DO SMALL BUSINESSES CREATE MORE JOBS? NEW EVIDENCE FROM THE
NATIONAL ESTABLISHMENT TIME SERIES

David Neumark
Brandon Wall
Junfu Zhang

Do Small Businesses Create More Jobs? New Evidence from the National Establishment Time
Series
David Neumark, Brandon Wall, and Junfu Zhang
NBER Working Paper No. 13818
February 2008
JEL No. J20,L25,L53

ABSTRACT

We use a new database, the National Establishment Time Series (NETS), to revisit the debate about
the role of small businesses in job creation. Birch (e.g., 1987) argued that small firms are the most
important source of job creation in the U.S. economy, but Davis et al. (1996a) argued that this conclusion
was flawed, and based on improved methods and using data for the manufacturing sector they concluded
that there was no relationship between establishment size and net job creation. Using the NETS data,
we examine evidence for the overall economy, as well as for different sectors. The results indicate
that small establishments and small firms create more jobs, on net, although the difference is much
smaller than what is suggested by Birch's methods. However, the negative relationship between establishment
size and job creation is much less clear for the manufacturing sector, which may explain some of the
earlier findings contradicting Birch's conclusions.

David Neumark
Department of Economics
University of California, Irvine
3151 Social Science Plaza
Irvine, CA 92697-5100
and NBER
dneumark@uci.edu

Brandon Wall
Department of Economics
Stanford University
579 Serra Mall
Stanford, CA 94305-6072
brandon.wall@gmail.com

Junfu Zhang
Deparment of Economics
Clark University
Worcester, MA 01610
juzhang@clarku.edu

1. Introduction

Ever since the provocative work of David Birch (1979, 1981, and 1987), researchers have paid

considerable attention to the role of small businesses in job creation.[1] The central thesis of Birch's

research was that small firms are the most important source of job creation in the U.S. economy. Birch

(1979, 1981) provided the first evidence in support of the argument that small businesses are the primary

engines of job growth, claiming that 66% of all net new jobs in the United States during 1969-1976 were

created by firms with 20 or fewer employees and 81.5% were created by firms with 100 or fewer

employees. Later, Birch (1987) argued that during the period 1981-1985, firms with fewer than 20

employees accounted for 82% of employment growth via expansion and contraction of existing firms

(Figure 1-3, p. 14), and 88.1% of overall employment growth (Figure 1-5, p. 16).

Birch's argument about the important role of small business in job creation captured the

imagination of both the popular press and politicians because it fit perfectly into the historical context.

The U.S. federal government had a long tradition of supporting small businesses. As early as 1953, the

U.S. Congress passed the Small Business Act with the intention to aid, counsel, assist, and protect the

interests of small businesses. Under this act, the federal government established the Small Business

Administration to ensure that a fair proportion of government contracts went to small businesses. The

Small Business Administration also makes direct loans and guarantees bank loans to small businesses,

based on the perspective that assisting small businesses is important to preserve free competitive

enterprise and to maintain and strengthen the U.S. economy.[2] Birch's findings fed into this thinking and

quickly became conventional wisdom.[3] Since then, his findings have often been cited as justification for

favorable government regulations, tax incentives, and support programs for small businesses.[4] Birch's

---

[1] See, for example, Baldwin and Picot (1995), Barnes and Haskel (2002), Broersma and Gautier (1997), Davis et al. (1996a, 1996b), Hohti (2000), and Wagner (1995).
[2] See http://www.sba.gov/aboutsba/index.html (viewed October 31, 2007).
[3] See Davis et al. (1996b) for a long list of quotes from the media and political speeches that recite this conventional wisdom.
[4] A series of legislative actions has continued the tradition of helping small businesses since Birch's work, including, for example: the Small Business Economic Policy Act of 1980; the Small Business Innovation Development Act of 1982; the Small Business Competitiveness Demonstration Program Act of 1988; the Small Business Technology Transfer Act of 1992; the National Small Business Act of 1996; the Small Business Job Protection Act of 1996; the

work appears to have had a lasting effect on U.S. policy. For example, the U.S. Small Business Administration has an Office of Advocacy (for small businesses) that still trumpets Birch's findings in trying to help small businesses with regard to regulatory constraints, taxation, etc.[5]

Birch's work also attracted considerable criticism. Brown et al. (1990) point out that the number of jobs created by an employer is not the only thing that matters. Spurred by Birch's findings and the attention they attracted, Brown et al. (1990) compare jobs created by small firms with those created by large firms. They argue that jobs created by small firms are less desirable because they tend to exist for a shorter period of time, pay lower wages, and have less generous fringe benefits in terms of health insurance, vacation, and pension plans. In addition, they argue that small firms tend to have poorer working conditions, provide less job training, and have higher job turnover rates. Based on these findings, they recommend that policymakers should be cautious about encouraging policies to help small firms. Although the argument of Brown et al. is an important one, it is not the focus of our paper.

Second, focusing more specifically on the job creation debate, Davis et al. (1996a) criticize the calculation used by Birch, which divides firms into size classes and then examines the variation of job growth across size classes. In particular, they argue that these calculations are subject to a "regression fallacy" that leads to upward bias in the estimated contribution of small firms to job growth. When they use a calculation that, they argue, is not prone to this bias, with data on the manufacturing sector from the Longitudinal Research Database (LRD) covering the period 1973-1988, they conclude that employment growth appears to have "no systematic relationship to average plant size" (p. 68).

There is a parallel, yet largely separate, industrial organization literature that considers a related question: How does employment growth vary by firm size? Much of this literature responds to what is

Small Business Regulatory Enforcement Fairness Act of 1996; the Small Business Reauthorization Act of 1997; the Small Business Franchise Act of 1999; the Small Business Procurement Competition Act of 2001; the Small Business Paperwork Relief Act of 2002; and the Small Business Health Fairness Act of 2005. In determining eligibility for most of these programs, the Small Business Administration specifies size standards on an industry-by-industry basis, most commonly based on number of employees, but sometimes based instead on annual receipts. The employee size thresholds range from 100 for wholesale trade to 1,500 for industries such as petroleum refineries, air transportation, and telecommunications. (See http://www.sba.gov/idc/groups/public/documents/sba_homepage/serv_sstd_tablepdf.pdf, viewed October 31, 2007, for a complete list.)
[5] See http://www.sba.gov/advo/25ann.html and http://www.sba.gov/advo/mission.html, viewed October 7, 2007.

known as "Gibrat's Law," which states that firm growth rates are independent of size (Gibrat, 1931). As a behavioral conjecture proposed to explain the observed firm size distribution, Gibrat's law has generated a large amount of research testing its validity; see the review in Sutton (1997). This research is motivated by the desire for a better understanding of the behavior of industrial organizations rather than by the goal of informing policy with respect to job creation by small versus large businesses. Nevertheless, it does have implications for the question of whether small firms create more jobs. For example, if firm growth is indeed independent of firm size, then on average the proportion of jobs created by small firms is solely determined by the employment share of small firms. In other words, if Gibrat's law holds and if large firms have a bigger employment share, it is impossible for small firms to create most of the new jobs.

In this paper, we use a new database, the National Establishment Time Series (NETS), which was constructed using the most recent waves of the Dun and Bradstreet (D&B) data. The availability of this new database provides a valuable opportunity to revisit the Birch thesis as well as the conclusions drawn by Davis et al. Given that Davis et al.'s results come solely from the manufacturing sector, it is important to ask whether their conclusions apply to the overall economy, and to other sectors of the economy, which we are able to do using the NETS data (although because of data constraints our analysis is limited to California).[6] Also, the answer to the questions addressed by this research may depend on the period studied. In particular, innovative start-ups appeared to have played an important role in the Internet revolution and the boom of the "new economy" in the late 1990s (Audretsch and Thurik, 2001), raising

---

[6] Indeed, there are reasons to believe that growth patterns in the manufacturing sector may not apply to the rest of the economy. As Klepper and Graddy (1990) observed, the number of firms in an industry will grow first, then decline sharply, and level off eventually. Thus, new and mature industries tend to show different industrial structures and growth patterns. For example, one may find a large number of new (and thus small) firms in a new industry, which grow fast and contribute a large share of employment growth. On the contrary, a mature industry may be shrinking and usually it is the smaller firms that exit first and thus constitute a large share of job loss. One of the most prominent features of the U.S. manufacturing sector is that its total employment (relative to the U.S. total) has been continuously declining since 1950 (Sachs and Shatz, 1994). Many manufacturing industries are mature ones and it is not surprising to find a limited role of small firms in job creation in that sector. But non-manufacturing sectors are different in that there are many new industries in which small firms may well create a larger share of new jobs.

the possibility that in this period (and extending beyond the manufacturing sector) new (and thus mostly small) firms contributed more to employment growth.[7,8]

2. Previous Research

2.1 Birch's hypothesis and Davis et al.'s critique

Birch's work prompted numerous researchers to consider his methods and to test his hypothesis in different countries.[9] Focusing first on methods, Davis et al. (1996a) criticized the statistical analysis leading to Birch's conclusions, pointing out that when businesses are classified by size for a given base year, two types of firms will "accidentally" fall into a smaller size category: firms that are not small but have just experienced a transitory negative shock to their employment and appear to be small for the moment; and firms that are not small but are mistakenly classified as small due to random measurement errors. If such transitory shocks or random measurement errors are not highly serially correlated, both types of "small" firms will "grow" fast over the next year only because of regression to the mean. A reverse argument can be made for large firms. That is, some large firms are categorized as large only because of positive transitory shocks and random measurement errors. These firms tend to experience negative growth if the transitory shocks or the measurement errors are serially uncorrelated, leading to a downward bias in the estimated growth rate of large firms.[10,11] To avoid this statistical pitfall, Davis et al. propose computing job creation and destruction rates from the base period *t* to period *t+1* relative to the

---

[7] Although young firms and small firms overlap a great deal, they are conceptually distinct groups. Most research, following the tradition established by Birch, focuses on the dichotomy of small and large firms. See Neumark et al. (2006) for some discussion of the contribution to job growth by young and old firms.

[8] Differences associated with the "new economy" may be particularly pronounced in data from California.

[9] Much of the latter research appeared in *Small Business Economics*, the founding of which was partly inspired by Birch's work (Acs and Audretsch, 1989).

