

   (c) Generate instruments: $z_{m,t}^{(b)} = \hat{\mu}_{m}^{z} + \eta_{t}^{z} + \text{res}_{m,t,n+1:n+2}^{(b)}$.

4. Re-estimate the VAR: $\hat{A}_{IV}^{(b)}, \hat{A}_{IV}^{\rho^{(b)}}, \{\bar{\gamma}_{i}^{(b)}\}_{i}, \bar{V}^{(b)}$. Compute and save the statistics of interest.

Algorithm 2 Bootstrap with bias-correction

1. Run algorithm 1 with $B/2$ repetitions.

2. For $X \in \{|A_{1}, \ldots, A_{\ell}|, [\rho_{1}, \ldots, \rho_{n+n_{p}}]\}$:
   (a) Compute the average bias: $\Psi_{X} = \hat{E}[\hat{X} - X] = \frac{1}{B/2} \sum_{b=1}^{B/2} \hat{X}^{(b)} - \hat{X}$.

   (b) Compute the bias-corrected coefficient: $\hat{X} = \hat{X} - \Psi_{X}$.

3. Re-run Algorithm 1 with $B$ repetitions and $\{|\hat{A}_{1}, \ldots, \hat{A}_{\ell}|, [\hat{A}_{IV}^{\rho}, \hat{A}_{IV}^{\rho}], \{\hat{\rho}_{1}, \ldots, \hat{\rho}_{n+n_{p}+n_{p}}\}\}$ in step (3).
The VAR specification varies slightly with the available data. We consider 354 MSAs. We allow two lags in our annual VAR. Most samples cover 1986 to 2013, the era for which we have migration data with $k = 2$ lags. Discarding the late 1970s and early 1980s has the advantage of excluding observations of rapid decline in startup activity that may be due to inaccurate measurement in the early part of the sample. However, similar results hold for the longer sample. For constructing the instruments, we use industry weights lagged by $\tau = 5$ years. We measure proximity $D$ using the inverse Euclidean distance, implying that spillovers decay with the physical distance between labor markets. We also consider alternative proximity measures for robustness, such as sharing a state border or having correlated business cycles. However, distance is preferable in our application since distance determines the flow of people which links different labor markets.

We begin by discussing the model specification and then turn to analyzing the impulse-response functions, which are the main focus. We also discuss the variance decomposition and historical counterfactuals in our baseline model. This section concludes by discussing alternative specifications.

### 4.1 Specification tests

**Spatial correlation.** Recall that the spatial correlation coefficient, $\rho$, takes on values between -1 and +1. We find that the spatial autocorrelation is always positive. Table 1 lists the estimated correlation coefficients for each of the variables we study: the point estimates range from 0.08 for the change in the gross job creation rate to 0.76 for net migration rates and 0.79 for house price growth. Our bootstrap finds that these estimates are biased upward. The bias is absolutely larger for the smaller coefficients, but always small. The autocorrelation is significantly positive for all variables at the 95% level, even for the small autocorrelation for the change in the gross job creation rate. Note that our model fits significantly better than more restrictive models with common or no spatial autocorrelation.\footnote{Table F.1 compares our baseline specification with variable-specific spatial auto-correlation to two restricted specifications: (1) a version with common spatial autocorrelation. (2) a version without spatial correlation. The restrictions}
Table 1: Spatial autocorrelation: Coefficients estimates by variable.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Point estimate</th>
<th>Bias-corrected confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Avg. bias 5%</td>
<td>16%</td>
</tr>
<tr>
<td>∆ startup job creation rate</td>
<td>0.08</td>
<td>-0.00</td>
</tr>
<tr>
<td>Employment/Pop ratio</td>
<td>0.41</td>
<td>-0.02</td>
</tr>
<tr>
<td>Pop growth</td>
<td>0.55</td>
<td>-0.03</td>
</tr>
<tr>
<td>Wage growth</td>
<td>0.25</td>
<td>-0.01</td>
</tr>
<tr>
<td>∆ firm entry rate</td>
<td>0.50</td>
<td>-0.02</td>
</tr>
<tr>
<td>House price growth</td>
<td>0.79</td>
<td>-0.03</td>
</tr>
<tr>
<td>∆ young firm exit rate</td>
<td>0.34</td>
<td>-0.02</td>
</tr>
<tr>
<td>Net migration rate</td>
<td>0.76</td>
<td>-0.03</td>
</tr>
<tr>
<td>Firm exit rate (all)</td>
<td>0.54</td>
<td>-0.02</td>
</tr>
<tr>
<td>Startup size</td>
<td>0.23</td>
<td>-0.02</td>
</tr>
<tr>
<td>Bartik: Startup job creation</td>
<td>0.74</td>
<td>-0.03</td>
</tr>
<tr>
<td>Bartik: Employment</td>
<td>0.55</td>
<td>-0.02</td>
</tr>
</tbody>
</table>

The table shows the point estimate and bias-corrected confidence interval for the spatial correlation coefficients $\rho_s$ for all variables $s$. Prior to bias-correction, the estimated spatial correlation shows a small upward bias across all variables. House price data are from CoreLogic Solutions.

