The Impact of Venture Capital on Economic Growth*

Steven Poelhekke[1] and Benjamin Wache[2]

[1] *Vrije Universiteit Amsterdam, Tinbergen Institute, CEPR, CESifo*

[2] *CPB Netherlands Bureau for Economic Policy Analysis, Vrije Universiteit Amsterdam, Tinbergen Institute, IfW Kiel*

October 8, 2023

Abstract

Does venture capital (VC) investment yield economic growth? A large literature studies the effect of VC investments on firm-level activity, but its effects on economic growth are less well understood. We identify the effect of VC investment flows on destination county employment, wages, and establishment creation, using a novel instrument that captures the 'social connectedness' of counties to non-local major sources of VC investment. Using detailed data on investor-to-company VC flows, we find a large positive impact of VC investment, suggesting that strong social connections to large venture capital hubs are an important contributor to regional economic growth.

JEL Codes: R11, G24, G41

Keywords: Growth, venture capital, social connectedness

*We thank Eric Bartelsman, Kirill Borusyak, Don Davis, Jos van Ommeren, and David Garces Urzainqui for helpful comments and discussions.

Benjamin Wache (corresponding author): CPB Netherlands Bureau for Economic Policy Analysis, Bezuidenhoutseweg 30, 2594 AV Den Haag, Netherlands. Email: b.wache@cpb.nl.

Steven Poelhekke: Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, Netherlands. Email: steven.poelhekke@vu.nl.

1 Introduction

Venture capital (VC) investors are different from banks and other capital providers because they collect detailed knowledge about the firms they invest in and monitor them intensively. As such, "these inside investors can provide financing to young businesses that otherwise would not receive external funds" (Lerner, 1995, p.301). While it is clear that this has positive effects on the receiving firms (Kortum and Lerner, 2000; Da Rin et al., 2013; Bernstein et al., 2016; Chen and Ewens, 2021), relatively less is known about whether venture capital investment translates into measurable effects on local economic growth beyond the receiving firm. One of the first to address this question, using data that covers the years 1993 to 2002, are Samila and Sorenson (2011) who find positive effects of the number of investments made by VC firms on pay and firm births (but not on employment) in the same county as that of the investing VC firm. Motivated by the growth of the VC industry since 2002—to $45 billion invested in 2019 and $1,116 billion assets under management in 2022 (NVCA, 2023)[1]—and by recent evidence showing that VC investment tends to raise innovation in technologically related other firms (Schnitzer and Watzinger, 2022), we revisit the question whether VC investment contributes to overall (local) economic growth. Moreover, we measure VC investments in the *destination* county to address the fact that these investments are nowadays much more spread across space than the locations of investing VC firms and typical tech hubs.

We collect detailed investor- and company-level data on the value of VC investments made between 1986 and 2019 and combine this with panel data from the County Business Patterns (US Census) on destination county-level employment, total payrolls, and total number of establishments. In addition, we propose a novel instrument to address endogeneity concerns such as reverse causality and simultaneity. To instrument the amount of venture capital invested in a destination county, we construct a measure of social proximity to major sources of venture capital investment outside of the destination county's own metropolitan area. In particular, we construct the weighted sum of venture capital investments originating in all other source counties (excluding as sources the destination's metropolitan area), where

[1] See Crisanti et al. (2021) for a description of recent developments in the EU.

the weights are the county's strength of social ties toward those other counties. We call this index *social access to venture capital*.[2] Intuitively, this variable captures the availability of VC throughout the economy, weighted by how much social access a county has to the rest of the economy. The idea is to capture the likelihood that an entrepreneur in a given location is matched with a potential venture capital investor to support the entrepreneur's business, which arguably depends on social connections to major sources, and more so if the entrepreneur is located outside the major source counties themselves. Our proxy for these specific connections is the aggregate bilateral county-level degree of connectedness, as captured by the social connectedness index (SCI, see Bailey et al. 2018)—a variable derived from Facebook friendship data. The share of monthly active users represents 75% of the U.S. population in 2019, and correspond to real-world networks of relatives, colleagues, business partners, and friends. Where these networks are denser they increase county-to-county bank lending (Rehbein and Rother, 2022) and predict international and sub-national trade beyond the effect of distance and borders (Bailey et al., 2021). Moreover, Kuchler et al. (2022) show that institutional investors are more likely to invest in firms from regions to which they have stronger social ties, suggesting that "Facebook friendship links represents a broader type of network that is intrinsically different from networks based on factors such as shared neighborhood or education institutions".

By excluding VC flows where the destination county's own metropolitan area is a source of venture capital we avoid capturing simultaneous effects such as when a high concentration of VC firms in cities like San Francisco invest locally, because such cities may also be particularly active in patenting, start-up activity, host major universities, and have a high concentration of skilled labor (Bloom et al., 2021). In addition, we include a rich set of fixed and time-varying destination county effects to capture the presence of these factors. To make sure that our index does not affect economic growth directly (conditional on fixed effects), we also exclude from the instrument source-county VC investments into all counties of the destination county's own metropolitan area (including the destination county itself). To

[2] Samila and Sorenson (2011) instrument the number of first-round investments made by VC firms in a metropolitan area using the distance-weighted returns of institutional Limited Partners located in all other MSAs, who invest in VC funds and regularly re-balance their portfolios. While this effectively predicts the supply of VC investment in an origin location, it is not designed for our setting, which is the need to instrument the supply of VC investment in a destination location.

control for any remaining indirect geographic spillovers across larger spatial units, such as via trade links, we additionally control for a measure of 'social access' to county GDP.

Our empirical results indicate a large and positive impact of venture capital investments on local economies. In our preferred estimates that instrument venture capital investment with social access to venture capital sources, we find that a doubling of VC investment leads to a 7.5% increase in the number of establishments, and a 7.0% increase in the number of jobs and the amount of total payroll. In other words, each million dollar invested yields at least one establishment, 40 jobs, and $7 million in payroll wages (suggesting relatively high-paying jobs). We thus find that, while the number of investments had positive but modest effects on origin counties up to 2002 (Samila and Sorenson, 2011), that the intensive margin (the total amount invested into a county) has become much more important as a driver of economic growth, in destination counties during the 2000s and 2010s.

Several tests boost our confidence in these results: we include a broad set of fixed effects, including county and state-by-year effects, and alternatively state-by-year (or metropolitan area-by-year) and county-by-five year period fixed effects. In addition, we vary the definition of social access to venture capital in ways that include or exclude more counties in the vicinity of the destination county to rule out simultaneity and indirect effects, we exclude the main source states of California, Massachusetts, and New York, cluster at different levels, and we vary controls for market access. Our results are highly robust to these different specifications.

Our paper relates to a literature that studies drivers of regional clusters of economic growth, entrepreneurship, and innovation (Glaeser and Kerr, 2009; Glaeser et al., 2010a,b; Chatterji et al., 2014; Glaeser et al., 2015; Carlino and Kerr, 2015; Guzman et al., 2022), and a literature on the role of public entrepreneurial finance (Lerner, 2009; Bai et al., 2021) that finds that public involvement in VC financing is prevalent and discusses the circumstances under which it is efficient. We add to this literature by highlighting venture capital as a driving force of growth and by emphasizing the importance of social connections. Encouraging the entry of VC firms in local markets, as well as fostering social relations across space (e.g. via sector-specific trade fairs, or exchange programs), may be promising policies for regions looking to increase business dynamism.

We also contribute to the fast growing literature on the role of social networks in finance, with a recent collection of studies leveraging Facebook's social connectedness index. For example, Kuchler et al. (2022); Rehbein and Rother (2022); Wache (2023) have utilized this variable to look at the geographic spread of investments by institutional investors, cross-county lending, and spatial frictions in VC investment, respectively. Our study is the first to use the SCI to construct an instrument for local VC investment.

The rest of the paper is organized as follows. Section 2 explains the data and shows descriptive statistics. Section 3 explains the empirical strategy, and Section 4 presents and discusses the results. Section 5 concludes.

2 Data

The main dependent variables come from the County Business Patterns (CBP), published annually by the United States Census Bureau. In particular, we use annual observations of annual payroll, total employment, and the number of establishments, for all US counties between 1986 and 2019. *Employment* is the total of full- and part-time employees on the establishment's payroll; *Payroll* includes all forms of compensation paid during the reporting year in thousands of dollars; *Establishments* is the count of individual physical locations at which business is conducted. The final sample contains a balanced sample of 3,132 counties between 1986 and 2019.

The main explanatory variable is the total amount of venture capital (VC) equity investment into all companies in a given US county in a given year. This variable is constructed from data on all recorded individual venture capital investments at the firm-level in the ThomsonOne database, which is one of the standard sources on VC investments used in the literature (Da Rin et al., 2013). We observe 240,978 investment flows by 6,010 investors into 41,487 companies within the lower 48 states. Using this investment level information, we geocode the reported addresses of the investor (a venture capital firm) and the receiving company. We then aggregate the dollar amounts invested into companies located in each destination county in each year.[3] Let the total amount of venture capital equity investment

[3]We observe only one address of the receiving company: since they tend to be small and young firms, we assume that firms do not have other establishments outside of their headquarters' county.

(denoted in current millions of US dollars if not mentioned otherwise) into destination county i in year t be $VC_{i,t}$. Furthermore, let $VC_{ij,t}$ denote the bilateral VC investment flow into destination county i from origin county j in year t. We thus have that

$$VC_{i,t} = \sum_{j \in J} VC_{ij,t}$$

The Social Connectedness Index (SCI), a measure of social connections across US counties, was developed in collaboration with Facebook and published by Bailey et al. (2018). The index is based on Facebook friendships and measures the number of Facebook friendships between two counties divided by the product of the Facebook populations in the two counties. However, data on the county-to-county SCI is only available for a snapshot in August 2020. Previous studies (Kuchler et al., 2022; Wache, 2023) have shown that the influence of the SCI on several other variables is very stable over time, indicating that the SCI is a slow-moving object and a reasonable proxy for earlier years. Denote the social connectedness index between county i and county j as SCI_{ij}. The index is defined as:

$$SCI_{ij} = \frac{\text{Facebook Connections}_{i,j}}{\text{Facebook Users}_i * \text{Facebook Users}_j} \tag{1}$$

The measure of the Social Connectedness Index that is publicly reported is then scaled between a maximum of 1,000,000,000 and a minimum of 1. A doubling of the SCI means that the chance that two random people are friends (one from each given location) is roughly twice as large.

According to Bailey et al. (2018), more than 58% of US adults and 71% of the US online population used Facebook as of September 2014. Although Facebook usage is relatively similar across income, education, and racial groups, it declines with age. The authors also report that Facebook is primarily used as a platform for real-world friends and acquaintances to interact online, and that people typically only add connections they know in real life. Based on these findings, the authors argue that Facebook data provides a unique representation of US friendship networks on a large scale.