[10] Leonard (1987) reports some direct evidence consistent with regression to the mean in establishment size (Table 6.7), and suggests that it is attributable to real dynamics, rather than measurement error (p. 154).

[11] Davis et al. (1996a) and Kirchhoff and Green (1998) point out that many studies prior to Birch (1979), which did not use longitudinal data, could have either overstated or understated job growth due to smaller firms or larger firms, depending on whether firms were on average shrinking or growing. Lacking establishment-level data, these pre-Birch studies typically calculated the number of new jobs in each size category as the difference between the number of jobs in the size category this year and the number of jobs in the same size category in the previous year. This procedure wrongly assumes that firms do not move between size categories. Thus, for example, in periods in which firms are on average growing, this results in overstatement of job growth in large firms because jobs created when firms *grow into* a larger size class are mistakenly interpreted as job creation by firms *in* that size class.

average employment level in these two periods.[12] Davis et al. argue that this regression fallacy fully explains the relationship between establishment size and job growth. As noted in the Introduction, using the average instead of a base-year size measure, they find no systematic relationship between size and employment growth.[13] They also suspect that the measurement error in the D&B data used by Birch is more serious than in their LRD data, and thus that the regression fallacy bias is even more severe in Birch's results (Davis et al., 1996a, p. 70).[14]

However, in the broader research literature on firm (or establishment) size and job creation prompted by Birch's work, which covers many countries, analyses that use Davis et al.'s (1996a) method to avoid the regression fallacy more often than not confirm the finding that small establishments or firms contribute more to job growth. For example, Kirchhoff and Phillips (1988) examine the contribution of small and large firms to U.S. job growth. They find that firms with fewer than 100 employees are the major sources of net new job creation. In contrast, firms with more than 1,000 employees provided only 13% of all new jobs despite their 37% employment share in 1985. Using Canadian data on the manufacturing sector, Baldwin and Picot (1995) find that net job creation by small manufacturing establishments is greater than that of large establishments. Broersma and Gautier (1997) and Barnes and Haskel (2002), using data for the Netherlands and the United Kingdom, respectively, find that small

---

[12] Somewhat confusingly, they refer to this as the "current" establishment size measure. Davis et al. (1996b, 1996a) also define a historical average size as a weighted average of employment over all years of an establishment's existence in the sample period, using employment as the weight. They consider this as a measure of the business establishment's "intended scale of operations." By this definition, a business is small only if it is intended to be small and thus stays small for many years. From this perspective, employment fluctuations at the business establishment level are considered mostly random, as a result of changes in demand or other market conditions. Carree and Klomp (1996) cast doubt on the usefulness of this concept of long-run firm size. First, facing constant environmental changes, a firm should and will continuously adjust its optimal size. This reality is in conflict with the concept of long-run firm size that is assumed to be stable over a long period of time. Second, for young firms that just entered the market, it is simply impossible to know their intended long-run size. We do not use the latter measure, because over a longer period the definition of size can change. Our "average" measure corresponds to Davis et al.'s "current" measure, as "average" seems a more appropriate label.

[13] Davis et al. also extend their criticism to similar calculations included in annual reports to the President of the Small Business Administration (in the mid- to late-1980s).

[14] In addition to avoiding the regression fallacy, the average size is widely used also because it captures job creation and destruction symmetrically (e.g., Pivetz et al., 2001, pp. 14-15). For example, suppose a firm's number of employees is 1, 2, and 1 in three consecutive years. If we use base-year employment as the denominator to calculate growth rate, this firm has experienced 100% job growth followed by 50% job decline. Alternatively, if we use two-year average employment as the denominator, this firm has 67% job growth (= 1/1.5) followed by 67% job decline (= −1/1.5).

firms/establishments contribute more to net job creation than large ones. Voulgaris et al. (2005) also find that small firms create more jobs on net, using data from Greece. On the other hand, Hohti (2000), studying data for Finland, reaches a conclusion more in line with Davis et al. (1996a), that there is no clear relationship between establishment size and net job creation.[15]

Davidsson et al. (1998) consider more directly whether the regression fallacy can account for the apparent important role of small firms in job creation, by attempting to estimate the bias from the regression to the mean that underlies the regression fallacy. Based on evidence from Swedish data, they suggest that the regression fallacy accounts for little distortion in the relationship between firm size and net job growth. They also – echoing concerns raised by other researchers – question drawing general implications regarding the behavior of small establishments or firms in job creation from plant-level data from the declining manufacturing sector in the United States.

2.2 Gibrat's Law

As Mansfield (1962) pointed out, there are three versions of Gibrat's law: the law of proportionate growth may apply to both surviving and exiting firms, only to surviving firms, or only to firms that exceed a minimum scale for efficiency. Clearly, if one cares about net job creation by firms of different sizes, like Birch did, the first version of the Gibrat's law is the most relevant one. In particular, if Gibrat's law holds for both surviving and exiting firms, Birch's claim that small firms create more jobs cannot be true. Sutton (1997) provides a comprehensive summary of this literature, concluding that the bulk of the empirical evidence suggests that Gibrat's law holds only for firms above a certain size limit. That is, findings from this literature do not contradict Birch's conclusion.

---

[15] Some research has also documented that, whereas small firms create a large share of new jobs, they are also responsible for a large share of job loss. For example, using data on manufacturing establishments from Germany, Wagner (1995) finds that although net job creation rates are not systematically related to firm size, gross job creation and destruction rates tend to be higher the smaller the firm. Broersma and Gautier (1997), Hohti (2000), and Barnes and Haskel (2002) also find that small firms have both higher job creation and higher job destruction rates than large ones. Similar results are reported by Gallagher et al. (1991) for the United Kingdom, for the period 1985-1987, with small firms contributing 48% of job creation but also 43% of job destruction, although they constituted only 21% of all employment in 1985. For this period, there is some evidence that small firms create more jobs on net, although the relationship is not very clear (Table IX).

Audretsch et al. (2004) point out that the empirical literature on Gibrat's law discussed in Sutton (1997) is almost entirely within the domain of manufacturing, echoing the concern regarding the limitation of Davis et al.'s findings to the manufacturing sector.[16] They argue that the findings based on the manufacturing sector do not necessarily apply to other industries. Gibrat originally postulated that the growth opportunity captured by an active firm is proportional to its size (i.e., expected growth rate conditional on survival is independent of firm size). This assumption implies that the law holds for all (both surviving and exiting) firms if the rate of survival is independent of firm size. However, ample empirical evidence shows that depending on the relative importance of entry cost, scale economies, industry growth, and capital intensity, the relationship between survival rate and firm size varies a great deal from one industry to another.[17] Because of this difference in survival, it is natural that small firms may be found to grow faster than larger ones in one industry but not in another when one examines surviving as well as exiting firms. Indeed, using data from the Dutch services sector, Audretsch et al. show that firm growth in services obeys Gibrat's Law, in contrast to evidence from manufacturing.[18] Audretsch et al. give a compelling reason that the research on the relationship between job creation and firm size must also pay attention to the potential differences between manufacturing and other sectors.[19] In particular, even if small firms in manufacturing do not create more net new jobs than large firms, one has no reason to believe that the same is true for other sectors.

3. The National Establishment Time Series Database

The database used in this study, the National Establishment Time Series (NETS), is a relatively new longitudinal file created by Walls & Associates using D&B data (Walls & Associates, 2003). We currently have access to an extract of this data set that covers all business establishments that were ever

---

[16] See, e.g., Dunne et al. (1989), Evans (1987), and Hall (1987), all of which use data from the manufacturing sector.
[17] See the long list of references cited in Audretsch et al. (2004).
[18] Audretsch et al. (2004) used a methodology originally developed by Chesher (1979). Let $z_{t,i}$ be the deviation of the logarithm of the size of firm i at time t from the mean of the logarithms of firm sizes at time t, and regress $z_{t,i}$ on $z_{t-1,i}$. If Gibrat's Law holds, $z_{t-1,i}$ should have a coefficient equal to 1, which can be tested statistically.
[19] Using data from Italy, Lotti (2007) compares the patterns of firm entry, growth, and survival in manufacturing and services. He finds that the patterns in the services sector are qualitatively similar to the stylized facts from the manufacturing sector, although there are some differences quantitatively. For example, the positive effect of firm size on firm survival is larger in the services sector.

located in California between 1989 and 2004, their respective parent headquarters (regardless of location), and establishments in other states that report to these headquarters.[20] The NETS database does not contain a rich set of information about each establishment, but it does include the business name, a unique D&B establishment identifier (the DUNS number), the establishment location, an SIC code in each year (which we have converted to a NAICS code), the type of location (single location, headquarters, branch) in each year, the identifier of the firm's headquarters, and employment (as well as sales, which we do not use).

The unit of observation in the NETS is a business establishment, which is a business or industrial unit at a single physical location that produces or distributes goods or provides services – for example, a single store or factory. Of course many firms own or control more than one establishment, and those establishments may be located in different geographic areas and may be engaged in different industries. The NETS data indicate whether an establishment is a stand-alone firm or a branch of a multi-establishment firm, and in the latter case establishments of the same firm can be linked. We sometimes refer to an establishment as a "business," reserving the word "firm" to refer to what may be collections of many establishments with a common owner.

One highly desirable feature of the NETS database is that it covers essentially all establishments. This reflects the fact that it is designed to capture the universe rather than a sample of establishments. Over the sample period of 1989-2004, the database includes information each year on between 1.2 and 1.9 million establishments in California providing about 15 million to 18 million jobs per year. In total, more than 24 million establishment-years are covered in our extract of the NETS database. However, we regard the data prior to 1992 as less reliable, because only beginning in 1992 was D&B able to purchase information on business units (from the Yellow Pages) from the regional Bells. We therefore use data covering 1992-2004.[21]

---

[20] Purchasing the NETS database for the entire United States was prohibitively expensive, costing $200,000 (for a two-year license).
[21] We have previously used the NETS data to study business relocation (Neumark et al., 2007) and the contribution of new firms to employment growth (Neumark et al., 2006). Other work uses the NETS data to study changes in the

The data construction effort – including both the cross-sectional files and the longitudinal linking that tracks establishments over time – is a massive and complicated one. Details are provided in Neumark et al. (2007). As previously noted, the earlier D&B data (from thirty years ago) that were used in Birch's work came in for considerable criticism. In particular, Davis et al. (1996a, 1996b) assert that these data were not particularly suitable for studying job creation-destruction dynamics for two reasons. First, early D&B data gave a much higher total employment level than alternative data sources such as those from Census or BLS. Second, the early D&B data did a poor job in terms of capturing new businesses, as documented in previous research that Davis et al. cite (and discussed further in Neumark et al., 2007). However, information technology has improved dramatically during the decades since then, and data sources have changed. Consequently the quality of the D&B data may have improved substantially. For these reasons, we have undertaken a good deal of investigation to document and examine the quality of the NETS data, in order to assess their reliability and their potential limitations, and how these limitations might affect results of various analyses.[22] We do not repeat the details here, except to focus on those issues highlighted in the earlier work that are most relevant to the analysis in this paper.