Quality of instruments. The first-stage $F$-statistics suggest that our instruments predict the shocks well. After removing the predictable VAR-variation and identifying shocks we find $F$-statistics of 16.2 (Δ job creation rate) and 84.3 (employment/population ratio) in regressions of the identified shocks on the two instruments, see Table 2. The 68% confidence intervals lie at or above 9.5.

Table 2: First-stage $F$-statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Point estimate</th>
<th>Confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Avg. bias 5%</td>
<td>16%</td>
</tr>
<tr>
<td>∆ startup job creation rate</td>
<td>16.2</td>
<td>6.7</td>
</tr>
<tr>
<td>Employment/Pop ratio</td>
<td>84.3</td>
<td>32.5</td>
</tr>
</tbody>
</table>

The table reports the $F$-statistics of regressions of the identified structural shocks on the (spatially filtered) instruments with bootstrapped confidence intervals. The instruments in our baseline specification have $F$-statistics above 10 with a bootstrapped 90% confidence interval of (6.7, 26.0) and (34.3, 104.7), respectively.

4.2 Impulse responses

The figures displaying impulse-response functions, such as Figure 3 for our baseline results, show the point-wise median response as the black line, along with the 68% and 90% point-wise confidence intervals as gray shaded areas. We normalize signs so that the startup shock and the overall labor are rejected at the 1% level based on differences in the log quasi-likelihood and bootstrapped critical values.
demand shock increase the job creation rate and the employment-to-population ratio, respectively.\footnote{We compute IRFs for a counterfactual isolated MSA without spatial spillovers. With spatial autocorrelation in the errors, the spatial spillovers change the IRFs in each MSA by a constant factor of proportionality that depends only on the MSA and the variable of interest, but not on the horizon of the IRF. While these spatial spillovers vary with each variable via $\rho$ and each MSA we found them to be small even for variables with a high degree of spatial correlation. We thus abstract from spatial spillovers here, but results are available upon request.}

**Startup productivity shocks.** This section shows that startup productivity shocks raise employment at startups and, in aggregate, raise employment and population simultaneously. Figure 3 shows these responses on impact and for ten years after the shock. Consider the direct effects first: The upper left panel shows the response of $\Delta$ job creation rate in response to the startup productivity shock, i.e., the change in startup job creation, scaled by overall employment. A one standard deviation shock raises the job creation rate by 0.6 percentage point on impact. Subsequently, the change drops to around -0.2 percentage point, before becoming largely insignificant. This implies a cumulative increase in the job creation rate of around 0.4 percentage point relative to the time before the shock. Given the average employment to population ratio of 37.5\% in our sample,\footnote{The BEA’s employment-to-population ratio is around 60\%. Our denominator is larger, as it also includes institutionalized population and those under age 16, while our numerator excludes government employment.} the initial 0.6 percentage point increase translates to an increase of 0.2 percentage point relative to population, or two thirds of the 0.3 percentage point difference in average startup job creation between the 10th percentile of MSAs and the median MSA. The cumulative increase still accounts for half of this difference.

Turning to population growth, we find an initial increase in population growth of 0.4\%, which dies out gradually and adds up to an increase of about 1\%. As we argue in Appendix B.1, domestic migration seems to account for the population growth. The (overall) employment level only increases significantly with a delay, but eventually stabilizes at a level about 1\% higher, driven by the higher population level. The estimated confidence intervals are reasonably precise and, except for the impact response of the employment level, all three responses are significant at the (two-sided) 10\% level.

One narrative consistent with our estimates is that startups invigorate the local economy. Initially, jobs are largely reallocated from incumbents to startups, as in the estimates of Hombert et al. (2014) for the French economy. The persistent entry of startups makes the local economy more attractive, however, and migrants start arriving. As population increases, overall employment rises.

Turning to prices, the startup productivity shock does not lead to significant growth in the wage rate, but increases house prices. House prices grow significantly for the first three years after the shock, before house price growth reverts to zero. Average wage growth is not significantly different from zero.
The figure shows the estimated impulse-responses to a one standard deviation startup productivity shock, along with bootstrapped confidence intervals. The solid line is the median across the bootstrapped draws, while the shaded areas are the 68% and 90% confidence intervals. $\Delta$ job creation rate refers to the job creation rate by startups. The initial increase in startup job creation is only partly undone in subsequent years. Initially, overall employment does not increase. With a delay, the shock leads to population growth and higher overall employment. The underlying house price data are from CoreLogic Solutions.

Figure 3: Impulse-responses to startup productivity shocks: Baseline VAR, 1986 to 2013.

Both the house price growth and the flat wage rate are consistent with competition from migration keeping wages down, while driving up house prices. Our findings on house prices, population, and employment are in line with the effects of a state-wide entry rate increase reported in Gourio et al. (2016). They also find persistent increases in population and employment, and higher house prices.

Startup productivity shocks also raise firm entry and exit, as Figure 3 also shows. Following the startup productivity shock, firm entry rises. The change in the firm entry rate increases by 0.2 percentage point on impact, with virtually no offsetting effects thereafter. Exit of young firms also rises, but by less than 0.04 percent point.