Data on GDP at the county level (available from 2001) is taken from the Bureau of

Economic Analysis. Information on the geography of the United States, such as Metropolitan Statistical Areas (MSA), Combined Statistical Areas (CSA), and state membership at the county-level are taken from the website of the Census Bureau.[4] Descriptive statistics of all variables are provided in Online Appendix OA1.

In Figure OA1 we show the evolution of total VC investment and its share of GDP over the period 1986-2019. Clearly visible is the peak and bust of the dot-com bubble of 1998-2002: at its peak in 2000 VC investment reached close to 0.7% of GDP. By 2019 VC investment has recovered with ups and downs to about $45 billion and 0.2% of GDP.

Within our sample, VC investments were made by 6,010 individual investors: the top three source states are California (1,610 investors), New York State (1,115 investors), and Massachusetts (457 investors), reflecting that the VC industry is more clustered in space than the finance industry and entrepreneurship itself (Chen et al., 2010; Kraemer-Eis et al., 2018). However, our detailed data allows us to split origins and destinations: of the 240,978 observed investment flows, only 104,751 were done within the same state (39% of total value invested), and only 28,146 within the same county (10% of total value invested). In Figure 1 we show the geographical distribution of inward VC investment by *destination* county, averaged over 2016-2019. This shows that inward investment is especially large in the Seattle, San Francisco, Los Angeles, Manhattan, and Boston area. However, many other counties spread across the country receive VC investment of 63.2 million dollars on average among non-zero county-year observations. This geographic pattern contrasts the more geographically concentrated origin counties and motivates us to look at the effects of VC investment on destination counties.[5]

[4]The US Census Bureau defines an MSA as "areas [that] consist of the county or counties (or equivalent entities) associated with at least one urbanized area of at least 50,000 population, plus adjacent counties having a high degree of social and economic integration with the core as measured through commuting ties". A CSA is defined as a group of adjacent metropolitan and micropolitan areas that has a high degree of interchange. See https://www.census.gov/programs-surveys/metro-micro/about/glossary.html. As of March 2020, there were 384 Metropolitan and 543 Micropolitan Statistical Areas, which together form 172 Combined Statistical Areas.

[5]In Figure OA2 we show the total amount of outward VC investment by *origin* county.

3 Empirical Strategy

We gauge the performance of US counties by focusing on three measures of economic activity at the county-year level: total employment, total annual payroll, and number of establishments, and relate these to the total dollar amount of inward VC investment made in the previous year. We estimate the basic regression specification given in Equation (2):

$$Y_{i,t+1} = \beta VC_{i,t} + X_{i,t} + \mu_i + \nu_{r,t} + \varepsilon_{i,t}, \tag{2}$$

where $Y_{i,t+1}$ is the 1-year lead of annual data on annual payroll, employment, and number of establishments at the county level in the US. $VC_{i,t}$ is the total amount of VC equity funding invested into county i in year t. $X_{i,t}$ are control variables, μ_i are county fixed effects, and the $\nu_{r,t}$ are combinations of region and region times year fixed effects, where regions can be counties, states, or Combined Statistical Areas. In some specifications, $Y_{i,t+1}$ and/or $VC_{i,t}$ are transformed via the log or inverse hyperbolic sine (asinh) transformation. We always cluster standard errors at the county level (the level of treatment), which is robust to heteroskedasticity and autocorrelation.[6]

Despite the timing in Equation (2), estimating with OLS raises endogeneity concerns. For example, counties with high economic growth may be more attractive to invest in, which results in reverse causality, or local economic growth and VC investment may follow common trends. Our first step in addressing these concerns is to include a rich set of fixed effects. We report results for regressions including (i) county fixed effects, (ii) state-by-year fixed effects, and (iii) CSA-by-year fixed effects. Because our unit of observation is a county-year we cannot include county-year fixed effects that would absorb common trends such as the local business cycle stance. To come close to these we (iv) include county-by-five-year-period fixed effects.

Second, we propose an instrument for $VC_{i,t}$. In particular, we use *social access to venture capital* (SAVC) as an exogenous measure of supply of venture capital to a destination county. We first define the instrument in detail before discussing the identifying assumptions. The idea is to capture the likelihood that an entrepreneur in a given location is matched with

[6]Results are robust to clustering at the state, or CSA, or county & year (two-way) level instead.

a potential venture capital investor (from another location) to support the entrepreneur's business. Unlike banks, venture capital firms do not have branches where they might service 'walk-in' customers. Entrepreneurs and VC investors thus have to rely on their networks to find each other.[7] In this spirit, we approximate the potential success rate of this unobserved matching process by a county's vicinity to venture capital, as captured by the strength of social connections of a county to other counties, where we measure the importance of VC source counties by the total amount of venture capital that is invested by VC investors from that county in a given year. SAVC in a given county is defined as the sum of the total amount of venture capital invested by VC firms located in all other counties, weighted by the bilateral social connectedness index between the home county and all respective other counties. In order to exclude a potential mechanical influence of VC investments on a county's local economy, we exclude all VC investments made by any firm toward companies located in the *same MSA* as county i. The construction of the SAVC measure relies on investments by VC firms based in all origin counties, but is based only on observed investment flows toward other counties *not* located in the destination county's MSA. For example, the SAVC of New York County is the SCI-weighted sum over all counties, of the total outgoing VC investment from these sending counties to all other counties minus their investments into counties that belong to the same MSA as New York County (i.e. the New York-Newark-Jersey City MSA). Alternative definitions of this variable are discussed in Online Appendix OA4.

More formally, we define social access to venture capital as:

$$SAVC_{i,t} = \sum_{j \in J} SCI_{ij} \cdot \sum_{k \notin M_i} VC_{jk,t}$$

where $\sum_{k \notin M_i} VC_{jk,t}$ is the sum of flows from county j to all counties k that are not in M_i, where M_i is the set of counties in the same MSA as county i. SCI_{ij} is the SCI between counties i and j.

[7] See Gompers et al. (2020) for a taxonomy of the VC investment process, including a description of how venture capitalists 'meet' entrepreneurs to potentially invest in. Burg et al. (2022) review the mechanisms that entrepreneurs use to acquire funding and other resources via their social networks. See Ewens et al. (2022) for a recent example of a theoretical treatment of the search-and-matching process between venture capitalist and entrepreneur. In addition, a related literature studies financial flows across space more generally (Head and Ries, 2008; Okawa and van Wincoop, 2012; Pellegrino et al., 2023).

This instrumental variable is motivated by Wache (2023), who shows that in a gravity equation of county-to-county venture capital investments, the social connectedness index is the most robust predictor of investment flows, whereas physical distance, trade, migration, and commuting have much less predictive power after the inclusion of the social connections. We take this result one step further and use the strength of social connections to construct a measure of differential exposure to venture capital over time and space. The definition of SAVC is similar in spirit to the concept of *market potential* (Harris, 1954), defined as the sum over other markets' GDP divided by physical bilateral distances, as well as the more recent concept of *market access*, which takes into account different price levels across locations (e.g. Redding, 2022).

The key identifying exclusion restriction assumption in this setting states that access to venture capital affects local economic outcomes only through its effect on local VC investment. We argue for this restriction by: a) excluding all investments into a given county's own MSA when constructing access to venture capital; b) including fixed effects (county, state-year, county-five-year-period); and c) controlling for the effect of general improvements in market access. The first step prevents VC inflows from affecting outcomes indirectly by stimulating economic growth in the MSA of a county. For example, when constructing the SAVC of New York County, we exclude all investments from all locations toward companies based in the New York-Newark-Jersey City MSA. This is to avoid both a direct and an indirect relation between the measured VC investments with local county and local MSA economic conditions.

Figure 2 illustrates the variation in SCI in August 2020 for one of the most important source counties by VC investments made between 2015-2019: San Mateo County in California (south of and adjacent to San Francisco). San Mateo County is home to Menlo Park, one of the most prominent clusters of VC firms in the US. A log difference of 1 roughly corresponds to a doubling of the likelihood that two random people in the given counties are friends. Social connections are overall declining in physical distance with most connections to counties nearby, but not in a linear fashion: beyond Colorado connections roughly follow population density.

The most important source of VC investment is San Francisco (see also Figure OA2).

Other important sources are Chicago, Boston, New York and Los Angeles, each with total investments of a billion dollars or more in 2015, but there are many other source counties across the country with total investments in the range of 10-100 million dollars.

Our instrument thus aims to capture that an entrepreneur located in Reno, in Washoe County (Nevada), cannot get VC funding from local VC firms, because there are no VC firms active there (as shown by the yellow region in Figure OA2), but will benefit from the strength of social connections between Reno and counties in California, which increases the probability of being matched with VC firms in, for example, San Mateo county. Moreover, we exploit the spatial and temporal variation in the annual change in the amount of outward investment from origin counties (shown in Figures OA3 and OA4, for selected years): if investors from a particular origin county invest more in a particular year, then entrepreneurs in destination counties with good social connections to this origin county likely benefit more from the uptick in investment.

A second threat to identification is that venture capital investment in a location outside of the own MSA influences economic conditions in that other location, which in turn influences economic conditions in the location that is being considered through economic linkages (as discussed by Betz et al., 2018). To control for this possibility, we propose to use $SAGDP$ as an additional control, to capture variation in the general degree to which locations are socially connected to economic activity, and therefore benefit from high degrees of social connectedness and/or social proximity to economic activity. We define $SAGDP_{i,t}$ as:

$$SAGDP_{i,t} = \sum_{j \neq i} SCI_{ij} \cdot GDP_{j,t}$$

where $GDP_{j,t}$ is the real GDP of county j at time t. We create a set of dummies for each percentile of its distribution and include them in equation 2.

4 Impact of VC Investment on Local Economic Outcomes

How large is the impact of VC investments on local economic outcomes? We start by estimating Equation (2) by straightforward OLS, before moving to IV estimates.

4.1 OLS Results

For each of the outcomes employment, payroll, and establishments, we show in Table 1 three regressions where we vary the fixed effects in three ways. Column 1 shows, conditional on county and state-times-year fixed effects (that control for the relative location of a county and its state's business cycle stance, respectively), that receiving twice as much inward VC investment significantly increases employment by 1.1% in the following year. It also raises total payroll wages by 1.9% and the number of establishments by 1%. For comparison, these effects are two orders of magnitude larger than previously reported by Samila and Sorenson (2011) who use similar sources but aggregate up to the level of MSAs for the years 1993 to 2002 and measure outcomes in the origin county of VC investors. However, the venture capital industry has since grown in size and geographic spread. According to our data, annual VC investment in the US grew by 315% between 2002 and 2019, up from 0.13% of GDP in 2002 to 0.21% of GDP in 2019.[8] Because the local business cycle may deviate from that of the state and may drive both VC investments and local economic activity we replace in columns 2, 5, and 8 the county fixed effects by county $\times$ five-year-period fixed effects. In columns 3, 6, and 9 we instead replace the state $\times$ year fixed effects with CSA $\times$ year fixed effects. We find in all cases a significant positive effect of VC investment on the economic performance of counties, although the impact is now more attenuated by a factor two to five.