We assessed the reliability of the NETS data for measuring employment levels and changes by comparing the NETS with data from the Quarterly Census of Employment and Wages (QCEW), the Current Employment Statistics (payroll) survey (CES), and the Size of Business data (SOB). In these latter data sources we cannot identify establishments, but we can look at employment by industry, county, and size category. We find that the NETS does report higher aggregate employment levels, but that this appears to be mainly because it covers more very small businesses than alternative databases do. On the other hand, there is a fair amount of rounding of reported employment as well as imputation of missing data. As a result, the NETS data seem to show firm/establishment employment changes less frequently

---

geographic scope of firms' operations (Kolko and Neumark, forthcoming) and agglomeration economies in the high-tech sector (Wallace and Walls, 2004).

[22] See Neumark et al. (2007) as well as Kolko and Neumark (forthcoming).

than we would expect. As shown below, some of our analysis addresses possible biases from rounding and imputation of employment.[23]

It is important to track establishment births accurately in studying employment growth at small establishments. Two approaches were used to check the accuracy with which the NETS data capture new establishments: comparisons to new establishments identified from phonebooks in consecutive years, coupled with information elicited from these establishments on their actual start dates; and comparisons of the NETS data to a convenient listing of California biotech companies in the BioAbility database of U.S. biotech companies.[24] In both cases, Neumark et al. (2007) found that the NETS data identified establishment openings and their start dates very accurately.

Tracking births (and deaths) also requires distinguishing these events from relocations. For example, if an establishment first appears at an address in year t, it is important to know whether it is a new establishment or one that moved from elsewhere; in the latter case, this should not be considered a birth. Similarly, if an establishment moves to a new location, its absence at the old location should not be considered a death.[25] By comparing relocations in the NETS with those identified in Lexis-Nexis searches, we were able to confirm that the NETS captures inter-state relocations with considerable accuracy.

Finally, in this paper we focus on analysis at the establishment level, but also do some analysis at the firm level. This raises the question of how well the NETS links existing establishments and new establishments (births) to their parent firms. Assessing this requires another data source on the establishments belonging to a firm, including births of new establishments. Through some efforts, we were able to obtain such data for three large companies – a retail merchandiser, a manufacturer of computer equipment, and a restaurant chain. Overall, the tracking of firms' establishments works quite

---

[23] We doubt that any of the databases used in empirical research in this field are free of these rounding and imputation problems. Unfortunately, researchers have little information about the nature of the measurement errors created by rounding and imputation and know even less about how such measurement errors may bias the findings.
[24] This is a database of more than 2,000 U.S. biotech companies maintained by BioAbility, a biotech consulting firm. See http://www.bioability.com/us_biotech_companies.htm (viewed on September 14, 2005).
[25] These statements may depend on how the relevant "economy" is defined. For example, in some runs we focus only on establishments while they were in California, and treat moves out of the state as deaths and moves into the state as births. However, nearly all relocations are within state (Neumark et al., 2007).

well, although there are sometimes establishments from the alternative data sources that do not appear in the NETS, or vice versa. These "matching" errors appear to arise largely in the most recent years of the data set because the NETS sometimes detects new establishments with a delay.[26]

Overall, we find that the NETS database, like all others, is not without its flaws, and has both advantages and limitations. Its complete coverage of the whole economy annually is an advantage that we value most for our study here.[27] To answer the question whether small businesses create more jobs, we think it crucial to examine all sectors of the economy instead of just the manufacturing sector that most previous research has focused on. Earlier D&B data were criticized for reporting higher aggregate employment levels and capturing the births of businesses poorly. With the NETS database, which was constructed using more recent D&B data, we find that the former is not a serious concern and the latter is largely invalid.

4. Employer Size and Job Creation: Empirical Findings

We take two approaches to investigating whether small businesses create more jobs. First, following the tradition established by Birch (1987) and Davis et al. (1996a), we divide business establishments into different size categories and examine whether there is a significant difference in net job creation rates across size categories. Second, we non-parametrically estimate the relationship between employment growth and establishment size to avoid any potential biases or loss of information introduced by using arbitrary size boundaries.

4.1 Job creation and destruction by employer size

4.1.1 Measurement of employer size and job creation and destruction

It is common practice among researchers in this area to divide business establishments into different size categories and then investigate the variation in net job creation across size categories (see, e.g., Birch, 1987; Davis et al., 1996a, 1996b). Following this practice, we define the size of an establishment in the following two ways:

---

[26] For more details, see Kolko and Neumark (2007).
[27] For the purposes of research on other topics, advantages of the NETS data include the capture of businesses relocation within the U.S., the ability to disaggregate to a fine geographic level, as well as ease of access and the absence of confidentiality restrictions.

Base-year size: the size of the business establishment at the starting point of the interval over which employment growth is calculated. It is the definition used by Birch (1979, 1981, 1987), which was criticized by Davis et al. because it tends to overestimate the contribution of small businesses to job creation due to the "regression fallacy."

Average size: the simple average of business establishment size in base year ($t-1$) and current year (t). This definition was proposed by Davis et al. to avoid the regression fallacy.[28] Remember that the regression fallacy stems from the fact that some business establishments are categorized as small only due to random measurement errors or transitory employment declines that influence measured employment on a year-to-year basis and thus will regress to the mean in a short time. Using this average size measure, one will treat a business as small only if it stays small for at least two years. The underlying assumption is that random measurement errors and transitory employment declines do not last two years.

It is important to note that using this average size definition one can find a less important role of small businesses in job creation than using the base-year definition even if regression to the mean does not exist. Consider an example in which every business establishment grows with a positive rate and the measurement is always accurate. By construction, there will be no bias due to the regression fallacy in the calculation of job creation using the base-year size definition. However, this average size definition will still give a lower job creation rate among small businesses than the base-year size definition, because the fastest-growing businesses, even if they start small, will be less likely to be categorized as small under the average size definition. Therefore, although calculations based on average size tend to mitigate the potential biases from the regression fallacy, we should not interpret the difference between calculations based on these two size definitions as solely resulting from the regression fallacy.

Although the size definitions are discussed in terms of business establishments, they can also be applied to firms. One may argue that firm size is a more relevant measure to study than establishment size. Indeed, government regulations related to small businesses usually target small firms instead of

---

[28] As noted earlier, they call it "current" size.

small establishments.[29] Similarly, Gibrat's law is based on the idea that growth opportunities are

proportional to size – which may be more pertinent to firms than to establishments.[30] In this section, we

will examine job creation based on both establishment and firm sizes. For businesses in each size

category, we use the following measures to quantify their contributions to job creation, again following

Davis et al. (1996a):

Gross job creation: the sum of employment gains over all business establishments whose

employment level is greater than a year ago. Businesses just started are assumed to have zero

employment prior to opening, and thus this measure captures both job growth at expanding businesses

and job creation by business openings.

Gross job destruction: the sum of job loss over all establishments whose employment level is

smaller than a year ago. Businesses just closed are assumed to have zero employment after closing, and

thus this measure captures both job decline at contracting businesses and job loss due to business

closures.

Net job creation: the difference between gross job creation and gross job destruction. While this

net job creation is the most relevant measure, insofar as we care about growth in the number of jobs, it is

interesting to look at gross job creation and destruction to learn how the net growth is achieved. For

example, a 3% net job creation rate can be a result of either a 5% gross job creation rate net of a 2% gross

job destruction rate, or a 13% gross job creation rate net of a 10% gross job destruction rate. The latter

case involves much more reallocation of jobs among establishments, an issue of some importance in the

macroeconomic literature.[31]

---

[29] All federal agencies are required to use the business size standard developed by the U.S. Small Business
Administration to determine eligibility for federal contracts, programs, and regulations specifically designated for
small businesses. In this size standard, small business is defined either by number of employees or average annual
receipts and the calculation of size "includes the employees or receipts of all affiliates." See
http://www.sba.gov/services/contractingopportunities/sizestandardstopics/indexguide/index.html (viewed October 7,
2007). The U.S. Bureau of Labor Statistics (BLS) also uses firm size when it publishes statistics on employment
dynamics by employer size (Butani et al., 2006).
[30] Although Gibrat's work was motivated by the size distribution of manufacturing plants (establishments), later
empirical work by others was conducted at both establishment and firm levels (Sutton, 1997, p. 41 and 46).
[31] These gross job creation and destruction statistics calculated at the establishment level still underestimate *worker*
reallocation. For example, within an establishment some jobs (e.g., assemblers) may be destroyed and some others
(e.g., technicians) may be added, with no change in overall employment. Studying worker reallocation requires the

Note that these calculations are always conducted at the establishment level although size category may be defined at the establishment level or the firm level. At the firm level, the firm exhibits either overall net job creation or overall net job destruction. But studied at the establishment level, a firm may contribute to both job creation (at expanding and new establishments) and job destruction (at contracting and closing establishments).

4.1.2 Main results

The first two panels in Table 1 show the annual rates of gross job creation, gross job destruction, and net job creation by establishment size for all industries, using the base-year and average size definitions. The last column shows the employment share of establishments in different size categories. When the base-year size is used, we find that the rate of gross job creation decreases monotonically, and sharply, with establishment size. While the average annual rate of job creation in the 0-19 category is 24.0%, it is only 0.4% in the "5,000 or more" category. In contrast, the rate of gross job destruction has much lower variation (between 7.9 and 12.2%) and shows no clear relationship to establishment size. As the difference between these two, the net job creation rate also clearly decreases with establishment size. Total jobs at establishments with 19 or fewer employees grow by 13.4% annually. Establishments in all other size categories show a negative net job creation rate. Overall, the findings using the base-year size definition clearly suggest an important role of small businesses in job creation, consistent with Birch's (1979, 1981, and 1987) findings using the same method.