The response of the average size of startups corroborates our interpretation of this shock as a shock to startup productivity. In our simple model, a startup productivity shock is an increase in the entrepreneur’s span of control and thus increases the average startup size for a given wage rate. Given that wage growth is not significantly different from zero, the model would predict average size to rise, if the shock is indeed one to startup productivity. In our impulse-responses, we find that average size increases significantly following a startup productivity shock: On impact, this increase in size is just under 13.2 percentage point and persists over eight years.
The intensive margin – the average size of entrants – also accounts for most of the job creation. The estimated 0.6 percentage point increase in the startup job creation rate corresponds to 21.2% increase in startup employment on average. Our estimates for the 0.2 percentage point increase in the entry rate imply that 2.1 percentage points of the 21.2% increase in startup employment are due to the entry margin. In 1986, for example, the corresponding numbers for the median MSA are 15.5% due to changes in average size and 1.8% due to the entry margin. Alternatively, when we estimate the importance of the intensive margin directly, we find an initial increase in average size of 13.2% with a 68% confidence interval of 8.8% to 15.7%. While the estimates of the exact contribution of the intensive margin do not line up perfectly, overall they clearly imply that the intensive margin accounts for most of the increased job creation, in line with Sedlacek and Sterk (2017).

Overall labor demand shocks. Since our identification of startup productivity shocks relied on “stripping out” the effect of overall labor demand shocks, we now discuss these shock responses. We first argue that the estimated responses are plausible and our identifying assumptions thus appropriate. Second, we contrast responses to the overall labor demand shock to responses to the startup productivity shock to shed light on the underlying economic forces.

The overall labor demand shock leads to a large transient increase in employment and the average wage rate, with only small effects on population growth. Figure 4 shows these responses to the overall labor demand shock. A one standard deviation shock to overall labor demand increases the (overall) employment level by about 2.2%. The increase in employment is persistent with a half-life of about 6 years. Eventually, however, employment reverts to zero. Population growth is positive for up to two years after the shock and contributes initially about 0.2 percentage point to the employment increase, but then becomes mildly negative. The average wage rate increases by about 1.3% on impact, and edges up further in subsequent years. The simultaneous increase in wages and employment would be consistent with a persistent increase in overall TFP, and seems plausible as a response to the overall labor demand shock we set out to identify.

The qualitative response of wages, population, and house prices is also consistent with research on place-based policies. House prices increase only initially, then start falling two years after the initial shock to overall labor demand. While this could reflect either construction booms or endogenous changes in the form of more relaxed land regulation, we show below that only the pattern, and not the relative magnitudes of housing price growth, is robust to different specifications of the VAR. The
A robust pattern of initial appreciation and subsequent reversal is intuitive for a temporary shock. The pattern of wage increases with small population responses and no lasting house price increases also mirrors the finding in Busso et al. (2013) for place-based policies.

The figure shows the estimated impulse-responses to a shock to overall labor demand, along with bootstrapped confidence intervals. The solid line is the median across the bootstrapped draws, while the shaded areas are the 68% and 90% confidence intervals. ∆ job creation rate refers to the job creation rate by startups. Lower barriers to entry do not increase employment of age 0 startups significantly, but operate through persistently higher entry. Similar to startup productivity shocks, this results in higher population growth and employment. The underlying house price data are from CoreLogic Solutions.

Figure 4: Impulse-responses to overall labor demand shocks: Baseline VAR, 1986 to 2013.

The marked differences of the responses of employment, population growth, and wages between overall labor demand shocks and startup productivity shocks may be due to the higher persistence of startup shocks. Startup shocks cause the job creation rate by startups to increase by about 0.4 percentage point and lead to an initially small but lasting increase in overall employment. The lasting employment increase following startup shocks is driven by population growth, while it is largely due to an increase in the employment to population ratio following a labor demand shock. Last, the average wage rate is flat following a startup productivity shock, but rises following an overall labor demand shock. One explanation for these differences is that migration undoes the wage pressure that results as labor demand shifts up and moves along an increasing labor supply schedule of the current population. If migration is less likely in response to short-lived shocks, for example due to moving costs, migration will adjust to wage differences only in response to persistent shocks. Thus, the high persistence of the effects of startup productivity shocks might explain why only after startup shocks the average wage...
rate remains flat and population shows no sign of mean reversion.

Overall labor demand shocks lead to small and transient increases in startup activity. Job creation by startups rises initially by 0.1% of overall employment, before reverting to zero. Similarly, firm entry rises by 0.1% of the stock of firms, before turning slightly negative for the next ten years. As a consequence, the average size of startups is largely flat, although the 68% confidence interval around the estimated 2% initial increase excludes zero.

**An alternative startup shock: Barriers to entry shock.** We now examine whether our conclusions about the effect of startup shocks depend on the nature of the startup shock. Our startup productivity shock operates primarily through the intensive margin, i.e., the size of entrants. We now turn to an alternative startup shock, which affects startups primarily by lowering barriers to entry. To that end, we construct a Bartik instrument that uses the cross-industry variation in startup entry in terms of the number of firms, rather than the employment at young firms.\(^{20}\)

\[^{20}\text{We use this instrument in a VAR which includes firm entry and exit in the core VAR, rather than in the periphery.}\]
characteristics. Figure 5 shows the corresponding impulse-responses. A one standard deviation shock that lowers barriers to entry raises the firm entry rate on impact by 0.4 percentage point – twice as much as in response to a startup productivity shock. The increase in firm entry is lasting, since the initial increase in the number of new firms per period is not undone by subsequent reversals. In contrast to a startup productivity shock, the barriers to entry shock leaves average startup size approximately unchanged: The point estimates are zero or mildly negative, with wide confidence intervals. Indeed, job creation by startups during their first year of existence does not rise significantly. The exit rate of young firms drops initially, indicating that the barriers to entry shock could be associated with lower fixed costs of operating a business instead. However, the initial decrease in exit is subsequently reversed, consistent with less careful selection of potential entrepreneurs into entrepreneurship.