4.2 IV Results

Our rich set of fixed effects can nevertheless not absorb unobserved factors at the finer county-year level that may simultaneously attract VC investment and determine local economic

[8]Total inward VC investment reached \$45.7 billion in 2019, up from \$14.5 billion in 2002 (and only \$2.6 billion in 1986), but still below the dot-com peak of \$67.2 billion in 2000.

performance. To address this issue we now turn to instrumental variable regressions. Table 2 shows the results of instrumenting inward VC investments by SAVC, the social access of a county to the venture capital that is invested into other counties, which varies over time as VC firms from various source-counties investment more over time, and over space according to where they choose to invest and the strength of social ties between counties. Note that in predicting the total amount of inward VC investment in a given county, we explicitly exclude all inward investments made by all VC firms into the given county's MSA. In addition, we control for market access by including a set of fixed effects for each percentile of the distribution of SAGDP, the social access weighted GDP of other counties. This reduces the sample because GDP at the county level is only available from 2001.

Panel A shows the second stage results and Panel B the corresponding first stages. As before, we vary the set of fixed effects in three ways, for each of the outcomes employment, payroll, and establishments. Starting with Panel B, we find that SAVC significantly predicts inward VC investments: a one standard deviation increase in SAVC (one unit) increases investment by 6 to 8%, and cluster-robust Kleibergen-Paap first-stage Wald-statistics are comfortably above 10 in all specifications and samples. Panel A shows that the IV estimates are large: doubling VC investment increases employment —directly and indirectly —by more than 8%, payroll by more than 10% and the number of establishments by more than 5%.

To achieve growth in VC investment in a local county, policies may contribute to strengthen the social ties with important VC source counties such that SAVC increases. This relates to more traditional policies that aim to improve the access to markets such as by constructing infrastructure (e.g. Donaldson and Hornbeck, 2016; Hornbeck and Rotemberg, 2021). While we are agnostic about the specific policies that successfully create business-related social ties, one may think for example of trade shows. We therefore report in Table OA2 the reduced form analogue of Table 2. The results show that a one standard deviation increase in SAVC —equivalent to turning Boulder County in Colorado into San Francisco's Marin County (both in 2017) —stimulates employment, total payroll and the number of establishments by about 1% each.

The IV estimates are substantially larger than the OLS estimates: one possible explanation is measurement error, which may have attenuated the OLS estimates toward zero if one or more of our variables of interest are measured imprecisely. However, the outcome variables are based on censuses and should be accurate.

Kaplan and Lerner (2017) and Maats et al. (2011) discuss ThomsonOne as a data source on VC investments. They conclude respectively that 10 to 20% of financing rounds are missing with a deteriorating quality around the years of the global financial crisis, but that fund coverage is better in ThomsonOne than in the alternative database by VentureSource.[9] Da Rin et al. (2013, p577) conclude that "financing amounts are measured with a fair amount of error in both databases, but the amounts tend to be unbiased on average", which suggests that measurement error is classical and can be address by IV.

A second possibility is that the estimated local-average treatment effect is larger than the average treatment effect, for example when some counties attract relatively little VC investments, despite their good social access to important venture capital source counties (the 'non-compliers', see Angrist and Pischke, 2009). The estimated effect is then more heavily influenced by a subset of counties that do attract a lot of VC investment (the 'compliers'). These counties may also be more conducive to other forms of investment, leading to faster improvement in economic activity.

Third, it is possible that our estimates are influenced by the fact that the inverse hyperbolic sine transformation of our main variable of interest is not scale invariant (Bellemare and Wichman, 2020; Aihounton and Henningsen, 2021; Norton, 2022; Chen and Roth, 2022). The transformation is nevertheless attractive, because it avoids dropping zeros when taking logs and avoids adding an arbitrary number to VC investments before taking logs (such as typically adding 1). In our main specifications the scale of VC investments is in dollar millions, because it is also reported in millions by ThomsonOne.[10] In Online Appendix OA8 we report IV regressions where we instead denote VC investments in terms of unit dollars or thousands of dollars before applying the inverse hyperbolic sine transformation: estimates are of the same order of magnitude but become smaller as we reduce the unit of VC invest-

[9] Now part of CB-Insights, see: `https://www.cbinsights.com/research/team-blog/dow-jones-venturesource-valuations/`.

[10] Alternative scalings such as unit dollars or thousands would thus suggest a level of precision that is not there in the raw data.

ments. A doubling of VC investment still stimulates employment, payroll and establishments by 4% or more.

As an alternative, we present IV estimation results of *untransformed* variables in Table 3, which departs from the log-linearity assumption and has the appealing feature of showing the impact of each million dollars invested on employment in persons, dollars paid in wages, and establishments formed. However, a caveat is that in these regressions the first stages perform less well in predicting VC investment: while SAVC still significantly predicts VC investment, Kleibergen-Paap first-stage Wald statistics are now much smaller and below the rule-of-thumb of 10. We address this issue by following the state of the art recommendations by Andrews et al. (2019) to compute the Montiel Olea and Pflueger (2013) effective F-statistic (which corrects for non-homoskedastic errors and is reported in Panel B) and compute weak identification-robust and efficient Anderson-Rubin (AR) confidence intervals (reported in Panel A).[11] While the AR intervals could not find an upper bound (due to relatively weakness of the instrument in this untransformed setting), they do provide an identification-robust *lower* bound to the estimated effects of VC investment on county outcomes. Each million dollar invested yields at least one establishment, 40 jobs, and $7 million in payroll wages (suggesting relatively high-paying jobs).

4.3 Robustness

The results presented in Section 4.1 and Section 4.2 are robust to a set of alternative specifications and variable definitions. We summarize the results here, and present regression tables and details in the Online Appendix. In all robustness tests we present variations of the IV specification of Table 2, columns 3, 6, 9, that include CSA×year fixed effects.

Alternative definitions of SAVC. In the main results we define SAVC as the social capital (SCI) weighted distance to total outward VC investments originating in other counties but excluding flows to the own MSA. We vary the definition of SAVC in three ways: (i) the

[11] In columns 2, 5, and 8 the algorithm to compute the intervals did not converge due to large number of fixed effects. These columns therefor also report Kleibergen-Paap Wald statistics rather than Montiel Olea and Pflueger (2013) effective F-statistics. However, under the assumption of homoskedastic errors the two statistics are the same in the just-identified case.

SCI-weighted sum over all counties, of the outgoing VC investment to all counties except the given county, (ii) the SCI-weighted sum over all counties, of the outgoing VC investment to all counties except flows toward counties in the same CSA as the given county, and (iii) the SCI-weighted sum over all counties except source counties that belong to the same MSA as the given county, of the outgoing VC investment to all counties. Formulas are provided in the Online Appendix. Table OA4 presents the regressions where we vary SAVC. The results are highly robust to this exercise, where our most stringent definition of SAVC (the second version) has also the strongest predictive power for VC investment. The third definition of SAVC suggests substantially larger effects, but this may be due to the fact that this definition does not exclude the own MSA as destination when aggregating VC investment from other origin counties. As discussed, this may lead to indirect effects, when neighboring counties receive substantial VC investment and thereby stimulate economic activity indirectly, potentially violating the exclusion restriction. Comparing the baseline results with definitions one and two, we find that it matters relatively less whether we exclude flows from other counties to the home county (definition one), to the home MSA (baseline), or to the home CSA (definition two) when defining SAVC.

Clustering. In our baseline results we cluster standard errors by county, from the logic that this is also the level of treatment. In Online Appendix OA5 we cluster alternatively by state, by CSA, and two-way by county and year. Standard errors are somewhat larger when clustering by state, but all results (including first-stage F-stats) remain significant throughout these variations in clustering.

Alternative controls for market access. To control for market access we constructed dummies for each percentile of the distribution of SCI-weighted GDP. In Online Appendix OA6 we instead create 50 or 500 bins of the distribution, or exclude them altogether. Results are again highly robust, with the note that excluding the controls for market access (which we so far always included) yields coefficients that are about 50% larger in magnitude.

Excluding High VC Locations. Because the spatial distribution of inward VC investment is skewed toward the states of California, Massachusetts, and New York (see Figure

1), we test whether their counties are driving the results. In Online Appendix OA7 we restrict the sample, each time excluding one or more of these states. Compared to the baseline results, excluding California results in larger coefficients, although they are virtually unchanged when also excluding the other two states.

Scaling. Finally, we present in Online Appendix OA8 results where we change the scale of the investment variable from millions (which is the unit in the raw data and which we use in the baseline) to unit dollars or thousands. In the baseline we transform dollar amounts with the inverse hyperbolic sine to preserve the zeros in the sample. This is however not scale-invariant, resembling the arbitrary choice of adding a 1 or another number before taking logs. For completeness, we repeat Tables 1 and 2 with alternative scaling. The results show that coefficient magnitudes are sensitive to scaling, although not by orders of magnitude, and all results remain significant. In some specifications, however, the first-stage F-statistics becomes smaller. We attribute this to the fact that rescaling to smaller units magnifies the skewness in the distribution such that it becomes harder for the instrument to predict VC investment. Our preferred baseline estimates therefor use the unit that corresponds to the unit observed in the raw data.

5 Conclusion

This paper investigates the county-wide growth effects of venture capital investment in terms of employment, wages and establishment creation. Unlike the existing literature, it focuses on outcomes in destination counties and leverages a plausibly exogenous source of variation, namely the degree to which the populations of these counties are socially connected to sources of venture capital. The data show that the volume of VC investments has grown since the early 2000s and that VC investment is prevalent in an increasing number of counties across the United States.

Counties on the receiving end of investment show a positive response: they increase employment, payroll wages and the number of establishments. These findings are robust even after accounting for potential endogeneity and reverse causality through our instrumental variable approach.

We conclude that the effects of VC investment are positive and substantial: during the 2000s and 2010s, each million dollar invested yielded at least one establishment, 40 jobs, and $7 million in payroll wages. These magnitudes are much larger than the effects documented for earlier time periods and for source counties, suggesting that the VC industry has become more important for regional economic development over time.

Policymakers aiming to foster entrepreneurship and regional development should take into account the potential impact of VC investments, and the social connections that help to match entrepreneurs to investors.

References

Aihounton, Ghislain B D and Arne Henningsen, "Units of measurement and the inverse hyperbolic sine transformation," *Econometrics Journal*, June 2021, *24* (2), 334–351.

Andrews, Isaiah, James H. Stock, and Liyang Sun, "Weak Instruments in Instrumental Variables Regression: Theory and Practice," *Annual Review of Economics*, August 2019, *11* (1), 727–753.

Angrist, Joshua D. and Jörn-Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, 2009.