The second panel in Table 1 presents the results using the average size definition. Under this size definition, the rate of net job creation for small establishments falls dramatically to 1.4% (from 13.4%). The same rate for the largest establishments, on the other hand, becomes much higher (−2.1%, in contrast to the −11.4% calculated using the base-year size definition). However, overall there is still a clear

---

ability to track individual workers, as in the Longitudinal Employer-Household Dynamics data (http://lehd.did.census.gov/led/, viewed October 7, 2007).

negative relationship between the net job creation rate and establishment size.[32] In fact, only

establishments with fewer than 250 employees create jobs. And under either definition, all establishments

in higher size categories lose jobs. The role of business establishments with fewer than 250 employees in

overall job creation is even more significant given that these establishments also constitute a large share

of total employment (75.1%). Together, these establishments account for three-fourths of total

employment, which is true under either size definition.

Comparing Panels I and II in Table 1, we do see evidence consistent with the regression fallacy.

In particular, the base-year size definition tends to overestimate gross job creation by the smallest

establishments and to overestimate gross job destruction by large establishments. And, therefore, it

overestimates net job creation by small establishments and underestimates net job creation by large

establishments. In this sense, for these data the concern of Davis et al. over Birch's methodology is well

justified. However, our results in Panel II suggest that Birch's conclusion is still valid, though to a lesser

extent than is suggested by his methods.

We repeated the calculation of job creation and destruction rates based on firm instead of

establishment size and the results are shown in Panels III and IV of Table 1. These results reveal a

similar pattern as those in the first two panels: Under the base-year size definition, we see small firms

create substantially more jobs, on net, than large firms. This difference becomes less dramatic once we

use the average size definition, but there is still clear evidence that small firms have higher net job

creation rates.

Table 2 translates the gross job creation and destruction rates into shares of overall job creation

and destruction attributable to each size class. The table shows that establishments in the 0-19 category

have a higher share in both gross job creation and gross job destruction than its employment share implies

– but more so for job creation. Under the base-year size definition, establishments with 19 or fewer

employees account for 33.2% of total employment, but their share is 79.3% in gross job creation and

---

[32] Relative to the differences in the estimates under the base-year size definition, these differences in growth rates
among average size categories seem small. However, since these are annual growth rates, the differences are still
quite substantial.

37.4% in gross job destruction. Similarly, under the average size definition, these smallest establishments account for 34.1% of total employment, but represent 47.8% of gross job creation and 45.7% of gross job destruction. Thus job reallocation at small establishments tends to be much higher. Panels III and IV of Table 2 show similar results at the firm level.

The results in Tables 1 and 2 are different from the findings in Davis et al. (1996a, 1996b). In fact, Davis et al. find that once they use the average size definition instead of the base-year size definition, the important role of small businesses in job creation as emphasized by Birch (1979, 1981, and 1987) completely disappears, and there is instead either no relationship or a positive relationship between net job creation and establishment size.[33] In contrast, we still find that small businesses create more jobs even if the average size definition is used, although the results are much less striking. Therefore, these findings point to the importance of the regression fallacy in estimating the size-growth relationship. However, there still seems to be some truth to Birch's original conclusion regarding the important role of small businesses in job creation.

4.1.3 Sensitivity analysis

We have conducted a series of sensitivity analyses to assess the robustness of the main results. The results of these analyses are shown in the two panels of Table 3. For comparison purposes, the first column of Panel I shows the baseline result of net job creation rates for all establishments in California across average establishment size categories, which simply reproduces the third column of Panel II in Table 1.

The calculations in Table 1 (and Table 2) used all establishments in California, and treated interstate moves as births or deaths. In fact, however, some establishments originally started in other U.S. states and moved into California later, and others started in California but moved to other states, and we may want to include these establishments in the population we study.[34] It is quite possible that these

---

[33] See their Table 4.2. Just to clarify, this positive relationship emerges for their "current" size measure, which we label "average establishment size." The correlation between net job growth and the ranking of this size category is 0.75, and the correlation between net job growth and the midpoint of the size category (using 7,500 for the 5,000 or more category) is 0.46.

[34] However, interstate relocation is very minor (Neumark et al., 2007).

movers differ significantly from other establishments in terms of their survival chances or prospects for growth. Because we have information on establishments that moved, we can check whether the inclusion or exclusion of these movers significantly affects our baseline result. The second column of Panel I in Table 3 is calculated using all establishments that were ever located in California, including those that moved out. The third column uses establishments that were always in California, excluding both those that moved out of and those that moved into California. The results differ only negligibly from the baseline results. It is still clear, using either alternative sample, that small establishments show higher net job creation rates than do large ones.

Since Davis et al. used the LRD data, which cover only the manufacturing sector, their results could be unique to manufacturing. We therefore replicated our baseline calculations using the data for the manufacturing sector only; for comparison purposes, we also did this for the services sector. The results are shown in the fourth and fifth columns of Panel I in Table 3. Interestingly, the results for the manufacturing sector are different from the baseline results using all industries. First, manufacturing establishments in all size categories lose jobs, which contrasts with the baseline result that only establishments with 250 or more employees lose jobs, but is consistent with the fact that manufacturing has been a declining sector. Second, although manufacturing establishments in all size categories lose jobs, we do not see larger establishments losing a higher proportion of jobs. Manufacturing establishments in the two smallest size categories lose 1.1% of their jobs; the net job destruction rate is smaller for the next three size categories; it then jumps up in the next two size categories with establishments in the 1,000-2,499 group losing the most jobs, 3.7% a year; and finally, it comes down in the next category and goes up again in the largest size category. It is reasonable to say that under the average size definition there appears to be no clear relationship between net job creation and establishment size. In other words, even if Davis et al. were studying the manufacturing sector using the NETS data, they would likely have come to the same conclusion. That is, Birch's original conclusion is untenable under a different (and more reasonable according to Davis et al.) size definition in the

manufacturing sector, even in this quite different time period.[35] In contrast, the results for the services

sector in the last column are similar to the baseline results for all industries, in that small establishments

have higher net job creation rates. Therefore, these findings suggest that the conclusions that Davis et al.

reached may be unique to the manufacturing sector and cannot be generalized to the whole economy.

Given the ongoing declining share of the manufacturing sector in the U.S. economy, we believe that this

is a significant result.

As mentioned above and documented in Neumark et al. (2007), a considerable share of

employment data in the NETS are not actual but imputed. The NETS database always indicates whether

a particular employment number is imputed or not, making it possible to check whether our main result is

robust to the use of imputed employment data. We redid the calculations dropping establishments with

imputed employment data. This is carried out in four different ways: (1) dropping an establishment if its

employment is imputed in every year; (2) dropping an establishment if its employment is imputed in more

than 50% of years; (3) dropping an establishment if its employment is imputed in more than 25% of

years; and (4) dropping an establishment if its employment is imputed in any year. The results are

presented in the first four columns of Panel II of Table 3.

Overall, the less stringent the data exclusion criterion used, the more the results resemble the

baseline results. When we drop the establishments with all or more than half imputed employment data,

the results still support the conclusion that small establishments have higher net job creation rates. When

we exclude the establishments with more than one-quarter imputed employment data (the third column of

Panel II), the relationship is a bit less clear. In particular, the 0-19 size category has a lower net job

creation rate than the next three size categories, and it is negative. However, in four of the five largest

size categories the net job creation rate is lower still. When establishments with any imputed data are

dropped, however, there is no longer an indication that net job creation is higher at smaller

establishments. Clearly these last two analyses use two samples that are very restrictive: the first drops

57% of the establishments and the second drops 74% of the establishments. It is thus rather reassuring to

---

[35] Of course, our data are also limited to California.

see that the main result does not disappear completely even in the former case. It is also not entirely

surprising that the results are affected by the stringent criterion of dropping observations with any

imputed data; Neumark et al. (2007) show that imputation is most common for early observations on new

establishments, so that imposing a restriction of *no* imputed data might generate a non-random sample

that, for example, excludes small new establishments in certain industries that tend not to be picked up

right away by D&B.

As also documented in Neumark et al. (2007), the distribution of employment numbers in the

NETS data tends to be disproportionally concentrated on multiples of 10 or 100. We consider this a sign

that some of the employment data have been rounded. Naturally, we are concerned that the higher net job

creation rates of small establishments may be an artifact of rounded employment data, although there is

no obvious reason why rounding would create bias in this direction.

Although we do not report all the results here, we have investigated the problem of rounded

employment data in two ways. First, we drop the establishments that seem most likely to have rounded

employment numbers. Specifically, we exclude an establishment from our calculation if its employment

is a multiple of 10 or 100 in every year. The first results (based on multiples of 10) are shown in the last

column of Panel II of Table 3. The results are quite similar to the baseline, showing a monotonically

decreasing rate of net job creation with establishment size.[36] Second, instead of dropping establishments

that may be reporting rounded data, we instead consider ways in which the employment data might have

been rounded, and synthetically round all employment data in the same way. We then recalculate net job

creation rates across establishment size groups using the synthetic data. The idea is that if our results are

affected by systematic differences in which establishments have rounded employment data, then the

results from the synthetic data are likely to be significantly different from our baseline results. The

results, however, were insensitive to this synthetic rounding.[37] In all cases, the results still indicated that

---

[36] Dropping any establishment if its employment was always a multiple of 100 resulted in losing very few observations relative to the baseline, and the results were correspondingly very similar.
[37] We synthetically rounded the employment data in four different ways: (1) round every employment number to the nearest multiple of 10; (2) round every employment number *down* to the nearest multiple of 10; (3) round every employment number *up* to the nearest multiple of 10; and (4) round every employment number to the nearest

small establishments have higher net job creation rates, and as in the baseline results, only establishments with fewer than 250 employees were net job creators while establishments in all larger groups were estimated to destroy jobs on net.