While the implications of the barriers to entry shock for startup characteristics differ from those of a startup productivity shock, the economy-wide implications are similar. As Figure 5 also shows, population growth rises persistently, reaching a cumulative population increase close to one percent after about six years. The overall employment level rises, largely driven by the higher population. Last, average wage growth is insignificantly different from zero, while house prices rise significantly. While house price growth is slightly higher following the barriers to entry shock, the estimated effect of either startup shock is very similar. Following a one standard deviation shock, population growth increases initially by around 0.4 percentage point, employment is not significantly different from zero on impact, but then rises persistently by about 1%.

4.3 Variance decomposition

How much of startup activity do startup shocks explain? And how large are the MSA-wide effects of startup shocks? To answer these questions, we now turn to a forecast error variance decomposition. Table 3 reports the results for variance decomposition for a typical MSA over ten years.21

Startup shocks explain about half of the variation in the job creation rate by startups. Specifically, Table 3 shows, in the first line and column (1), that startup shocks account for 50.8% of the variation in the startup job creation rate. The 68% confidence interval ranges from 25.8% to 74.8%. The IRFs implied that startup productivity shocks operate largely through the average size margin and

---

21The results at different horizons are similar and available upon request. Since we found spatial spillovers to be small in the IRF analysis, we abstract from spatial spillovers here. Our results can thus be interpreted as characterizing an MSA without neighbors.
moved entry little. This is also true relative to the overall forecast error variance: The identified shock explains only 6.8% of the variation in the entry rate, with most of the variation in entry attributable to non-identified shocks or idiosyncratic entry components – also when including entry in the core VAR (see Table F.5(b)). However, the importance of the intensive margin in the IRFs also holds up in relative terms: 39.5% of the overall forecast error variance in average size is explained by the startup productivity shock, and 53.9% (=39.5%/73.3%) of the variation is explained by the core VAR.

Our finding that startup productivity shocks are important at the local level lines up with structural estimates of aggregate models. Sedlacek and Sterk (2017) attribute about 70% of the variance in startup employment to startups shocks other than entry cost and more than 80% of the variance of average startup size. Drautzburg (2019) attributes about three quarters the variation in (detrended) startup employment to shocks to the productivity distribution of new entrepreneurs, and about 30% of the movements in average size. Our 50.8% estimate for the contribution to employment using local data is broadly comparable, but slightly smaller, and the 68% confidence interval accommodates the structural estimates. Similarly, our estimates imply that startup productivity shocks play an important role for understanding average size, in between the estimates in aforementioned papers. At the same time, our estimates leave room for sizable effects of other shocks. These other shocks can include, for example, responses to demographic changes, as in Karahan et al. (2015) and Engbom (2019).

Consistent with the estimated IRFs, startup shocks are an important driver of population growth and migration. The startup shock accounts for 39.4% of the forecast error variance in population growth. Together, the core VAR explains only 44.0% (100%-56.0%) of the variation in migration rates, with much of the remainder likely measurement error. But startup shocks account for 15.8% of the total forecast error variance in migration rates and 35.9% (=15.8%/44.0%) of the forecast error accounted for by the core VAR, similar to the explained variance of population growth. Because in this model population growth is the only source of long-run employment growth, this implies that both startup shocks and the non-identified shocks are important sources of employment fluctuations.

Overall labor demand shocks contribute little to the variation in startup activity, again similar to our analysis of IRFs. However, overall labor demand shocks explain most (76.4%) of the variation in the employment-to-population ratio. The overall labor demand shock accounts for only around 2% of the variation in startup job creation and entry rates.

Half of the variation in labor demand by startups is driven by the identified labor demand shock:
Table 3: Variance decomposition in our baseline VAR. 68% confidence interval.