Bai, Jessica, Shai Bernstein, Abhishek Dev, and Josh Lerner, "The Dance Between Government and Private Investors: Public Entrepreneurial Finance around the Globe," *SSRN Electronic Journal*, 2021.

Bailey, Michael, Abhinav Gupta, Sebastian Hillenbrand, Theresa Kuchler, Robert Richmond, and Johannes Stroebel, "International trade and social connectedness," *Journal of International Economics*, 2021, *129*, 103418.

_ , Rachel Cao, Theresa Kuchler, Johannes Stroebel, and Arlene Wong, "Social Connectedness: Measurement, Determinants, and Effects," *Journal of Economic Perspectives*, August 2018, *32* (3), 259–280.

Bellemare, Marc F. and Casey J. Wichman, "Elasticities and the Inverse Hyperbolic Sine Transformation," *Oxford Bulletin of Economics and Statistics*, February 2020, *82* (1), 50–61.

Bernstein, Shai, Xavier Giroud, and Richard R. Townsend, "The Impact of Venture Capital Monitoring: The Impact of Venture Capital Monitoring," *Journal of Finance*, August 2016, *71* (4), 1591–1622.

Betz, Timm, Scott J. Cook, and Florian M. Hollenbach, "On the Use and Abuse of Spatial Instruments," *Political Analysis*, October 2018, *26* (4), 474–479.

Bloom, Nicholas, Tarek Alexander Hassan, Aakash Kalyani, Josh Lerner, and Ahmed Tahoun, "The Diffusion of Disruptive Technologies," Working Paper 28999, National Bureau of Economic Research July 2021.

Burg, Elco, Tom Elfring, and Joep P. Cornelissen, "Connecting content and structure: A review of mechanisms in entrepreneurs' social networks," *International Journal of Management Reviews*, April 2022, *24* (2), 188–209.

Carlino, Gerald and William R. Kerr, "Agglomeration and Innovation," in "Handbook of Regional and Urban Economics," Vol. 5, Elsevier, 2015, pp. 349–404.

Chatterji, Aaron, Edward Glaeser, and William Kerr, "Clusters of Entrepreneurship and Innovation," *Innovation Policy and the Economy*, January 2014, *14*, 129–166.

Chen, Henry, Paul Gompers, Anna Kovner, and Josh Lerner, "Buy local? The geography of venture capital," *Journal of Urban Economics*, January 2010, *67* (1), 90–102.

Chen, Jiafeng and Jonathan Roth, "Log-like? Identified ATEs defined with zero-valued outcomes are (arbitrarily) scale-dependent," mimeo December 2022.

Chen, Jue and Michael Ewens, "Venture Capitalists' Access to Finance and Its Impact on Startups," mimeo September 2021.

Crisanti, Andrea, Simone Signore, Elitsa Pavlova, and Julien Krantz, "VC Factor," Technical Report, European Investment Fund and Invest Europe September 2021.

Donaldson, Dave and Richard Hornbeck, "Railroads and American Economic Growth: A "Market Access" Approach," *Quarterly Journal of Economics*, May 2016, *131* (2), 799–858.

Ewens, Michael, Alexander Gorbenko, and Arthur Korteweg, "Venture capital contracts," *Journal of Financial Economics*, January 2022, *143* (1), 131–158.

Glaeser, Edward L. and William R. Kerr, "Local Industrial Conditions and Entrepreneurship: How Much of the Spatial Distribution Can We Explain?," *Journal of Economics & Management Strategy*, September 2009, *18* (3), 623–663.

_ **, Sari Pekkala Kerr, and William R. Kerr**, "Entrepreneurship and Urban Growth: An Empirical Assessment with Historical Mines," *Review of Economics and Statistics*, May 2015, *97* (2), 498–520.

_ **, Stuart S. Rosenthal, and William C. Strange**, "Urban economics and entrepreneurship," *Journal of Urban Economics*, January 2010, *67* (1), 1–14.

_ **, William R. Kerr, and Giacomo A.M. Ponzetto**, "Clusters of entrepreneurship," *Journal of Urban Economics*, January 2010, *67* (1), 150–168.

Gompers, Paul A., Will Gornall, Steven N. Kaplan, and Ilya A. Strebulaev, "How do venture capitalists make decisions?," *Journal of Financial Economics*, January 2020, *135* (1), 169–190.

Guzman, Jorge, Inara Tareque, and Dan Wang, "The Impact of High Skilled Immigration on Regional Entrepreneurship," *SSRN Electronic Journal*, 2022.

Harris, Chauncy D., "The Market as a Factor in the Localization of Industry in the United States," *Annals of the Association of American Geographers*, December 1954, *44* (4), 315.

Head, Keith and John Ries, "FDI as an outcome of the market for corporate control: Theory and evidence," *Journal of International Economics*, January 2008, *74* (1), 2–20.

Hornbeck, Richard and Martin Rotemberg, "Growth Off the Rails: Aggregate Productivity Growth in Distorted Economies," mimeo December 2021.

Kaplan, Steven N. and Josh Lerner, "Venture Capital Data: Opportunities and Challenges," in "Measuring Entrepreneurial Businesses: Current Knowledge and Challenges," University of Chicago Press, 2017, pp. 227–246.

Kortum, Samuel and Josh Lerner, "Assessing the Contribution of Venture Capital to Innovation," *RAND Journal of Economics*, 2000, *31* (4), 674–692.

Kraemer-Eis, Helmut, Simone Signore, and Dario Prencipe, "The impact of EIF on the VC ecosystem," EIF Working Paper 34, European Investment Fund 2018.

Kuchler, Theresa, Yan Li, Lin Peng, Johannes Stroebel, and Dexin Zhou, "Social Proximity to Capital: Implications for Investors and Firms," *The Review of Financial Studies*, June 2022, *35* (6), 2743–2789.

Lerner, Josh, "Venture Capitalists and the Oversight of Private Firms," *Journal of Finance*, March 1995, *50* (1), 301–318.

Lerner, Joshua, *Boulevard of broken dreams: why public efforts to boost entrepreneurship and venture capital have failed - and what to do about it* Kauffman foundation series on innovation and entrepreneurship, Princeton, N.J.: Princeton Univ. Press, 2009.

Maats, F., A. Metrick, A. Yasuda, B. Hinkes, and S. Vershovski, "On the Consistency and Reliability of Venture Capital Databases," 2011. Unpublished working paper, University of California at Davis.

Montiel Olea, José Luis and Carolin Pflueger, "A Robust Test for Weak Instruments," *Journal of Business & Economic Statistics*, July 2013, *31* (3), 358–369.

Norton, Edward C., "The inverse hyperbolic sine transformation and retransformed marginal effects," *Stata Journal*, September 2022, *22* (3), 702–712.

NVCA, *National Venture Capital Association Yearbook* 2023.

Okawa, Yohei and Eric van Wincoop, "Gravity in International Finance," *Journal of International Economics*, July 2012, *87* (2), 205–215.

Pellegrino, Bruno, Enrico Spolaore, and Romain Wacziarg, "Barriers to Global Capital Allocation," Technical Report w28694, National Bureau of Economic Research, Cambridge, MA March 2023.

Redding, Stephen J., "Trade and Geography," in "Handbook of International Economics," Vol. 5, Elsevier, 2022, pp. 147–217.

Rehbein, Oliver and Simon Rother, "Social Connectedness (and Distance) in Bank Lending," mimeo September 2022.

Rin, Marco Da, Thomas Hellmann, and Manju Puri, "A Survey of Venture Capital Research," in "Handbook of the Economics of Finance," Vol. 2, Elsevier, 2013, pp. 573–648.

Samila, Sampsa and Olav Sorenson, "Venture Capital, Entrepreneurship, and Economic Growth," *Review of Economics and Statistics*, February 2011, *93* (1), 338–349.

Schnitzer, Monika and Martin Watzinger, "Measuring the Spillovers of Venture Capital," *Review of Economics and Statistics*, March 2022, *104* (2), 276–292.

Wache, Benjamin, "Social Connectedness and the Geography of Venture Capital," mimeo 2023.

Figures and tables

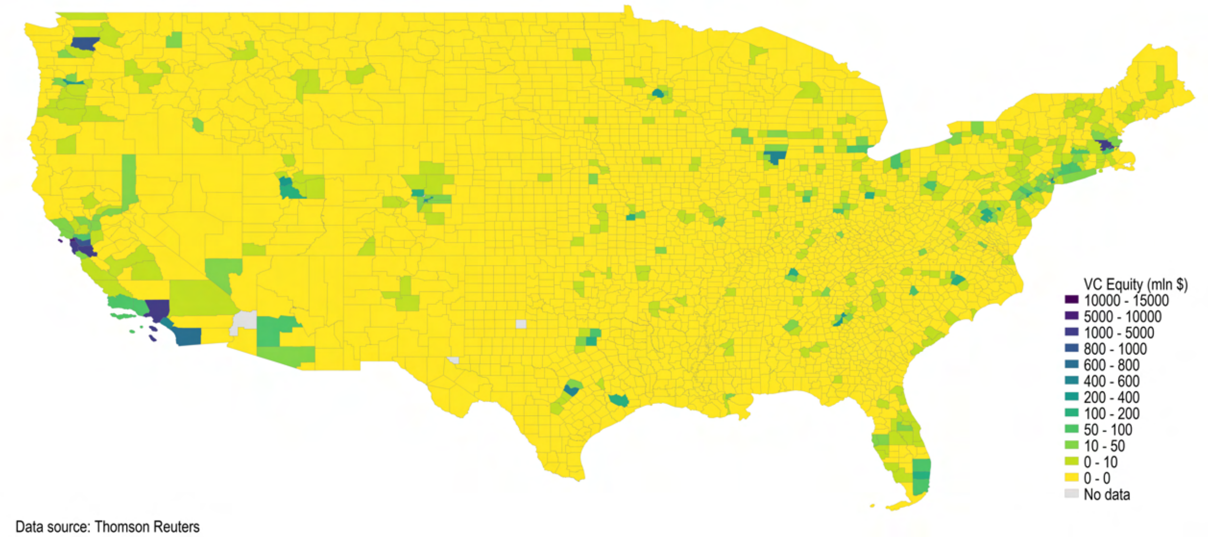


Figure 1: Inward VC investments by destination county, average of 2016-2019

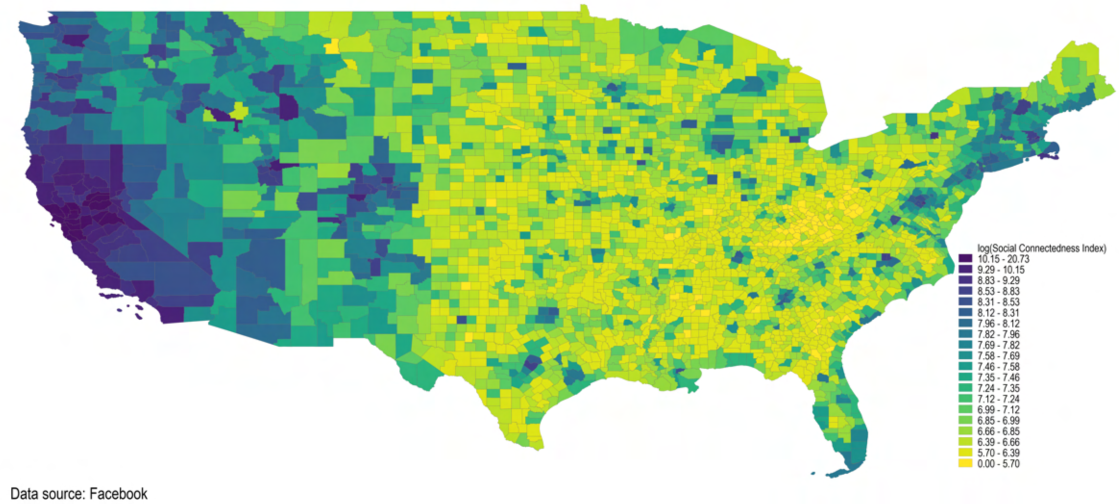


Figure 2: SCI of San Mateo County, CA, (FIPS code: 06081), August 2020

Note: This figure plots the SCI of San Mateo County (06081) with respect to all other counties in August 2020. The figure show twenty quantiles of the distribution of SCI.