4.2 Job creation and employer size: non-parametric estimation

In the previous section, we followed existing research by dividing business establishments into different size categories to examine the relationship between the net job creation rate and employer size. The choice of the size boundaries – while consistent with the ones used in the existing literature – is nevertheless arbitrary.[38] The particular employer size-class boundaries that are chosen could obscure the relationship between employer size and the employment growth rate. In order to check the robustness of this relationship, and to offer point estimates for the complete range of establishment sizes, we estimate the employment growth rate for every establishment size using a non-parametric locally-weighted regression.[39] Since the "regression fallacy" criticism also applies to these regression analyses, we still use the average establishment size instead of the base-year size as our independent variable.

Locally-weighted regressions are computationally intensive because the amount of calculations rises exponentially with the number of observations. Our large amount of data (over 20 million observations) would require performing over 20 million weighted mean calculations, each of which would use a substantial proportion of the total observations (depending on the particular bandwidth chosen). It is not feasible for us to do this. Instead, we devised a less computationally-intensive method. We first create a mean net job creation rate and an associated frequency weight for each unique value of establishment size. This reduces the number of observations enormously while still maintaining the relevant information. Then, we compute the locally-weighted smoothed means using weights that are a function of the locality and frequency of each observation. We use a 0.1 (10%) bandwidth kernel and

---

multiple of 10 if it is under 100, and to the nearest multiple of 100 otherwise. All of these methods except the third resulted in substantially fewer observations because of small establishments that had their data rounded to zero.

[38] Davis et al. (1996a) note that the choice of size-class boundaries is one of the factors that influence the magnitude of the bias associated with the regression fallacy (p. 68).

[39] Many researchers have tested Gibrat's law based on regressions of firm growth on firm size instead of a definition of firm-size boundaries. See, for example, Chesher (1979) and Audretsch et al. (2004). Farinas and Moreno (2000) use a non-parametric approach that is similar to ours, although there are some differences in the estimation, and their model is estimated using panel data on only about 2,200 manufacturing firms in Spain over a 5-year period.

frequency-adjusted tricubic weights. This estimation procedure is described in greater detail in the Appendix. While this method is not computationally equivalent to the standard approach, such smoothing would otherwise be infeasible.

To visualize the nonparametric regression results, we plot the estimated mean net job creation rate against average establishment size. Note that the number of jobs and its change are necessarily integers; the smallest unit is one. When the establishment size is close to zero, even a one-unit change (the smallest possible non-zero growth/decline) in total employment amounts to very large growth or decline in percentage terms. Therefore, the variance of the job growth rate is many times larger when establishment size is close to zero. Although this is essentially an artifact of the data at the individual establishment level, it obscures the picture of the relationship between job growth rate and establishment size. As a result, if we plot the relationship over the whole range of establishment sizes, it always appears that the change in job growth rate is very large at the smallest sizes and indiscernible beyond that. To show the relationship more clearly, in each figure we present the results in two panels: Panel I depicts the relationship over the whole range of establishment sizes and Panel II focuses on the relationship in the range of 5-2,000 employees.

As Figure 1 shows, the net job creation rate is clearly negatively correlated with employer size, meaning that smaller business establishments have higher net job creation rates.[40] We see, more clearly in Panel II, that the smallest establishments have a positive net job creation rate. This rate decreases as establishment size increases. At a size close to 450 employees, the net job creation becomes zero. It then becomes negative and stays negative for all larger establishment sizes.

A numerical example may help illustrate the economic significance of the difference in job creation rates. Panel II of Figure 1 suggests that an establishment with about 100 employees has a net job creation rate close to 0.4%, whereas an establishment with about 1,000 employees has a net job creation rate close to −0.8%. Consider two regions that each begin with employment of 100,000, one with 1,000 smaller establishments each of which has 100 employees, and the other with 100 larger establishments

---

[40] In this figure we otherwise use the data the same way as in Table 1.

21

each of which has 1,000 employees. In a decade, the region with the smaller establishments is expected to create 4,100 jobs; the region with the larger establishments is expected to lose 7,700 jobs, leading to a net difference of 11,800 jobs after a decade.[41] The smallest establishments (in the 0-19 category), of course, have much higher growth rates. However, as they grow bigger their growth rates are expected to decline sharply.

Parallel to the results presented in Tables 2 and 3, Figures 2 and 3 show the results from sensitivity analyses. Figure 2 plots the relationship between net job creation rate and average establishment size estimated using different samples. We first used all establishments that were ever located in California, including those that started in California but later moved to other states and those that started in other states but later moved into California (without treating these moves as births or deaths). We then used the sample of establishments that were always in California, excluding all establishments that moved into or out of California. Compared to those shown in Figure 1, the results in Figure 2 using these two alternative samples are almost identical. We also estimated the relationship using the manufacturing and services sectors separately. The differences between these two subsamples are clear. The services sector results are almost indistinguishable from the results using the whole sample, but the manufacturing sector is very different. First, manufacturing establishments show a consistently lower net job creation rate over the whole range of sizes than the overall sample, and the difference seems always greater than 0.5% a year. Again, this is consistent with the fact that the manufacturing sector was declining. Second, and perhaps more important, the negative relationship between net job creation rate and establishment size is much less clear in the manufacturing sector. The net job creation rate shows a few ups and downs in the range of 5-500 employees, although it is clearly declining in the range of 500-2,000 employees. Based on these results, one may still argue that the manufacturing sector gives no clear evidence that small establishments create more jobs, consistent with what Davis et al. (1996a, 1996b) found. Yet again, we emphasize the differences between manufacturing and the overall economy and

---

[41] These are based on calculations that assume a constant employment growth rate over a relatively small size range: $1,000 \cdot 100 \cdot (1 + 0.004)^{10} = 104,100$ and $1,000 \cdot 100 (1 - 0.008)^{10} = 92,300$.

point out that, overall, small establishments do have higher net job creation rates, which is more in line with the findings of Birch (1979, 1981, and 1987).

Finally, Figure 3 shows the results with the various restrictions on establishments with imputed data, as well as the sample dropping establishments that clearly had data rounded to multiples of 10. Except for the two cases of the data that exclude very large numbers of observations, as discussed earlier, the figure indicates that the results are robust to various treatments of imputed and rounded data in pointing to declining net job creation with establishment size.[42]

5. Conclusions

It has long been debated whether small businesses create more jobs than larger ones. The debate attracts a great deal of attention both in academia and in the media because it has important implications for policymaking at both federal and state levels. David Birch's book (1987) was probably the most prominent catalyst for this debate, arguing that small businesses play a much more important role in job creation than larger businesses – a view that quickly became the conventional wisdom. Yet this work was criticized, most notably by Davis et al. (1996a), who argued that Birch's findings are essentially a result of a flawed methodology applied to some unsuitable data. This debate also speaks to earlier (and continuing) research in industrial organization on Gibrat's Law, which posits that firm growth rates are independent of firm size.

Using newly-available data on businesses establishments in California, we revisit the issue of establishment and firm size and job creation. The data are drawn from the NETS database, which was recently constructed using raw data collected by D&B. Our evaluation of the NETS data in separate work suggests that these data are rather reliable, and a clear improvement upon early versions of the D&B data used by David Birch in the 1970s and 1980s. We analyze these data using two methods: one based on size classes but following the methods of Davis et al. to avoid the regression fallacy that they argue plagued Birch's work; and the other a nonparametric method that does not need to define arbitrary employer size categories.

---

[42] Figures for the other methods of treating rounded data, discussed in sub-section 4.1.3 are not shown. These figures similarly demonstrated the robustness of the results, and are available from the authors upon request.

Like Birch, we find that small establishments (and small firms as well) create more jobs. At the same time, we confirmed that the concern of Davis et al. regarding Birch's methods is well grounded, and the base-year size definition used by Birch indeed badly overestimates net job creation rates for small establishments. However, the result that small establishments create more jobs, on net, survives even after we use the methodology proposed by Davis et al. to avoid the regression fallacy. Although we still find that small establishments create more jobs, the difference is much smaller than that originally suggested by Birch. We find similar results at the firm level – results that are inconsistent with Gibrat's Law.

We also find that the negative relationship between establishment size and job creation is much less clear for the manufacturing sector. This likely explains why Davis et al. (1996a) conclude that the difference in job creation between small and large firms highlighted by Birch is nonexistent, as they used the LRD data, which cover the manufacturing sector only. The NETS data (for a later period, and for California) yield similar results for manufacturing. However, the overall economy, unlike the manufacturing sector, does show a higher net job creation rate among small establishments.

References

Acs, Zoltan J. and David B. Audretsch, "Editors' introduction," *Small Business Economics* 1, 1989, 1-5.

Audretsch, David B., Luuk Klomp, Enrico Santarelli, and A. Roy Thurik, "Gibrat's Law: Are the Services Different?," *Review of Industrial Organization* 24, 2004, 301–324.

Audretsch, David B. and Roy Thurik, "What's New about the New Economy? Sources of Growth in the Managed and Entrepreneurial Economies," *Industrial and Corporate Change* 10, 2001, 267-315.

Baldwin, John and Garnett Picot, "Employment Generation by Small Producers in the Canadian Manufacturing Sector," *Small Business Economics* 7, 1995, 317-331.

Barnes, Matthew and Jonathan Haskel, "Job Creation, Job Destruction and the Contribution of Small Businesses: Evidence for UK Manufacturing," Working Paper No. 461, Queen Mary, University of London, 2002.

Birch, David L., *The Job Generation Process*, Unpublished report prepared by the MIT Program on Neighborhood and Regional Change for the Economic Development Administration, U.S. Department of Commerce, Washington, D.C., 1979.

Birch, David L., "Who Creates Jobs?" *The Public Interest* 65, 1981, 3-14.

Birch, David L., *Job Creation in America: How Our Smallest Companies Put the Most People to Work*, Free Press, New York, 1987.

Broersma, Lourens and Pieter Gautier, "Job Creation and Job Destruction by Small Firms: An Empirical Investigation for the Dutch Manufacturing Sector," *Small Business Economics* 9, 1997, 211-224.

Brown, Charles, James Hamilton, and James Medoff, *Employers Large and Small*, Harvard University Press, Cambridge, MA, 1990.

Butani, Shail J., Richard L. Clayton, Vinod Kapani, James R. Spletzer, David M. Talan, and George S. Werking Jr., "Business Employment Dynamics: Tabulations by Employer Size," *Monthly Labor Review* 129, 2006, 3-22.