<table>
<thead>
<tr>
<th>10 year horizon</th>
<th>(1) Startup shock</th>
<th>(2) Overall labor demand</th>
<th>(3) Other VAR shocks</th>
<th>(4) Idiosyncratic shock</th>
</tr>
</thead>
<tbody>
<tr>
<td>∆ startup job creation rate</td>
<td>50.8 (25.8, 74.8)</td>
<td>1.9 (0.4, 3.5)</td>
<td>47.3 (23.9, 72.2)</td>
<td>0.0 (0.0, 0.0)</td>
</tr>
<tr>
<td>Employment/Pop ratio</td>
<td>2.9 (1.1, 4.9)</td>
<td>76.4 (70.6, 82.3)</td>
<td>20.8 (14.8, 26.5)</td>
<td>0.0 (0.0, 0.0)</td>
</tr>
<tr>
<td>Pop growth</td>
<td>39.4 (18.4, 61.8)</td>
<td>11.4 (7.6, 15.2)</td>
<td>49.2 (26.3, 70.1)</td>
<td>0.0 (0.0, 0.0)</td>
</tr>
<tr>
<td>Wage growth</td>
<td>6.6 (0.6, 12.7)</td>
<td>20.9 (15.8, 26.3)</td>
<td>72.4 (63.8, 81.3)</td>
<td>0.0 (0.0, 0.0)</td>
</tr>
<tr>
<td>∆ firm entry rate</td>
<td>6.8 (4.5, 9.0)</td>
<td>2.4 (1.7, 3.1)</td>
<td>3.3 (1.5, 5.1)</td>
<td>87.4 (86.0, 88.8)</td>
</tr>
<tr>
<td>House price growth</td>
<td>1.1 (0.4, 1.8)</td>
<td>11.0 (9.9, 12.2)</td>
<td>3.9 (2.8, 5.0)</td>
<td>83.9 (83.0, 84.8)</td>
</tr>
<tr>
<td>∆ young firm exit rate</td>
<td>1.1 (0.7, 1.6)</td>
<td>1.7 (1.4, 1.9)</td>
<td>1.3 (0.9, 1.8)</td>
<td>95.9 (95.4, 96.3)</td>
</tr>
<tr>
<td>Net migration rate</td>
<td>15.8 (7.1, 24.1)</td>
<td>7.3 (5.2, 9.4)</td>
<td>20.8 (11.0, 30.3)</td>
<td>56.0 (51.3, 60.6)</td>
</tr>
<tr>
<td>Firm exit rate (all)</td>
<td>1.4 (0.4, 2.4)</td>
<td>3.5 (2.9, 4.0)</td>
<td>3.2 (2.0, 4.4)</td>
<td>91.9 (90.6, 93.2)</td>
</tr>
<tr>
<td>Startup size</td>
<td>39.5 (21.1, 57.0)</td>
<td>1.3 (0.2, 2.4)</td>
<td>32.5 (15.1, 50.7)</td>
<td>26.7 (25.3, 28.0)</td>
</tr>
</tbody>
</table>

The table shows the fraction of the forecast error variance accounted for by the shocks in the VAR and, for the variables in the periphery, by the “idiosyncratic shocks” that affect only the variable of interest. Startup shocks explain about half of the variation in the job creation rate by startups and about 40% of population growth. Overall labor demand shocks contribute little to the variation in the job creation rate by startups, but explain most of the variation in the employment-to-population ratio. Except for migration, the VAR shocks explain relatively little of the peripheral variables.

But Table F.5 shows that the results are similar when we include more variables in the core VAR. House price data are from CoreLogic Solutions.

see row 1 and column (1) in Table 3. The point estimate indicates that startup shocks account for 50.8% of the variation in the startup job creation rate, but the confidence interval is wide. In addition, non-identified shocks can account for as much as 72.2% of the variation in the startup job creation rate, according to the upper bound of the 68% confidence interval shown in column (3) of the table.

Turning to overall labor demand shocks shown in column (2), the point estimate indicates that these shocks account for 76.4% of the variation in the employment-to-population ratio. Moreover, the overall labor demand shock shown is tightly estimated, as other shocks account only for 14.8% to 26.5% of the variation in the employment-to-population ratio.

Our results do not imply that other shocks, which may include labor supply, are not important in local economies. Subtracting the confidence interval for the contribution of the unidentified shocks from 100 shows that the two identified shocks together account for 27.8% to 76.1% of the variation in population growth and less than 36.2% of the variation in wage growth. This is consistent with a potentially important role for labor supply shocks, which Karahan et al. (2015) emphasize.

4.4 Counterfactual histories

We now examine the historical implications of our VAR as a way to assess the plausibility of its results. To do so, we back out the historical shocks for each MSA. We analyze the panel of historical shocks first, before zooming in on Philadelphia, an MSA for which we have local knowledge.
The figure shows binned scatter plots and regression lines relating growth in the number of venture capital funded firms to the median estimated shock realization in the MSAs. Weights are Stata’s \texttt{rreg} weights, i.e., combining Cook’s distance with downweighting of observations with large residuals. \( p \)-values based on the \( t \)-distribution with 12 degrees of freedom and standard errors clustered by MSA and year.

Figure 6: Relationship between time in VC growth and startup shock from 2000 to 2013.

To assess the historical shock series across MSAs, we compare them to a proxy for the financing costs of startups. The rationale for this comparison is that in our simple model, and with the available data, we cannot differentiate between a shock that lowers startup-specific input costs and a shock that raises new entrepreneur’s span of control: Only the ratio of startup productivity \( \phi_{i,m,t} \) and the financing cost of inputs \( \xi_{i,m,t} \) affect labor demand. While both affect profits, and thus entry decisions, to different degrees, both predict responses of the same signs. The comparison with the financing cost proxy is, thus, a useful test of our identification of startup productivity shocks in the motivating model.

Despite the noisy data on the financing cost proxy, our startup shock correlates significantly with this variable. Figure 6 shows a binned scatter plot of the identified startup shock against VC growth. It averages observations across MSAs and time, thereby reducing noise, and displays a statistically significant and positive relationship between VC growth and the startup shock. The \( p \)-value is 0.06, based on a \( t \)-distribution with 12 degrees of freedom and standard errors clustered by MSA and year. Using weights that weigh down outliers, computed using \texttt{rreg} in Stata, the \( p \)-value drops to less than 0.01. Various regression specifications confirm this finding (see Table F.2 in the Appendix).