Table 1: OLS Regressions of County-level Aggregate Outcomes on VC Investments

Dependent variable →	log Employment			log Payroll			log Establishments		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Inward VC Investment, asinh, t-1	0.011***	0.002***	0.005*	0.019***	0.004***	0.012***	0.010***	0.001***	0.004*
	(0.002)	(0.000)	(0.003)	(0.003)	(0.001)	(0.003)	(0.002)	(0.000)	(0.002)
Observations	102,604	102,589	102,637	102,736	102,722	102,769	103,092	103,092	103,125
County FEs	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
State×Year FEs	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
County×Five Year Period FEs	No	Yes	No	No	Yes	No	No	Yes	No
CSA×Year FEs	No	No	Yes	No	No	Yes	No	No	Yes

Note: Standard errors clustered at the county level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Employment* is the total of full- and part-time employees on the establishment's payroll; *Payroll* includes all forms of compensation paid during the reporting year in thousands of dollars; *Establishments* is the count of individual physical locations at which business is conducted. *Inward VC Investment* is the inverse hyperbolic sine (asinh) of all inward VC equity investments in millions of dollars. County×Five Year Period FEs are county by period fixed effects where periods are five-year long non-overlapping periods starting with 1986-1990. CSA×Year FEs are combined statistical area times year fixed effects.

Table 2: IV Regressions of County-level Aggregate Outcomes on Inward VC Investments

Panel A: Second stage

Dependent variable →	log Employment			log Payroll			log Establishments		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Inward VC Investment, asinh, t-1	0.140***	0.083***	0.098***	0.149***	0.119***	0.098***	0.145***	0.047***	0.105***
	(0.038)	(0.025)	(0.032)	(0.041)	(0.036)	(0.034)	(0.038)	(0.011)	(0.027)

Panel B: First stage

Dependent variable →	Inward VC Investment, asinh								
Social Access to VC (SAVC)	0.076***	0.061***	0.084***	0.076***	0.061***	0.084***	0.076***	0.061***	0.084***
	(0.016)	(0.015)	(0.016)	(0.016)	(0.015)	(0.016)	(0.016)	(0.015)	(0.016)
Observations	55,854	55,840	55,872	55,986	55,973	56,004	56,230	56,229	56,248
First-stage F-stat	22.2	17.3	27.5	22.2	17.3	27.5	22.3	17.4	27.4
County FEs	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
State×Year FEs	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
County×Five Year Period FEs	No	Yes	No	No	Yes	No	No	Yes	No
CSA×Year FEs	No	No	Yes	No	No	Yes	No	No	Yes
SAGDP FEs (100 bins)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors clustered at the county level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Employment* is the total of full- and part-time employees on the establishment's payroll; *Payroll* includes all forms of compensation paid during the reporting year in thousands of dollars; *Establishments* is the count of individual physical locations at which business is conducted. *Inward VC Investment* is the inverse hyperbolic sine (asinh) of all inward VC equity investments in millions of dollars. County×Five Year Period FEs are county by period fixed effects where periods are five-year long non-overlapping periods starting with 1986-1990. CSA×Year FEs are combined statistical area times year fixed effects. SAGDP FEs are 100 fixed effects, one for each percentile of SAGDP. The first-stage F-stat is the cluster robust Kleibergen-Paap Wald statistic.

Table 3: IV Regressions of Untransformed Variables

Panel A: Second stage									
Dependent variable →	Employment			Payroll			Establishments		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Inward VC Investment, t-1	74.1**	49.3**	80.9**	16,149.3**	6,761.4***	17,335.9**	2.104*	0.925**	2.901**
(instrumented)	(35.1)	(20.1)	(36.2)	(8,136.6)	(2,463.4)	(8,329.1)	(1.139)	(0.459)	(1.442)
Anderson-Rubin 95% C.I.	[39.7;...]		[46.1;...]	[6,910.6;...]		[8,698.0;...]	[0.996;...]		[1.524;...]
Panel B: First stage									
Dependent variable →	Inward VC Investment								
Social Access to VC (SAVC)	73.2*	55.9**	62.4**	73.2*	55.9**	62.4**	73.2*	55.8**	62.3**
	(38.0)	(25.2)	(30.9)	(38.0)	(25.2)	(30.9)	(38.0)	(25.2)	(30.9)
Observations	55,836	55,822	55,836	55,968	55,955	55,968	56,212	56,211	56,212
First-stage F-stat	3.7	4.9	4.1	3.7	4.9	4.1	3.7	4.9	4.1
County FEs	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
State×Year FEs	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
County×Five Year Period FEs	No	Yes	No	No	Yes	No	No	Yes	No
CSA×Year FEs	No	No	Yes	No	No	Yes	No	No	Yes
SAGDP FEs (100 bins)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors clustered at the county level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Employment* is the total of full- and part-time employees on the establishment's payroll; *Payroll* includes all forms of compensation paid during the reporting year in thousands of dollars; *Establishments* is the count of individual physical locations at which business is conducted. *Inward VC Investment* is all inward VC equity investments in millions of dollars. County×Five Year Period FEs are county by period fixed effects where periods are five-year long non-overlapping periods starting with 1986-1990. CSA×Year FEs are combined statistical area times year fixed effects. SAGDP FEs are 100 fixed effects, one for each percentile of SAGDP. The first-stage F-stat is the Montiel Olea and Pflueger (2013) statistic in columns 1, 3, 4, 6, 7, and 9, and the Kleibergen-Paap statistic in columns 2, 5, and 8. Anderson-Rubin 95% C.I. are coefficient confidence intervals that are robust to potential weak identification and are efficient in the one-instrument case (Andrews et al., 2019). When the algorithm did not find an (upper) bound we report "...".

Online Appendix (not for publication)

for "The Impact of Venture Capital on Economic Growth"

Steven Poelhekke Benjamin Wache

October 8, 2023

Table of Contents

OA1 Descriptive Statistics

Table OA1: Descriptive Statistics

	N	Coverage	Mean	Median	SD	Min	Max
Dependent Variables							
Employment	105,685	1986 - 2019	34,311.1	6,142	129,562	3	4,007,163
log(Employment)	105,685	1986 - 2019	8.8	8.7	1.7	1.1	15.2
Annual Payroll ($1,000)	105,817	1986 - 2019	1,273,263.1	148,347	6,405,727	31	275,585,585
log(Annual Payroll)	105,817	1986 - 2019	12.0	11.9	1.9	3.4	19.4
Establishments	106,182	1986 - 2019	2,226.2	531	7,603	2	285,383
log(Establishments)	106,182	1986 - 2019	6.4	6.3	1.5	0.7	12.6
Explanatory Variables							
VC Equity Investment ($ mln)	106,182	1986 - 2019	5.7	0.0	103.6	0	11,570.3
asinh(VC Equity Investment)	106,182	1986 - 2019	0.2	0.0	0.9	0	10.0
Social Access to VC (SAVC)	106,179	1986 - 2019	0.5	0.2	1	0.007	59.3
Social Access to GDP (SAGDP)	59,334	2001 - 2019	2.1	2.0	1	0.65	49.0

Note: This table shows descriptive statistics for the main sample used to estimate Equation (2) in Section 4. The unit of observation is a county-year.

OA2 Additional Figures OA1 and OA2, and Table OA2

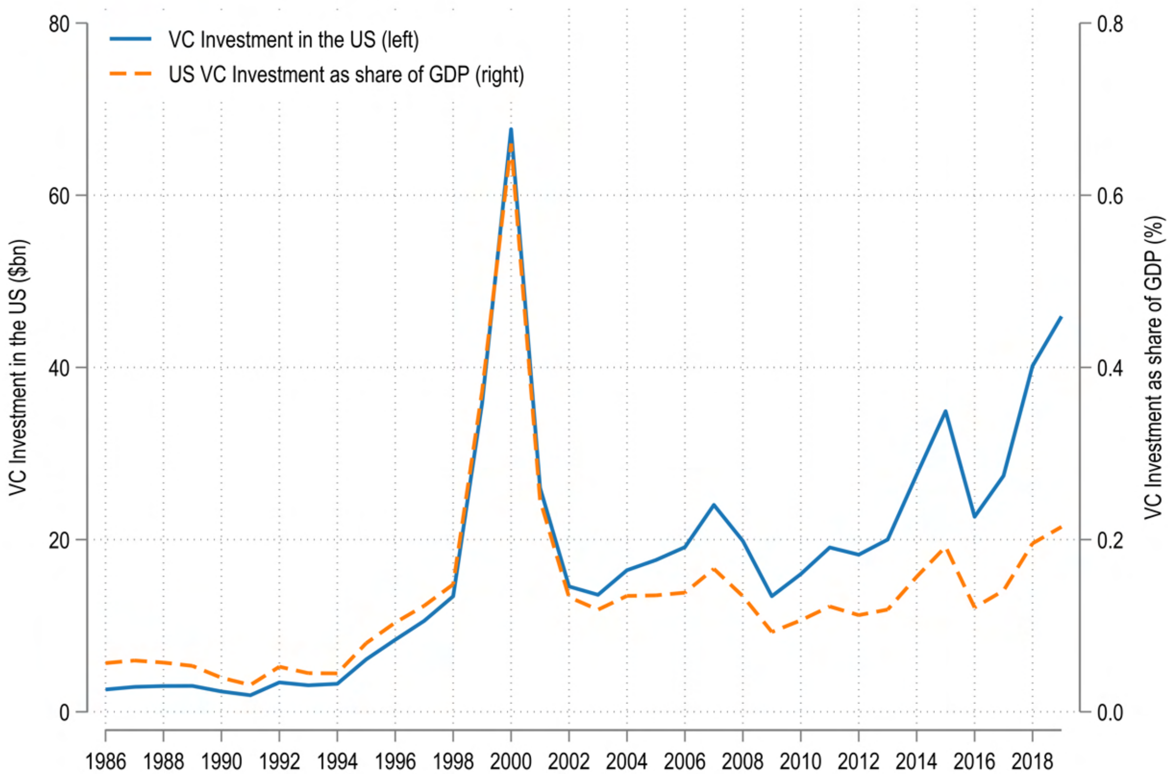


Figure OA1: VC investment, 1986-2019

Description: The left axis of this figure shows the total amount of annual VC investment in the US ($ bn). The right axis shows VC investment as a share of GDP. Data on nominal GDP is taken from the OECD.