Carree, Martin and Luuk Klomp, "Small business and job creation: A comment" *Small Business Economics* 8, 1996, 317-322.

Chesher, Andrew, "Testing the Law of Proportionate Effect," *Journal of Industrial Economics* 27, 1979, 403-411.

Cleveland, William S., "Robust Locally Weighted Regression and Smoothing Scatterplots," *Journal of the American Statistical Association* 74, 1979, 829-836.

Davis, Steven J., John C. Haltiwanger, and Scott Schuh, *Job Creation and Destruction*, The MIT Press, Cambridge, MA, 1996a.

Davis, Steven J., John Haltiwanger, and Scott Schuh, "Small Business and Job Creation: Dissecting the Myth and Reassessing the Facts," *Small Business Economics* 8, 1996b, 297-315.

Dunne, Timothy, Mark Roberts, and Larry Samuelson, "The Growth and Failures of U.S. Manufacturing Plants," *Quarterly Journal of Economics* 104, 1989, 671-698.

Evans, David S., "The Relationship between Firm Growth, Size, and Age: Estimates for 100 Manufacturing Industries," *Journal of Industrial Economics* 35, 1987, 567-581.

Farinas, Jose C., and Lourdes Moreno, "Firms' Growth, Size, and Age: A Nonparametric Approach," *Review of Industrial Organization* 17, 2000, 249-265.

Gallagher, Colin C., Daly, Michael J., and Thomason, Jeremy C., "The Growth of U.K. Companies and Their Contribution to Job Generation, 1985-87." *Small Business Economics* 3, 1991, 269-286.

Gibrat, Robert, *Les Inégalités Économiques*, Librairie du Recueil Sirey, Paris, 1931.

Hall, Bronwyn H., "The Relationship between Firm Size and Firm Growth in the US Manufacturing Sector," *Journal of Industrial Economics* 35, 1987, 583-606.

Heshmati, Almas, "On the Growth of Micro and Small Firms: Evidence from Sweden," *Small Business Economics* 17, 2001, 213–28.

Hohti, Satu, "Job Flows and Job Quality by Establishment Size in the Finnish Manufacturing Sector 1980–94," *Small Business Economics* 15, 2000, 265–81.

Kirchhoff, Bruce and Bruce Phillips, "The Effect of Firm Formation and Growth on Job Creation in the United States," *Journal of Business Venturing* 3, 1988, 261-272.

Kirchhoff, Bruce A. and Patricia G. Greene, "Understanding the Theoretical and Empirical Content of Critiques of U.S. Job Creation Research," *Small Business Economics* 10, 1998, 154-169.

Klepper, Steven and Elizabeth Graddy, "The Evolution of New Industries and the Determinants of Market Structure," *RAND Journal of Economics* 21, 1990, 27-44.

Kolko, Jed, and David Neumark, *Business Location Decisions and Employment Dynamics in California*, Public Policy Institute of California, San Francisco, 2007.

Kolko, Jed, and David Neumark, "Changes in the Location of Employment and Ownership: Evidence from California," forthcoming in *Journal of Regional Science*.

Leonard, Jonathan S., "In the Wrong Places at the Wrong Time: The Extent of Frictional and Structural Unemployment," in Kevin Lang and Jonathan S. Leonard (eds.), *Unemployment and the Structure of Labor Markets*, Blackwell, 1987, 141-163.

Lotti, Francesca, "Firm Dynamics in Manufacturing and Services: A Broken Mirror?" *Industrial and Corporate Change* 16, 2007, 347-369.

Neumark, David, Junfu Zhang, and Brandon Wall, "Where the Jobs Are: Business Dynamics and Employment Growth," *Academy of Management Perspectives* 20, 2006, 79-94.

Neumark, David, Junfu Zhang, and Brandon Wall, "Employment Dynamics and Business Relocation: New Evidence from the National Establishment Time Series," *Research in Labor Economics* 26, 2007, 39-83.

Pivetz, Timothy R., Michael A. Searson, and James R. Spletzer, "Measuring Job and Establishment Flows with BLS Longitudinal Microdata," *Monthly Labor Review*, 2001, 13-20.

Sachs, Jeffrey D. and Howard J. Shatz, "Trade and Jobs in U.S. Manufacturing," *Brookings Papers on Economic Activity*, 1994, 1-84.

Sutton, John, "Gibrat's Legacy," *Journal of Economic Literature* 35, 1997, 40-59.

Voulgaris, Fotini, Theodore Papadogonas, and George Agiomirgianakis, "Job Creation and Job Destruction in Greek Manufacturing," *Review of Development Economics* 9, 2005, 289–301.

Wagner, Joachim, "Firm Size and Job Creation in Germany," *Small Business Economics* 7, 1995, 469-474.

Wallace, Nancy E., and Donald Walls, "Agglomeration Economies and the High-Tech Computer," Fisher Center Working Paper No. 292, University of California, Berkeley, 2004.

Walls & Associates, "NETS: National Establishment Time-Series Database," Oakland, California, 2003.

Appendix: Technical Notes for the Non-parametric Analysis

Our non-parametric analysis uses a method that is a slight modification of the locally-weighted scatterplot smoothing (LOWESS) method originally proposed by Cleveland (1979).

Let y_i be the employment growth rate of observation i (an establishment over a two-year period), x_i the size of observation i measured using the average size definition, and N the total number of observations. The standard implementation of locally-weighted mean smoothing would proceed as follows. Order the data such that $x_i \leq x_{i+1}$ for all $i = 1, \ldots, N-1$. For each y_i, choose the subset of the data that is indexed by $i_- = \max(1, i - k)$ through $i_+ = \min(i + k, N)$, where $k = [(N \cdot h - 0.5)/2]$ and h is the prespecified bandwidth that indicates the proportion of the data used in the calculations of the smoothed value $\hat{y}_i$. Choose a function that assigns a weight w_i to each observation $j = i_-, \ldots, i_+$; observations outside of this range are given no weight. For example, one may choose a tricubic weight function (the kernel), in which case the smoothed value $\hat{y}_i$ is calculated as:

$$\hat{y}_i = \frac{\sum_{j=i_-}^{i_+} \left(w_j \cdot y_j \right)}{\sum_{j=i_-}^{i_+} w_j}, \text{ where } w_j = \left\{ 1 - \left(\frac{\left| x_j - x_i \right|}{\Delta} \right)^3 \right\}^3 \text{ and } \Delta = 1.0001 \cdot \max(x_{i+} - x_i, x_i - x_{i-}).$$

Given the large size of our dataset, this method is computationally infeasible. It would involve making approximately 20 million weighted calculations each using the share of observations determined by the specified bandwidth. Instead, we utilize the following method where we first compute an average value y for each unique value of x, and then calculate a smoothed value $\hat{y}_{i'}$ from the reduced dataset.

We use the following procedure. First, order the data such that $x_i \leq x_{i+1}$ for all $i = 1, \ldots, N-1$. For each unique value of x_i, create a $z_{i'} = x_i$. Let the total number of $z_{i'}$ be M and order all of them such that $z_{i'} < z_{i'+1}$ for all $i' = 1, \ldots, M-1$. Then, let $y_{i'} = \dfrac{\sum_{y_i \in \widetilde{Y}_{i'}} y_i}{C(\widetilde{Y}_{i'})}$ for all $i' = 1, \ldots, M$, where $\widetilde{Y}_{i'} = \{ (y_i, x_i) : x_i = z_{i'} \}$

and $C(\widetilde{Y}_{i'})$ is the cardinality of $\widetilde{Y}_{i'}$. Now apply the standard smoothing procedure to observations $(y_{i'}, z_{i'})$,

except that the weight function is adjusted using the frequency of $y_{i'}$. Again using a tricubic weight function, this amounts to calculating the following smoothed value:

$$\hat{y}_{i'} = \frac{\sum_{j'=i'_-}^{i'_+}\left(w_{j'}\cdot y_{j'}\right)}{\sum_{j'=i'_-}^{i'_+}w_{j'}} \qquad \text{where } w_{j'} = C(\widetilde{Y}_{i'})\cdot\left\{1-\left(\frac{\left|z_{j'}-z_{i'}\right|}{\Delta}\right)^3\right\}^3 .$$

These two methods essentially use the same information in the data although they usually assign slightly different smoothed values to different observations. Whereas the standard method gives multiple predicted values for each $z_{i'}$ in cases where there are multiple x_i such that $x_i = z_{i'}$, our method only returns one predicted value per unique value of x.[43]

---

[43] As a consequence, another advantage of this procedure, aside from shrinking the data set, is that it avoids the arbitrariness in the treatment of ties in LOWESS estimation.