Since few firms receive VC funding, it may be surprising that our shock correlates with VC growth, if the growth in VC funded firms reflects credit supply. However, if VC funds the largest startups, this may free up funding from other lenders for smaller startups, thus also pulling up smaller startups,
which are more likely to be captured in our data. Moreover, our measure of startup shocks is based on startup employment and thus tilted towards larger firms.

Our proxy for the financing costs of startups is the growth in the number of VC funded firms. Since VC focuses on young firms it is reasonable to associate this measure with costs of financing startups. Our preferred interpretation is that the cross-sectional variation in VC capital reflects credit supply, and thus financing costs. The data on the growth in the number of VC funded firms come from the National Venture Capital Association. We have up to 13 observations for the 250 MSAs that we matched by name. The average MSA has seven observations. The venture capital data are noisy, with annual growth rates of ± 100% in the sample.

Figure 7: Historical startup productivity shocks, shock correlations, and counterfactuals: Philadelphia, 1986–2013.

Panel (a) shows the median identified startup shock for the Philadelphia MSA with 68% confidence intervals (gray area) and the identified labor demand shock with 68% confidence intervals (red dashed lines). Panel (b) relates a 2-year moving average of the startup shocks to the number of gene therapy startups whenever a non-zero number is reported. Panels (c) and (d) report the data on startup job creation and population growth and the counterfactual with zero startup shocks (gray area) or zero labor demand shocks (red dashed lines), along with the 68% confidence interval.

Figure 7 shows the corresponding startup shock time series for Philadelphia, which also aligns with the isomorphism between startup productivity and financing costs and the lack of price data, i.e., data on \( \xi \) directly, the increase in the number of VC funded startups could also reflect increased demand for funds following a startup productivity shock that shifts the distributions of startups to the right.
broadly with our narrative. Evidently, as Panel (a) suggests, Philadelphia experienced a large, positive startup shock in 1986 and smaller shocks thereafter – with an average standard deviation of only 0.4, as opposed to 0.9 for all MSAs. Reassuringly, the shock exhibits no significant autocorrelation, with a $t$-statistic of -0.9. The large spike in 1986 coincides with the Reagan era tax reform that could have affected startup activity. Our model suggests that the 1986 startup shock was more important in Philadelphia than in the average MSA. Afterward, startup shocks were on average negative with the exception of two spells in the early 1990s. Since 2008, startup shocks have been positive. After accounting for the noisy data on the VC financing, we also find the positive relationship between the cost of financing startups and our startup shock in the data for Philadelphia.\textsuperscript{23}

Our motivating model emphasizes shocks to entrepreneurial productivity, and our empirical strategy proxies for these using past industrial structure. We now provide an example of how this strategy applies to Philadelphia. Philadelphia has a strong presence in medicine, with several teaching hospitals and a presence of large pharmaceutical firms. Recently, the \textit{Philadelphia Inquirer} ran a piece on how startups in the Philadelphia area have been working on the commercialization of gene therapy. The first NIH-approved and successful application in humans took place in 1989. The \textit{Inquirer} lists 13 startups founded in our sample period from 1988 to 2012, with several more in more recent years. Since we cannot tell with precision when the startups listed by the \textit{Inquirer} showed up in our sample, which is based on mid-March payroll information, we compute 2-year averages of our startup shock and the number of gene therapy startups. The correlation of these measures in the 13 years with non-zero gene therapy startup formation is 0.72 with a heteroskedasticity robust $t$-statistic of 4.3.\textsuperscript{24}

Last, the historical analysis allows us to quantify specific episodes during which startup shocks were important for population growth in Philadelphia. Specifically, we compute counterfactual job creation – Panel (c) in Figure 7 – and population growth – Panel (d) – in Philadelphia without startup shocks. Without the large 1986 shock, Philadelphia would have seen job creation by startups and population growth almost one percentage point lower than in the data. For much of the 1990s and early 2000s, in contrast, population growth and job creation was held back by the largely negative job creation shocks. For example, in 2002 population growth would have increased 0.32 percentage points instead of the

\textsuperscript{23}Using the same procedure to compute outlier-robust regression weights as Figure 6, but applying it to the data for Philadelphia only, reveals a strong positive relationship between the annual fluctuations in VC funded firms in Philadelphia, and the startup shock since 2001, when the data begin: The correlation is 0.45 with a heteroskedasticity robust $t$-statistic of 2.9.

\textsuperscript{24}A Tobit regression between the identified startup shocks and reported gene therapy startup formation yields a positive coefficient and a heteroskedasticity-robust $t$-stat of 1.6.
observed 0.12 percentage point if it were not for startup shocks. Since 2008, job creation shocks have pulled up startup job creation and population growth, although the effect is significant only in 2012 and 2013. In 2013, population growth would have been increased by 0.10 percentage point instead of 0.29 percentage point absent the positive startup shocks.

4.5 Alternative specifications

In this subsection, we summarize variations of our baseline specification, detailed in online Appendix F.2.