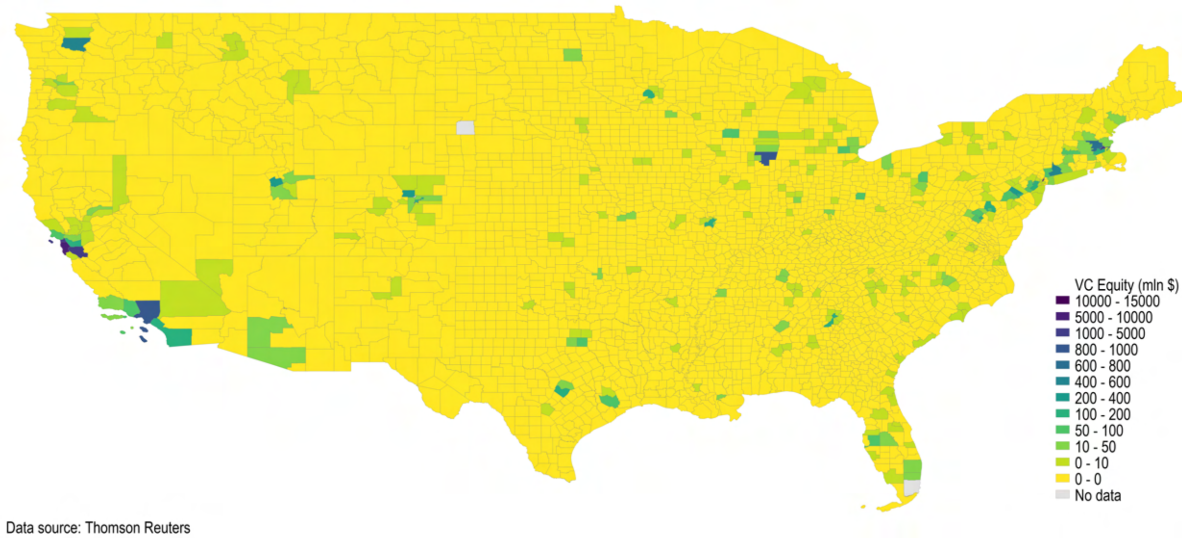


Figure OA2: Outward VC investments by origin county, average of 2016-2019

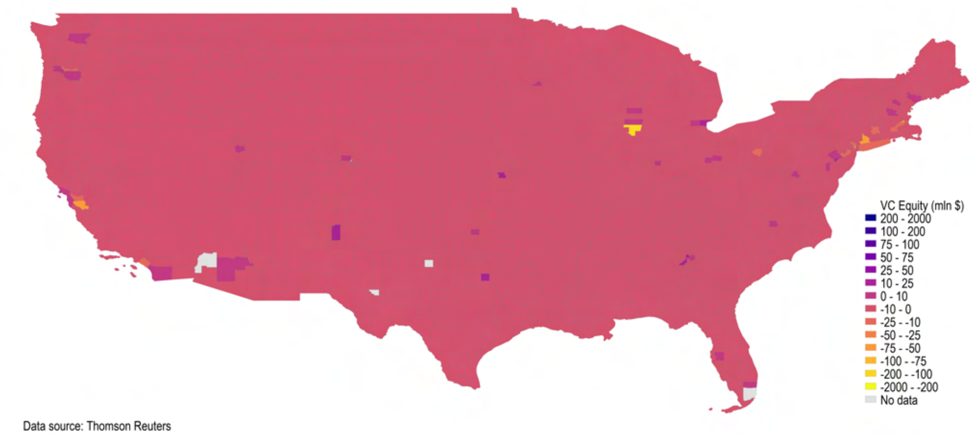
Table OA2: Reduced-form regressions of County Outcomes on Social Access to VC

Dependent variable →	log Employment			log Payroll			log Establishments		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Social Access to VC (SAVC), t-1	0.011***	0.005***	0.008***	0.011***	0.007***	0.008***	0.011***	0.003***	0.009***
	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.001)	(0.002)
Observations	55,854	55,840	55,872	55,986	55,973	56,004	56,230	56,229	56,248
County FEs	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
State×Year FEs	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
County×Five Year Period FEs	No	Yes	No	No	Yes	No	No	Yes	No
CSA×Year FEs	No	No	Yes	No	No	Yes	No	No	Yes
SAGDP FEs (100 bins)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

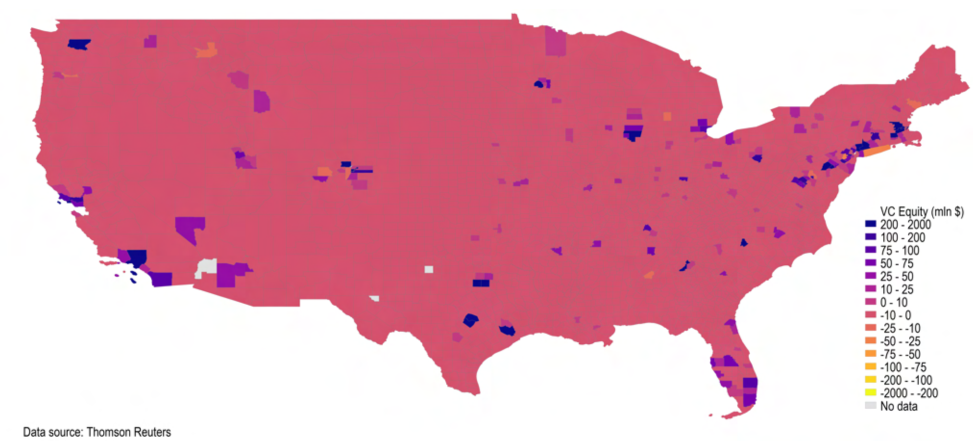
Note: Standard errors clustered at the county level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Employment* is the total of full- and part-time employees on the establishment's payroll; *Payroll* includes all forms of compensation paid during the reporting year in thousands of dollars; *Establishments* is the count of individual physical locations at which business is conducted. *SAVC* is the social access to venture capital. County×Five Year Period FEs are county by period fixed effects where periods are five-year long non-overlapping periods starting with 1986-1990. CSA×Year FEs are combined statistical area times year fixed effects. SAGDP FEs are 100 fixed effects, one for each percentile of SAGDP.

OA3 Annual Change in Outward VC Investment by Origin County (Figures OA3 and OA4)

1990



2000



2001

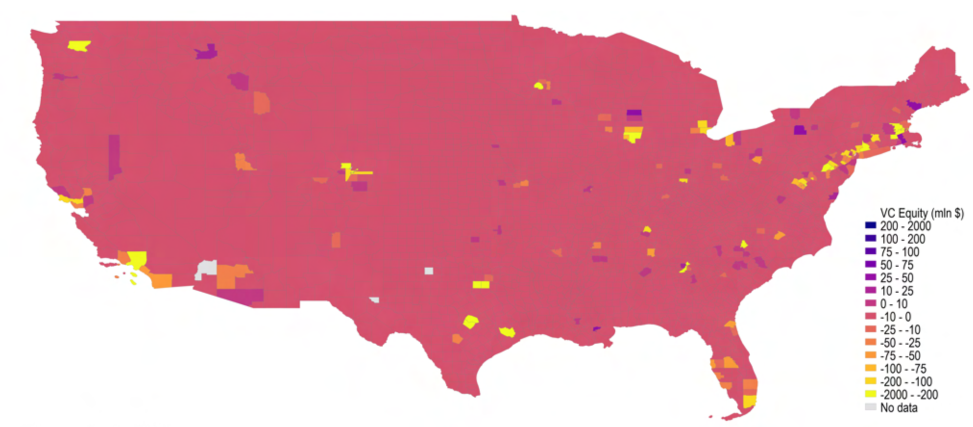
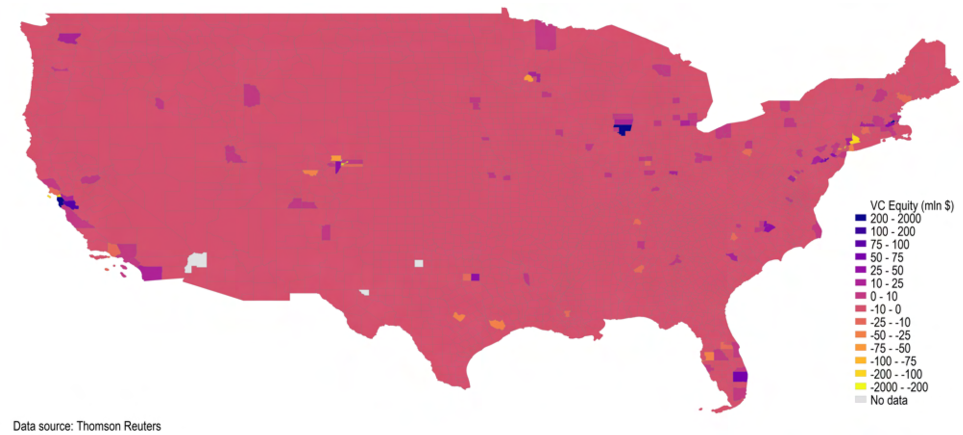
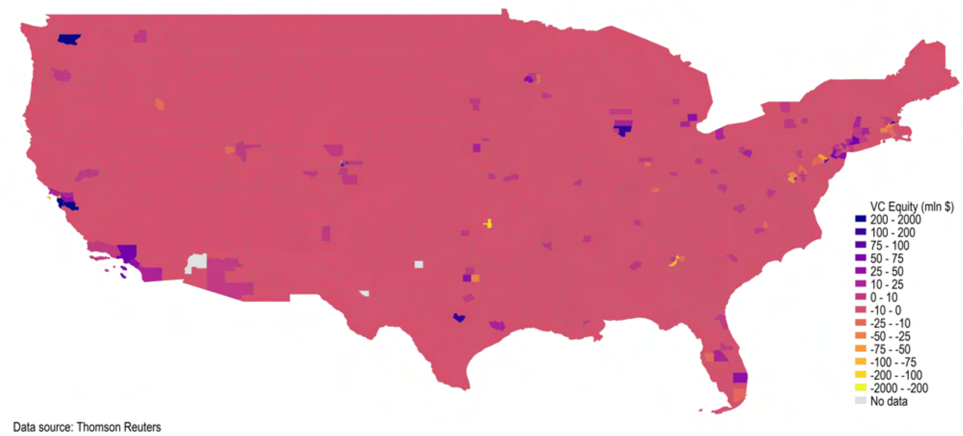