TABLE 1
Rates of job creation and destruction by employer size, 1992-2004
(All industries, excluding government*)

Size class (employees)	Gross job creation	Gross job destruction	Net job creation	Employment share
I. Base-year establishment size**				
0 to 19	24.0%	10.6%	13.4%	33.2%
20 to 49	4.5%	7.9%	−3.5%	15.2%
50 to 99	3.6%	7.9%	−4.2%	12.3%
100 to 249	3.3%	8.3%	−5.0%	13.7%
250 to 499	2.6%	9.4%	−6.8%	7.3%
500 to 999	1.9%	10.2%	−8.3%	5.5%
1,000 to 2,499	1.5%	11.8%	−10.2%	6.6%
2,500 to 4,999	1.0%	12.2%	−11.2%	3.4%
5,000 or more	0.4%	11.8%	−11.4%	2.9%
II. Average establishment size***				
0 to 19	14.0%	12.6%	1.4%	34.1%
20 to 49	9.0%	8.1%	0.8%	15.2%
50 to 99	7.8%	7.1%	0.7%	12.3%
100 to 249	7.8%	7.4%	0.4%	13.5%
250 to 499	7.8%	8.2%	−0.4%	7.2%
500 to 999	9.1%	10.3%	−1.2%	5.5%
1,000 to 2,499	7.5%	10.0%	−2.5%	6.4%
2,500 to 4,999	7.1%	8.3%	−1.2%	3.2%
5,000 or more	3.7%	5.7%	−2.1%	2.7%
III. Base-year firm size**				
0 to 19	32.5%	10.4%	22.1%	8.5%
20 to 49	12.2%	7.7%	4.5%	3.6%
50 to 99	13.2%	8.1%	5.1%	3.1%
100 to 249	15.5%	9.2%	6.3%	4.2%
250 to 499	16.3%	10.1%	6.2%	3.5%
500 to 999	15.4%	10.7%	4.7%	4.2%
1,000 to 2,499	13.9%	11.1%	2.8%	7.4%
2,500 to 4,999	10.3%	11.9%	-1.5%	7.1%
5,000 to 9,999	9.4%	11.9%	-2.5%	8.4%
10,000 to 24,999	7.7%	11.4%	-3.6%	13.1%
25,000 to 49,999	6.3%	10.5%	-4.3%	11.3%
50,000 or more	4.5%	7.5%	-3.0%	25.5%
IV. Average firm size***				
0 to 19	15.1%	11.9%	3.2%	8.6%
20 to 49	10.5%	7.8%	2.7%	3.6%
50 to 99	9.6%	7.6%	2.0%	3.1%
100 to 249	10.2%	8.6%	1.6%	4.2%
250 to 499	10.2%	9.1%	1.0%	3.5%
500 to 999	10.4%	9.3%	1.0%	4.2%
1,000 to 2,499	9.9%	9.4%	0.4%	7.5%
2,500 to 4,999	9.3%	9.5%	−0.2%	7.2%
5,000 to 9,999	9.2%	10.0%	−0.9%	8.5%
10,000 to 24,999	8.5%	9.8%	−1.4%	13.0%
25,000 to 49,999	7.9%	9.6%	−1.8%	11.3%
50,000 or more	7.5%	9.0%	−1.6%	25.3%

* Establishment calculations use all observations in California and treat interstate moves as births or deaths; firm calculations use all branches of firms with at least one establishment ever located in California.

** Equal to the number of employees in the initial year of the interval over which the growth rate is calculated.

*** Average business size for base year (t−1) and current year (t); used to avoid regression fallacy described by Davis et al. (1996a).

TABLE 2
Shares of gross job creation and destruction by employer size, 1992-2004
(All industries, excluding government*)

Size class (employees)	Job creation share	Job destruction share	Employment share
I. Base-year establishment size**			
0 to 19 employees	79.3%	37.4%	33.2%
20 to 49	6.9%	12.7%	15.2%
50 to 99	4.6%	10.2%	12.3%
100 to 249	4.7%	11.8%	13.7%
250 to 499	1.9%	7.1%	7.3%
500 to 999	1.1%	5.8%	5.5%
1,000 to 2,499	1.0%	7.7%	6.6%
2,500 to 4,999	0.3%	4.0%	3.4%
5,000 or more	0.1%	3.3%	2.9%
II. Average establishment size***			
0 to 19 employees	47.8%	45.7%	34.1%
20 to 49	13.8%	13.1%	15.2%
50 to 99	9.5%	9.0%	12.3%
100 to 249	10.6%	10.5%	13.5%
250 to 499	5.6%	6.1%	7.2%
500 to 999	5.0%	5.7%	5.5%
1,000 to 2,499	4.7%	6.2%	6.4%
2,500 to 4,999	2.1%	2.4%	3.2%
5,000 or more	0.9%	1.4%	2.7%
III. Base-year firm size**			
0 to 19 employees	25.7%	9.5%	8.5%
20 to 49	4.1%	3.0%	3.6%
50 to 99	3.7%	2.6%	3.1%
100 to 249	6.0%	4.0%	4.2%
250 to 499	5.1%	3.7%	3.5%
500 to 999	6.0%	4.6%	4.2%
1,000 to 2,499	9.7%	8.4%	7.4%
2,500 to 4,999	6.8%	8.6%	7.1%
5,000 to 9,999	7.2%	10.0%	8.4%
10,000 to 24,999	9.0%	14.8%	13.1%
25,000 to 49,999	6.3%	11.8%	11.3%
50,000 or more	10.3%	19.0%	25.5%
IV. Average firm size***			
0 to 19 employees	14.2%	11.7%	8.6%
20 to 49	4.2%	3.2%	3.6%
50 to 99	3.2%	2.5%	3.1%
100 to 249	4.6%	3.9%	4.2%
250 to 499	3.8%	3.5%	3.5%
500 to 999	4.6%	4.1%	4.2%
1,000 to 2,499	7.9%	7.5%	7.5%
2,500 to 4,999	7.2%	7.2%	7.2%
5,000 to 9,999	8.5%	8.9%	8.5%
10,000 to 24,999	11.8%	13.2%	13.0%
25,000 to 49,999	9.6%	11.1%	11.3%
50,000 or more	20.4%	23.2%	25.3%

* Establishment calculations use all observations in California and treat interstate moves as births or deaths; firm calculations use all branches of firms with at least one establishment ever located in California.
** Equal to the number of employees in the initial year of the interval over which the growth rate is calculated.
*** Average business size for base year ($t-1$) and current year (t); used to avoid regression fallacy described by Davis et al. (1996a).

TABLE 3
Net job creation rate under alternative empirical specifications, by average establishment size, 1992-2004
(All industries, excluding government; all establishments in California unless otherwise noted)

I. Robustness to scope of observations

			Sample		
Size class (employees)	Baseline results (using all establishments in California)	Using establishments ever in California, including those that moved in or out	Using establishments always in California, excluding those that moved in or out	Using establishments in the manufacturing sector only*	Using establishments in the services sector only*
0 to 19 employees	1.4%	1.4%	1.4%	−1.1%	1.8%
20 to 49	0.8%	0.9%	0.9%	−1.1%	0.9%
50 to 99	0.7%	0.7%	0.7%	−0.6%	0.5%
100 to 249	0.4%	0.5%	0.5%	−0.7%	0.2%
250 to 499	−0.4%	−0.3%	−0.4%	−0.8%	−0.4%
500 to 999	−1.2%	−1.1%	−1.1%	−2.1%	−0.8%
1,000 to 2,499	−2.5%	−2.4%	−2.6%	−3.7%	−2.0%
2,500 to 4,999	−1.2%	−0.9%	−1.1%	−0.7%	−1.7%
5,000 or more	−2.1%	−2.0%	−2.2%	−1.9%	−1.9%
No. of observations	21,076,927	21,194,984	20,935,749	1,327,157	12,249,318

II. Robustness to use of imputed/rounded employment data

|			Sample		
Size class (employees)	Drop any establishment with employment imputed in *every* year	Drop any establishment with employment imputed in more than 50% years	Drop any establishment with employment imputed in more than 25% of years	Drop any establishment with employment imputed in *any* year	Drop any establishment if its employment is always a multiple of 10
0 to 19 employees	1.1%	0.8%	−0.5%	−2.8%	1.4%
20 to 49	1.0%	1.0%	0.5%	−0.7%	1.1%
50 to 99	0.8%	0.9%	0.6%	−0.2%	1.1%
100 to 249	0.5%	0.7%	0.5%	−0.2%	1.1%
250 to 499	−0.3%	−0.2%	−0.6%	−1.4%	0.4%
500 to 999	−1.2%	−0.8%	−1.0%	−1.7%	−0.1%
1,000 to 2,499	−2.4%	−1.6%	−1.7%	−1.9%	−1.8%
2,500 to 4,999	−1.1%	0.3%	0.0%	−0.6%	−1.4%
5,000 or more	−2.6%	−2.4%	−2.6%	−1.3%	−2.1%
No. of observations	16,730,011	12,982,772	9,089,282	5,582,543	20,351,872

* We treat the 2% of the establishments whose NAICS code changed over time as belonging to the industry in which they are classified for the most number of years.
In the event that an establishment is classified in two industries for an equally long period of time, the more recent of the two industries is chosen.

Figure 1: Net job creation rate vs. establishment size (10% bandwidth), all establishments in California