Identifying assumptions. In the baseline VAR, we assume that the standard Bartik instrument loads only on the overall labor demand shock. In Figure F.3 we report the results when we relax this assumption using a conditional Cholesky factorization. The results are similar to baseline VAR.

Instrument construction. First, we also construct the Bartik-style instruments using constant 1974 weights instead of using time-varying industry weights. The exclusion restriction may be more readily satisfied with time-invariant weights computed further in the past; see Figure F.4. Second, we address concerns that other shocks in the MSA may be reflected in the instrument since we cannot drop each MSA’s own startup employment when constructing the startup-Bartik instruments. To address this concern, we split the estimation into the 18 MSAs with more than 1% of the population in 1977 and the other 336 MSAs, while accounting for spatial correlation between all MSAs; see Figure F.5. Third, we swap the employment-based Bartik instrument for the overall labor demand shock for an instrument that uses TFP growth of incumbents. This addresses concerns that the exclusion restriction for the overall labor demand shock could be violated because overall MSA employment includes that by startups; see also Figure F.5. Neither variant changes our baseline results qualitatively.

Bootstrap algorithm. We show that our block bootstrap produces results, shown in Figure F.6, that are very similar to an $iid$ bootstrap and to spatial blocking.

Distance measure. In Figure F.7, we show robustness with respect to three different proximity measures: (1) We consider a binary proximity measure that classifies MSAs as neighbors when they share a common state. (2) In the spirit of Crone (2005), we consider the correlation of business cycles, measured as the correlation of HP-filtered (log) employment, as a proximity measure. (3) We use
a weighted average of distance matrices \( D(\lambda) = D_1 \lambda + D_2 (1 - \lambda), \) where \( D_1 \) is the inverse distance proximity matrix, \( D_2 \) is the one based on common states, and \( \lambda \) is also estimated.

**Lag length.** Our results are comparable when we extend lag length to three years. Both the quality of our first stage and the main results for the IRFs are largely unchanged. A VAR with a single lag appears misspecified due to non-stationarity; see Figure F.8.

**Information set.** Moving variables from the periphery to the core VAR has few effects on the impulse-responses. In particular, we include entry rates and house price growth in the main VAR. This could be important, because Mehrotra and Sergeyev (2017) find that house price growth matters for new firms as a proxy for credit risk. However, the larger VAR leaves the responses of the core variables largely unchanged, although the first stage \( F \)-stat drops; see Figure F.9.

**Sample period.** When we begin the estimation in 1980, we find qualitatively similar results. However, the effects of the startup shock on employment and population growth are noisier. Since startup activity was high but quickly declining at the beginning of the sample, the early sample period may reflect a different type of dynamic than the remainder of the sample; see Figure F.10.

**Metro characteristics** Allowing for the VAR to differ by ex ante MSA characteristics, we find that our results are driven by the bulk of MSAs, which exclude the densest MSAs (Figure F.11), the MSAs with the highest startup rates (Figure F.12), and the most regulated MSAs (Figure F.13). But we cannot say whether the results differ because we have little power in the sample of MSAs with higher density, more regulation, or higher startup rates, or whether the MSA characteristics matter.

5 **Conclusion**

In our analysis of local U.S. economies, startup shocks are sizable, account for a substantial part of the forecast error variance, and have lasting effects. On impact, a positive one standard deviation shock closes raises job creation, relative to population, by about two thirds of the difference between the 10th percentile of MSA and the median MSA. With a delay, this leads to a lasting increase in population and employment: While larger or new startups initially are largely associated with a reallocation of
jobs from incumbent businesses, their arrival is also associated with a persistent increase in the arrival of migrants, and population growth. This population growth leads to higher employment.

Our results are comparable to those of structural models of entrepreneurship estimated for the U.S. as a whole. Similar to Sedlacek and Sterk (2017) and Drautzburg (2019), we find that shocks that move the average size of startups are important for understanding job creation by startups, even though we use local data. While our results also leave room for other drivers of startup activity, they should encourage researchers to develop models with a prominent role for direct startup shocks, which particularly drive the intensive margin.

Local policymakers who seek policies to promote entrepreneurship may also be encouraged by our findings. However, the startup shock we focus on is a shock to startup productivity. Our paper is silent on which policies would increase startup productivity. Our results do show, however, that lower barriers to entry encourage more startups to enter and have similar economy-wide effects than shocks to startup productivity. This suggests that policies designed to lower barriers to entry may be attractive to local policymakers. However, our results are based on a reallocation of people between MSAs. If many policymakers engaged in similar policies, one would expect the effects to be weaker.
References


Chatterji, Aaron, Edward Glaeser, and William Kerr, Clusters of Entrepreneurship and Innovation, University of Chicago Press, April 2013