Figure OA3: First-difference of outward investment, by origin county (1990, 2000, 2001)

2005



2010



2015

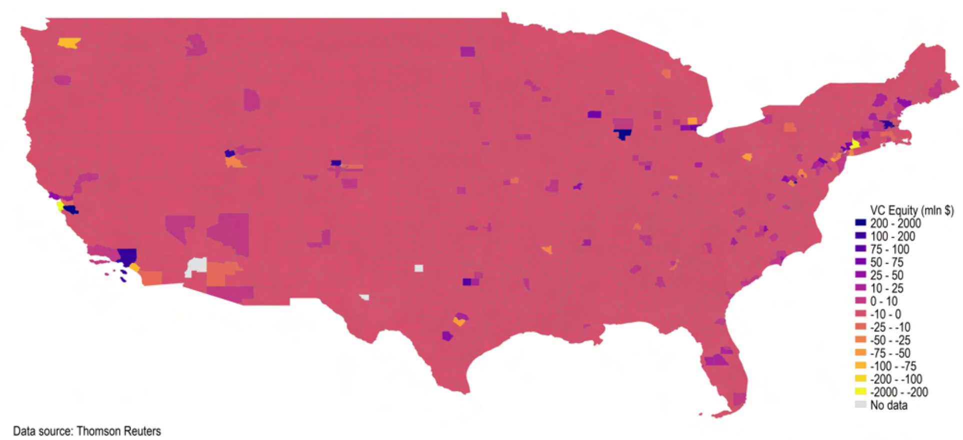


Figure OA4: First-difference of outward investment, by origin county (2005, 2010, 2015)

OA4 IV Regressions with Alternative SAVC Definitions

In this section we present instrumented regressions of county-level economic outcomes on inward VC investment, using alternative definitions of the social access to venture capital (SAVC) instrument (see columns 3, 6, and 9 of Table 2 for the baseline results). The results are robust to the use of alternative definitions.

The definition of SAVC for a given county that we use in the main analysis is given in Section 3, and is, in words: the SCI-weighted sum over all counties, of the outgoing VC investment flows from those counties to all counties except flows towards counties in the same MSA as the given county. Here we use the following alternative definitions: (i) the SCI-weighted sum over all counties, of the outgoing VC investment to all counties except the given county, (ii) the SCI-weighted sum over all counties, of the outgoing VC investment to all counties except flows towards counties in the same CSA as the given county, and (iii) the SCI-weighted sum over all counties except source counties that belong to the same MSA as the given county, of the outgoing VC investment to all counties.

Written out as formulas, the alternative definitions are as follows. SAVC excluding investments to given county i we call $SAVC^h_{i,t}$, and define as:

$$SAVC^h_{i,t} = \sum_{j \in J} SCI_{ij} \cdot \sum_{k \neq i} VC_{jk,t}$$

where $\sum_{k \neq i} VC_{jk,t}$ is the sum of flows from county j to all counties k except to county i. SCI_{ij} is the SCI between counties i and j.

SAVC excluding investments to the given county i's CSA we call $SAVC^c_{i,t}$, and define as:

$$SAVC^c_{i,t} = \sum_{j \in J} SCI_{ij} \cdot \sum_{k \notin C_i} VC_{jk,t}$$

where $\sum_{k \notin C_i} VC_{jk,t}$ is the sum of flows from county j to all counties k that are not in C_i, and where C_i is the set of counties in the same CSA as county i. SCI_{ij} is the SCI between counties i and j.

SAVC excluding investments from counties in the given county i's MSA we call $SAVC^m_{i,t}$, and define as:

$$SAVC_{i,t}^{m} = \sum_{j \notin M_i} SCI_{ij} \cdot \sum_{k \in K} VC_{jk,t}$$

where $\sum_{k \notin M_i} VC_{jk,t}$ is the sum of flows from county j to all counties k, and where M_i is the set of counties in the same MSA as county i. SCI_{ij} is the SCI between counties i and j.

Table OA3 presents the results. Columns 1, 4, and 7 use $SAVC_{i,t}^{h}$ as an instrument; columns 2, 5, and 8 use $SAVC_{i,t}^{c}$ as an instrument; columns 3, 6, and 9 use $SAVC_{i,t}^{m}$ as an instrument. All regression specifications include county fixed effects, and CSA×year fixed effects.

Table OA3: IV Regressions of County-level Aggregate Outcomes on Inward VC Investments

Panel A: Second stage									
Dependent variable →	log Employment			log Payroll			log Establishments		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Inward VC Investment, asinh, t-1	0.086***	0.101***	0.238***	0.109***	0.096***	0.169**	0.097***	0.106***	0.236***
	(0.029)	(0.032)	(0.075)	(0.037)	(0.032)	(0.073)	(0.027)	(0.028)	(0.064)
Panel B: First stage									
Dependent variable →	Inward VC Investment, asinh								
SAVC	0.082***	0.079***	0.069***	0.082***	0.079***	0.069***	0.082***	0.079***	0.069***
	(0.018)	(0.014)	(0.017)	(0.018)	(0.014)	(0.017)	(0.018)	(0.014)	(0.017)
Observations	55,872	55,872	55,872	56,004	56,004	56,004	56,248	56,248	56,248
First-stage F-stat	19.9	30.3	17.5	19.9	30.3	17.5	19.9	30.3	17.4
Instrument →	$SAVC^h$	$SAVC^c$	$SAVC^m$	$SAVC^h$	$SAVC^c$	$SAVC^m$	$SAVC^h$	$SAVC^c$	$SAVC^m$
County FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CSA×Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
SAGDP FEs (# of bins)	100	100	100	100	100	100	100	100	100

Note: Standard errors are clustered at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Employment* is the total of full- and part-time employees on the establishment's payroll; *Payroll* includes all forms of compensation paid during the reporting year in thousands of dollars; *Establishments* is the count of individual physical locations at which business is conducted. *Inward VC Investment* is the inverse hyperbolic sine (asinh) of all inward VC equity investments in millions of dollars. CSA×Year FEs are combined statistical area times year fixed effects. SAGDP FEs are 100 fixed effects, one for each percentile of SAGDP. The first-stage F-stat is the Kleibergen-Paap statistic. In columns 1, 4, 7, $SAVC^h$ is used as the instrument; in columns 2, 5, 8, $SAVC^c$ is used as the instrument; in columns 3, 6, 9, $SAVC^m$ is used as the instrument.

OA5 IV Regressions with Alternative Clustering

In this section we present instrumented regressions of county-level economic outcomes on inward VC investment, using alternative ways of clustering the standard errors (see columns 3, 6, and 9 of Table 2 for the baseline results).

Table OA4 presents the results. Standard errors in columns 1, 4, and 7 are clustered at the state level; in columns 2, 5, and 8 are clustered at the CSA level; in columns 3, 6, and 9 two-way clustered at the county and year level. All regression specifications include county fixed effects, CSA×year fixed effects, and 100 SAGDP bucket fixed effects.

All second stage results remain statistically significant, and all first stage F-tests are comfortably above 10.

Table OA4: IV Regressions of County-level Aggregate Outcomes on Inward VC Investments

Panel A: Second stage									
Dependent variable →	log Employment			log Payroll			log Establishments		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Inward VC Investment, asinh, t-1	0.098*	0.098**	0.098**	0.098**	0.098***	0.098**	0.105**	0.105***	0.105**
	(0.049)	(0.040)	(0.038)	(0.039)	(0.029)	(0.038)	(0.045)	(0.036)	(0.038)
Panel B: First stage									
Dependent variable →	Inward VC Investment, asinh								
SAVC	0.084***	0.084***	0.084***	0.084***	0.084***	0.084***	0.084***	0.084***	0.084***
	(0.012)	(0.015)	(0.020)	(0.012)	(0.015)	(0.020)	(0.012)	(0.015)	(0.020)
Observations	55,872	55,872	55,872	56,004	56,004	56,004	56,248	56,248	56,248
First-stage F-stat	45.1	32.3	18.0	45.1	32.3	18.0	44.9	32.1	17.9
County FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CSA×Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
SAGDP FEs (100 bins)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clustering	State	CSA	County & Year	State	CSA	County & Year	State	CSA	County & Year

Note: Standard errors are clustered as follows. Columns 1, 4, and 7 are clustered at the state level. Columns 2, 5, and 8 are clustered at the CSA level. Columns 3, 6, and 9 are two-way clustered at the county and year level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Employment* is the total of full- and part-time employees on the establishment's payroll; *Payroll* includes all forms of compensation paid during the reporting year in thousands of dollars; *Establishments* is the count of individual physical locations at which business is conducted. *Inward VC Investment* is the inverse hyperbolic sine (asinh) of all inward VC equity investments in millions of dollars. County×Five Year Period FEs are county by period fixed effects where periods are five-year long non-overlapping periods starting with 1986-1990. CSA×Year FEs are combined statistical area times year fixed effects. SAGDP FEs are 100 fixed effects, one for each percentile of SAGDP. The first-stage F-stat is the Kleibergen-Paap statistic.

OA6 IV Regressions with Alternative SAGDP Quantiles

In this section we present instrumented regressions of county-level economic outcomes on inward VC investment, using alternative ways of controlling for SAGDP (see columns 3, 6, and 9 of Table 2 for the baseline results).

Table OA5 presents the results. Columns 1, 4, and 7 do not include any explicit controls for SAGDP; columns 2, 5, and 8 include 50 SAGDP fixed effects bins (one for each 50th quantile part of the distribution); columns 3, 6, and 9 include 500 SAGDP fixed effects bins (one for each 500th quantile part of the distribution). All regression specifications include county fixed effects, and CSA×Year fixed effects.

Including the dummies yields more conservative results, but it matters less whether we include 50, 100, or 500 dummies of the distribution of SAGDP.