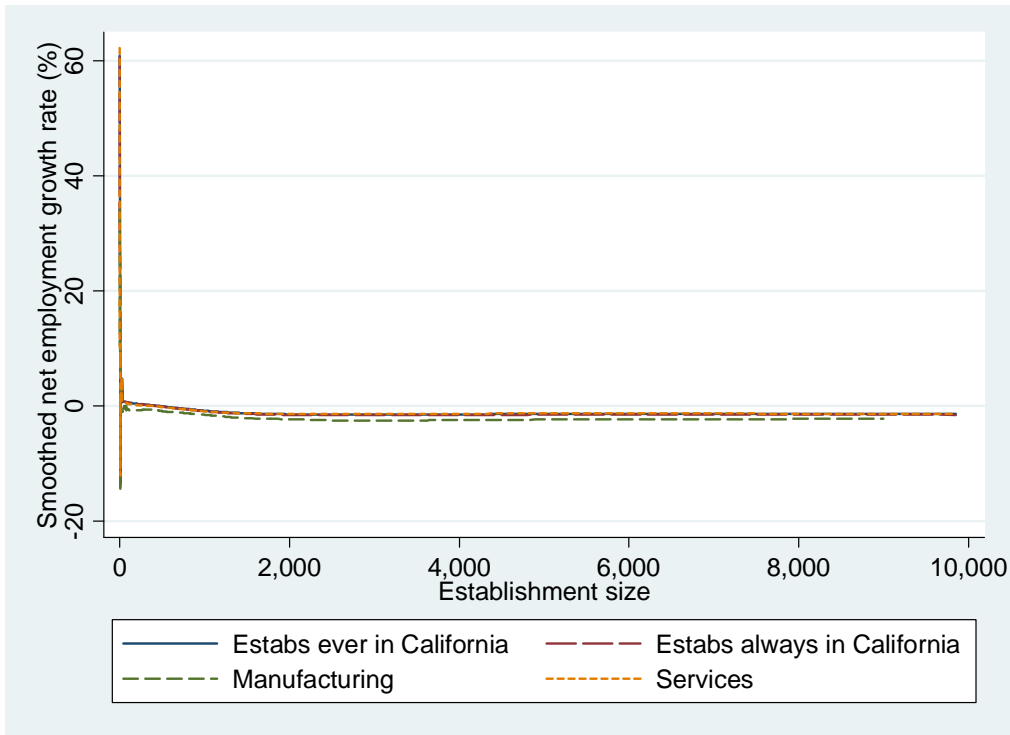**(I) Establishments up to 10,000 employees**
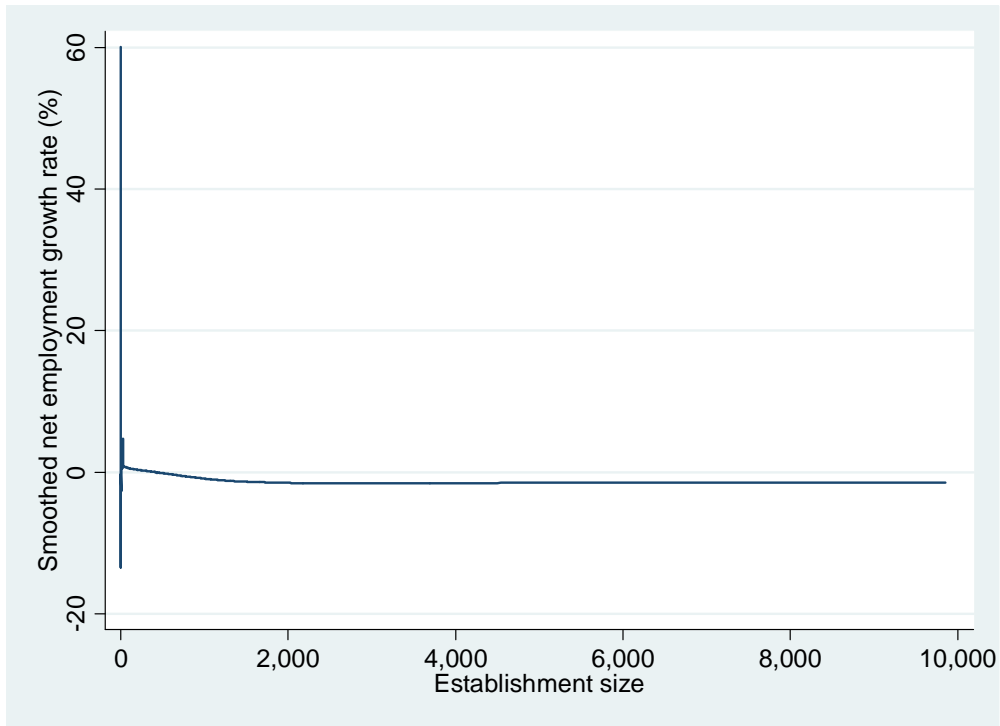
(II) Zoomed in to the range of 5-2,000 employees
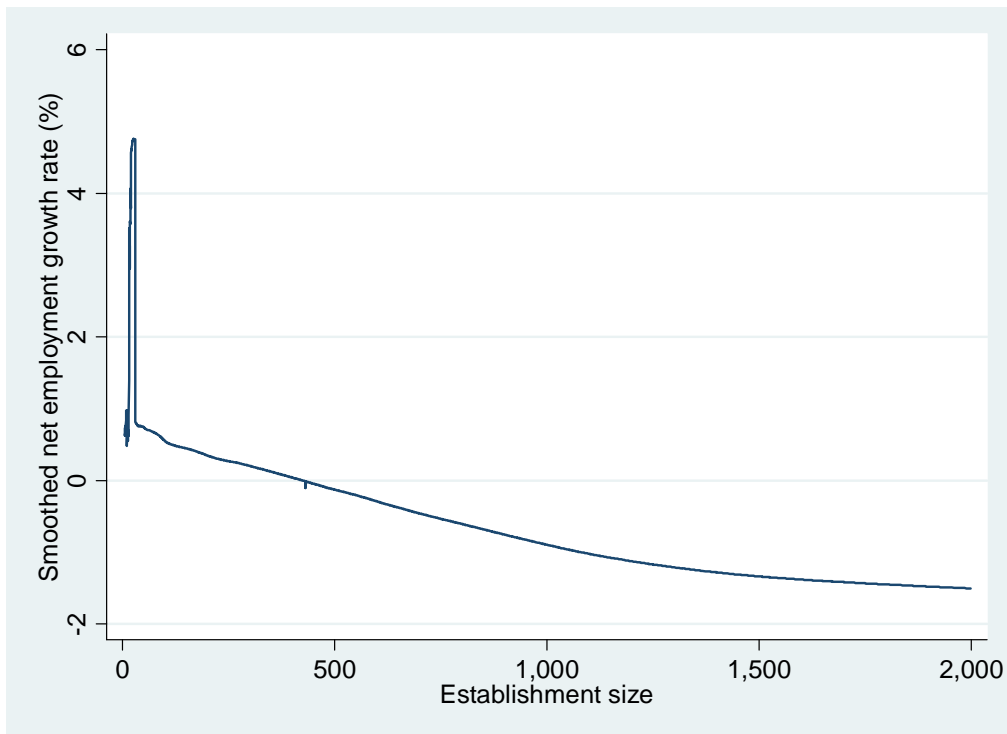
Figure 2: Net job creation rate vs. establishment size (10% bandwidth), robustness to scope of establishments

(I) Establishments up to 10,000 employees

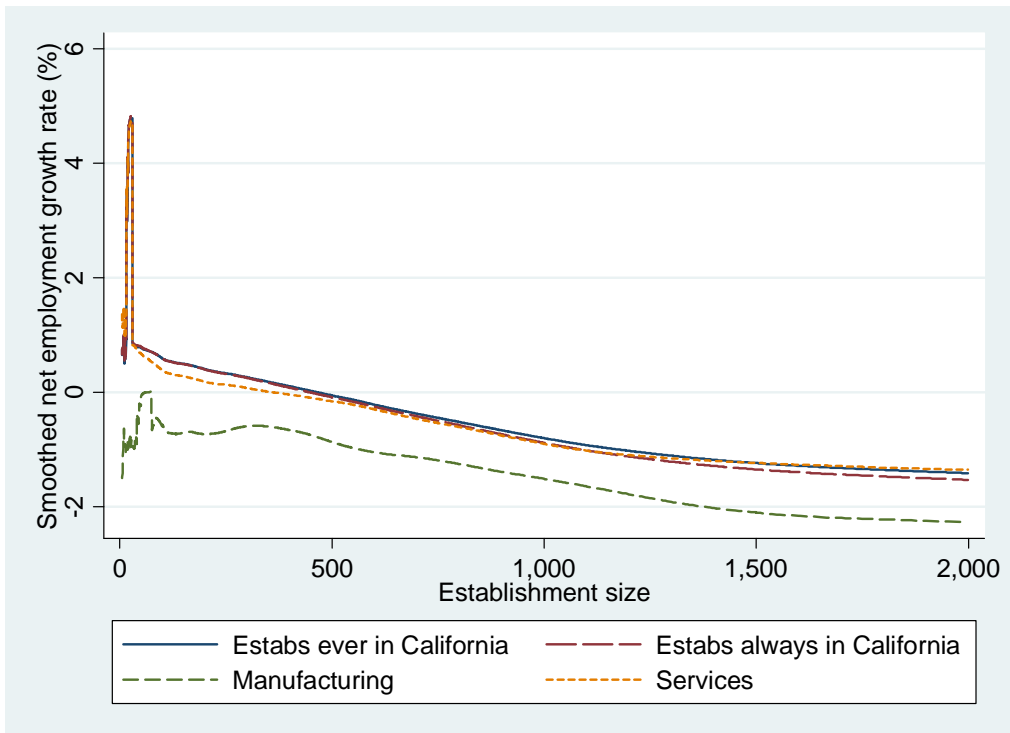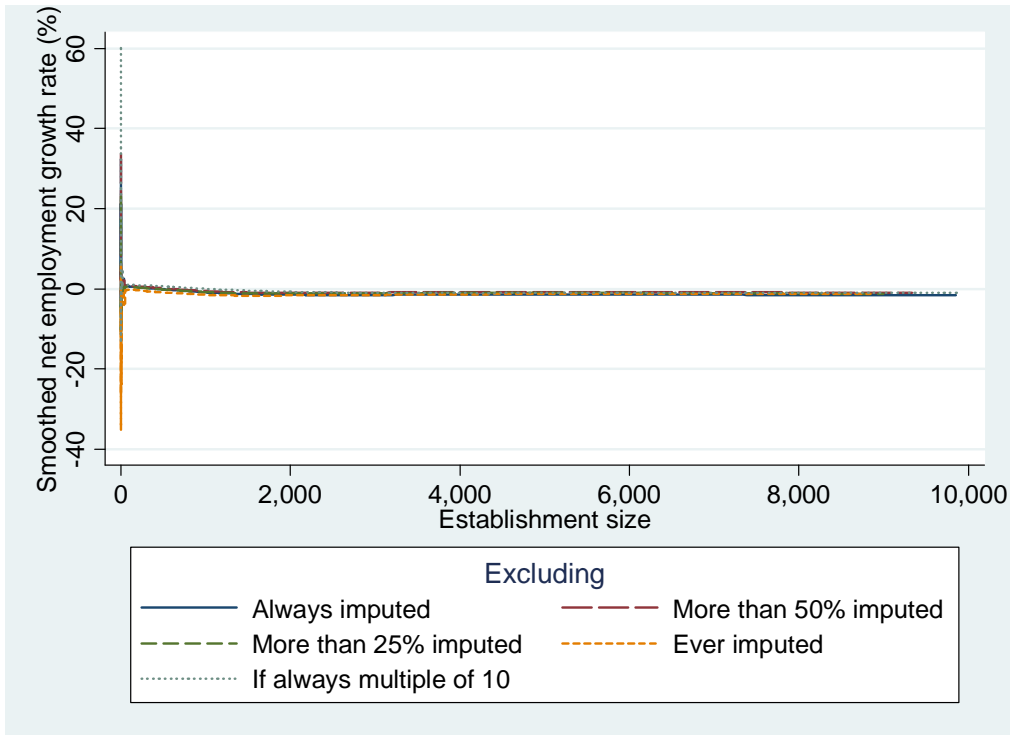**(II) Zoomed in to the range of 5-2,000 employees**

**Figure 3: Net job creation rate vs. establishment size (10% bandwidth),
excluding establishments with imputed employment data**

(I) Establishments up to 10,000 employees

(II) Zoomed in to the range of 5-2,000 employees