Appendix

A Data

County-industry employment imputation. We build on Autor et al. (2013) for our imputation. Their code uses the county of establishments within industry-size brackets as well as employment totals at higher levels to impute county-level employment at the four digit SIC and six digit NAICS level. Intuitively, the algorithm computes a year-industry specific mapping of the binned size distribution of establishments to total year-industry employment. The algorithm runs repeatedly until estimates on the pooled disclosed and imputed data converge. We then aggregate the county-industry-level employment to the metropolitan level. Following Autor et al. (2013), we begin with OLS imputation, imposing lower and upper bounds for average employment by size bracket after the estimation. This procedure always converges for the decadal data analyzed in Autor et al. (2013), but not in all years. When the OLS analysis with ex-post bounds does not converge, we switch to non-linear least squares that imposes the bounds during the estimation. After imputing employment according to the prevailing classification scheme in each year, we use cross-walks from the 1977 SIC classification to the 1987 SIC classification and from future NAICS classifications to the 1997 six-digit NAICS classification, which we then, in turn, transform to the 1987 four-digit SIC classification and aggregate up to the three-digit level.\footnote{For the NAICS to SIC crosswalk, we use the crosswalk from Autor et al. (2013). We could not find a comprehensive crosswalk for the minor within-SIC and within-NAICS changes. To that end, we first use correspondence tables to identify the mapping between sectors. For some industries, this identifies the mapping uniquely, i.e., 100\% of one or more industries map into a single industry. If one industry maps into more than one industry, we compute the weights in the crosswalk by regression: We regress the share in the originating industry in the last year of the old classification on the shares of the receiving industries in the new classification at the county-industry level. In our baseline specification, we use OLS and set negative coefficients to zero before normalizing weights to add up to unity. A non-linear LS procedure that respects these constraints yields similar results, but can become unwieldy in the rare cases when a large number of industries are the receiving industries, e.g., in the case for some wholesale sectors.} We also use these data to compute sectoral weights to predict startup activity.

Main variable definitions.

• Net migration rate: We define the net migration rate as the difference between inflows and outflows of IRS exemptions, divided by the population level in the prior period. The number of exemptions on tax returns is, typically, the number of household members.

\[
\text{net migration rate}_{m,t} = \frac{\text{No. exemptions (inflow)}_{m,t} - \text{No. exemptions (outflow)}_{m,t}}{\text{Population}_{t-1}} \tag{A.1}
\]

• Job creation rate: We define the job creation rate as the change in job creation by firms aged 0, divided by the average of overall private employment in the current and prior year.

\[
\Delta \text{job creation rate}_{m,t} = \frac{\Delta \text{job creation by firms aged 0 in MSA}_{m,t}}{\frac{1}{2}(\text{MSA employment}_{m,t} - 1 + \text{MSA employment}_{m,t})} \tag{A.2}
\]

The numerator follows Haltiwanger et al. (2013).

• \(\Delta\) Firm entry rate: We define the firm entry rate as the change in the number of firms aged 0, divided by the average of the number of firms of any age in the current and prior year.

\[
\Delta \text{firm entry rate}_{m,t} = \frac{\Delta \text{Firms aged 0 in MSA}_{m,t}}{\frac{1}{2}(\text{All firms}_{m,t} - 1 + \text{All firms}_{m,t})} \tag{A.3}
\]
• Δ Firm exit rate: We define the firm exit rate as the change in the number of firms aged 1 that exit, divided by the average of the number of firms of any age in the current and prior year.

\[
\Delta \text{firm exit rate}_{m,t} = \frac{\text{Firm deaths of firms aged 1 in MSA}_{m,t}}{\frac{1}{2}(\text{All firms}_{m,t-1} + \text{All firms}_{m,t})} \tag{A.4}
\]

• Δ Overall firm exit rate: We define the overall firm exit rate as the change in the number of firms of any age that exit, divided by the average of the number of firms of any age in the current and prior year.

\[
\text{overall firm exit rate}_{m,t} = \frac{\Delta \text{Firm deaths of any firm in MSA}_{m,t}}{\frac{1}{2}(\text{All firms}_{m,t-1} + \text{All firms}_{m,t})} \tag{A.5}
\]

• Employment-to-population ratio: We use overall employment from the County Business Patterns to compute the log growth rate. This measure agrees closely with BDS employment; see Figure B.1. It enters the analysis in logs.

• Population growth: We compute the log growth rate. The log growth rate has the advantage of being additive to compute level changes, from which we can back out the change in the employment level.

• Growth of average wages: We compute the log growth rate of the average wage rate in the County Business Patterns.

• House price growth: We compute the log growth rate of first quarter house prices.

• TFP growth: See Appendix E.

Definition of instruments.

• Overall labor demand shock proxy:

\[
Z^{\text{overall}}_{m,t} = \sum_{i} \omega^{\text{SIC3}}_{m,i,t-5} \Delta(\log(emp_{i,t} - emp_{m,i,t}) \tag{A.6}
\]

• Startup productivity shock proxy:

\[
Z^{\text{startup}}_{m,t} = \sum_{i} \omega^{\text{sector}}_{m,i,t-5} \Delta \text{job creation rate}_{i,t} \tag{A.7}
\]

• Barriers to entry shock proxy:

\[
Z^{\text{barriers}}_{m,t} = \sum_{i} \omega^{\text{sector}}_{m,i,t-5} \Delta \text{firm entry rate}_{i,t} \tag{A.8}
\]

• Overall TFP-based labor demand shock proxy, where the SIC classification follows Table E.1:

\[
Z^{\text{overall}}_{m,t} = \sum_{i} \omega^{\text{SIC}}_{m,i,t-5} \Delta \log(TFP_{i,t}) \tag{A.9}
\]