Table OA5: IV Regressions of County-level Aggregate Outcomes on Inward VC Investments

Panel A: Second stage									
Dependent variable →	log Employment			log Payroll			log Establishments		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Inward VC Investment, asinh, t-1	0.144***	0.096***	0.102***	0.179***	0.096***	0.104***	0.131***	0.103***	0.106***
	(0.042)	(0.034)	(0.033)	(0.050)	(0.035)	(0.034)	(0.034)	(0.028)	(0.028)
Panel B: First stage									
Dependent variable →	Inward VC Investment, asinh								
SAVC	0.084***	0.084***	0.084***	0.084***	0.084***	0.084***	0.084***	0.083***	0.083***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Observations	55,872	55,872	55,872	56,004	56,004	56,004	56,248	56,248	56,248
First-stage F-stat	26.9	27.6	27.2	26.9	27.6	27.2	26.8	27.6	27.1
County FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CSA×Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
SAGDP FEs (# of bins)	-	50	500	-	50	500	-	50	500

Note: Standard errors are clustered at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Employment* is the total of full- and part-time employees on the establishment's payroll; *Payroll* includes all forms of compensation paid during the reporting year in thousands of dollars; *Establishments* is the count of individual physical locations at which business is conducted. *Inward VC Investment* is the inverse hyperbolic sine (asinh) of all inward VC equity investments in millions of dollars. CSA×Year FEs are combined statistical area times year fixed effects. SAGDP FEs are no/50/500 fixed effects, respectively, one for each quantile of SAGDP. When included, the bins are created across the entire sample period. The first-stage F-stat is the Kleibergen-Paap statistic.

OA7 IV Regressions without High VC Locations

In this section we present instrumented regressions of county-level economic outcomes on inward VC investment, while we exclude states from the sample that could potentially have a distorting effect on the whole sample (see columns 3, 6, and 9 of Table 2 for the baseline results).

Table OA6 presents the results. Columns 1, 4, and 7 exclude California; columns 2, 5, and 8 exclude New York state, and Massachusetts; columns 3, 6, and 9 exclude all three states from the sample. All regression specifications include county fixed effects, and CSA×Year fixed effects.

Table OA6: IV Regressions of County-level Aggregate Outcomes on Inward VC Investments

Panel A: Second stage									
Dependent variable →	log Employment			log Payroll			log Establishments		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Inward VC Investment, asinh, t-1	0.153***	0.113***	0.230***	0.135***	0.117***	0.202***	0.153***	0.112***	0.212***
	(0.056)	(0.040)	(0.085)	(0.049)	(0.043)	(0.077)	(0.048)	(0.033)	(0.073)
Panel B: First stage									
Dependent variable →	Inward VC Investment, asinh								
SAVC	0.078***	0.074***	0.061***	0.078***	0.074***	0.061***	0.078***	0.073***	0.061***
	(0.018)	(0.017)	(0.017)	(0.018)	(0.017)	(0.017)	(0.018)	(0.017)	(0.017)
Observations	54,828	54,505	53,461	54,960	54,637	53,593	55,204	54,880	53,836
First-stage F-stat	19.0	19.5	13.1	19.0	19.5	13.072	18.9	19.4	13.0
County FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CSA×Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
SAGDP FEs (# of bins)	100	100	100	100	100	100	100	100	100
States excluded	CA	NY, MA	CA, NY, MA	CA	NY, MA	CA, NY, MA	CA	NY, MA	CA, NY, MA

Note: Standard errors are clustered at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Employment* is the total of full- and part-time employees on the establishment's payroll; *Payroll* includes all forms of compensation paid during the reporting year in thousands of dollars; *Establishments* is the count of individual physical locations at which business is conducted. *Inward VC Investment* is the inverse hyperbolic sine (asinh) of all inward VC equity investments in millions of dollars. CSA×Year FEs are combined statistical area times year fixed effects. SAGDP FEs are 100 fixed effects, respectively, one for each percentile of SAGDP. The first-stage F-stat is the Kleibergen-Paap statistic. Columns 1, 4, and 7 exclude California; columns 2, 5, and 8 exclude New York and Massachusetts; columns 3, 6, and 9 exclude all three states from the sample.

OA8 Regressions with Alternative Scaling

OLS. In this section we present OLS regressions of county-level economic outcomes on inward VC investment, using alternative scalings of the explanatory variable *Inward VC Investment*.

Table OA7 presents results with *Inward VC Investment* expressed in unit dollars, and Table OA8 presents results with *Inward VC Investment* expressed in thousands of dollars.

Compared to the results in Table 1, which measures investment in millions in keeping with how investment is reported in the raw data, the below tables show that scaling matters in terms of magnitudes, but also that the results are robust in terms of significance.

Table OA7: OLS Regressions of County-level Aggregate Outcomes on VC Investments: investment rescaled to unit dollars

Dependent variable →	log Employment			log Payroll			log Establishments		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Inward VC Investment, asinh, t-1	0.002***	0.000***	0.001***	0.002***	0.000	0.002***	0.001***	0.000**	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	102,604	102,589	102,637	102,736	102,722	102,769	103,092	103,092	103,125
County FEs	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
State×Year FEs	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
County×Five Year Period FEs	No	Yes	No	No	Yes	No	No	Yes	No
CSA×Year FEs	No	No	Yes	No	No	Yes	No	No	Yes

Note: Standard errors clustered at the county level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Employment* is the total of full- and part-time employees on the establishment's payroll; *Payroll* includes all forms of compensation paid during the reporting year in thousands of dollars; *Establishments* is the count of individual physical locations at which business is conducted. *Inward VC Investment* is the inverse hyperbolic sine (asinh) of all inward VC equity investments in dollars. County×Five Year Period FEs are county by period fixed effects where periods are five-year long non-overlapping periods starting with 1986-1990. CSA×Year FEs are combined statistical area times year fixed effects.

IV. Table OA9 presents results for the iunstrumented regressions, with *Inward VC Investment* expressed in dollars. Table OA10 presents IV results with *Inward VC Investment* expressed in thousands of dollars.

Compared to the results in Table 2, which measures investment in millions in keeping with how investment is reported in the raw data, the below tables show again that scaling matters in terms of magnitudes, but also that the results are robust in terms of significance. However, the first stages show that the instrument has more difficulty predicting investment,

Table OA8: OLS Regressions of County-level Aggregate Outcomes on VC Investments: investment rescaled to thousands of dollars

Dependent variable →	log Employment			log Payroll			log Establishments		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Inward VC Investment, asinh, t-1	0.003***	0.000***	0.002***	0.005***	0.000**	0.004***	0.003***	0.000***	0.002***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)
Observations	102,604	102,589	102,637	102,736	102,722	102,769	103,092	103,092	103,125
County FEs	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
State×Year FEs	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
County×Five Year Period FEs	No	Yes	No	No	Yes	No	No	Yes	No
CSA×Year FEs	No	No	Yes	No	No	Yes	No	No	Yes

Note: Standard errors clustered at the county level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Employment* is the total of full- and part-time employees on the establishment's payroll; *Payroll* includes all forms of compensation paid during the reporting year in thousands of dollars; *Establishments* is the count of individual physical locations at which business is conducted. *Inward VC Investment* is the inverse hyperbolic sine (asinh) of all inward VC equity investments in thousands of dollars. County×Five Year Period FEs are county by period fixed effects where periods are five-year long non-overlapping periods starting with 1986-1990. CSA×Year FEs are combined statistical area times year fixed effects.

with some F-tests dropping below 10.

Table OA9: IV Regressions of County-level Aggregate Outcomes on Inward VC Investments: investment rescaled to unit dollars

Panel A: Second stage

Dependent variable →	log Employment			log Payroll			log Establishments		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Inward VC Investment, asinh, t-1	0.059***	0.034*	0.042***	0.063***	0.049*	0.042**	0.062***	0.019**	0.044***
	(0.019)	(0.018)	(0.016)	(0.021)	(0.026)	(0.016)	(0.019)	(0.009)	(0.014)

Panel B: First stage

Dependent variable →	Inward VC Investment, asinh								
SAVC	0.178***	0.147*	0.197***	0.178***	0.147*	0.197***	0.178***	0.148*	0.197***
	(0.045)	(0.077)	(0.051)	(0.045)	(0.077)	(0.051)	(0.045)	(0.077)	(0.051)
Observations	55,854	55,840	55,872	55,986	55,973	56,004	56,230	56,229	56,248
First-stage F-stat	15.6	3.7	14.8	15.6	3.7	14.8	15.6	3.7	14.8
County FEs	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
State×Year FEs	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
County×Five Year Period FEs	No	Yes	No	No	Yes	No	No	Yes	No
CSA×Year FEs	No	No	Yes	No	No	Yes	No	No	Yes
SAGDP FEs (100 bins)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors clustered at the county level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Employment* is the total of full- and part-time employees on the establishment's payroll; *Payroll* includes all forms of compensation paid during the reporting year in thousands of dollars; *Establishments* is the count of individual physical locations at which business is conducted. *Inward VC Investment* is the inverse hyperbolic sine (asinh) of all inward VC equity investments in dollars. County×Five Year Period FEs are county by period fixed effects where periods are five-year long non-overlapping periods starting with 1986-1990. CSA×Year FEs are combined statistical area times year fixed effects. SAGDP FEs are 100 fixed effects, one for each percentile of SAGDP. The first-stage F-stat is the Kleibergen-Paap statistic.

Table OA10: IV Regressions of County-level Aggregate Outcomes on Inward VC Investments: investment rescaled to thousands of dollars

Panel A: Second stage

Dependent variable →	log Employment			log Payroll			log Establishments		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Inward VC Investment, asinh, t-1	0.085^{***}	0.050^{**}	0.059^{***}	0.091^{***}	0.072^{**}	0.059^{***}	0.088^{***}	0.028^{***}	0.063^{***}
	(0.025)	(0.024)	(0.021)	(0.027)	(0.033)	(0.022)	(0.024)	(0.011)	(0.018)

Panel B: First stage

Dependent variable →	Inward VC Investment, asinh								
SAVC	0.124^{***}	0.100^{**}	0.138^{***}	0.124^{***}	0.100^{**}	0.138^{***}	0.124^{***}	0.100^{**}	0.138^{***}
	(0.027)	(0.044)	(0.030)	(0.027)	(0.044)	(0.030)	(0.027)	(0.044)	(0.030)
Observations	55,854	55,840	55,872	55,986	55,973	56,004	56,230	56,229	56,248
First-stage F-stat	21.1	5.2	20.8	21.1	5.2	20.8	21.2	5.3	20.8
County FEs	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
State×Year FEs	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
County×Five Year Period FEs	No	Yes	No	No	Yes	No	No	Yes	No
CSA×Year FEs	No	No	Yes	No	No	Yes	No	No	Yes
SAGDP FEs (100 bins)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors clustered at the county level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Employment* is the total of full- and part-time employees on the establishment's payroll; *Payroll* includes all forms of compensation paid during the reporting year in thousands of dollars; *Establishments* is the count of individual physical locations at which business is conducted. *Inward VC Investment* is the inverse hyperbolic sine (asinh) of all inward VC equity investments in thousands of dollars. County×Five Year Period FEs are county by period fixed effects where periods are five-year long non-overlapping periods starting with 1986-1990. CSA×Year FEs are combined statistical area times year fixed effects. SAGDP FEs are 100 fixed effects, one for each percentile of SAGDP. The first-stage F-stat is the Kleibergen-Paap statistic.