

IZA DP No. 9267

**Finance and Growth at the Firm Level:
Evidence from SBA Loans**

J. David Brown
John S. Earle

August 2015

Finance and Growth at the Firm Level: Evidence from SBA Loans

J. David Brown

*U.S. Census Bureau
and IZA*

John S. Earle

*George Mason University,
Central European University and IZA*

Discussion Paper No. 9267
August 2015

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Finance and Growth at the Firm Level: Evidence from SBA Loans*

We analyze linked databases on all Small Business Administration (SBA) loans, on all SBA lenders, and on all U.S. employers to estimate the effects of financial access on employment growth. Our methods combine regressions with matching on firm age, size, industry, year, and employment history, and with instrumental variables capturing ease of access to SBA lenders. The estimation results imply an increase of 3-4 jobs for each million dollars of loans, suggesting that credit constraints impede small business growth prior to loan receipt. We also investigate the variation in estimated employment effects for the SBA 504 versus 7(a) programs, and with respect to the business cycle, local credit conditions, and within-county versus non-SBA county-industry control firms. Finally, for loans issued over the 1992-2007 period, we estimate total job creation of 1.0-2.1 million and the government's cost per job of \$8,200-\$18,000 measured five years after the loan year.

JEL Classification: D04, G21, G28, H32, H81, J23, L53

Keywords: finance, financial access, loans, employment, growth,
small business administration, small business finance

Corresponding author:

John S. Earle
School of Public Policy
George Mason University
3351 Fairfax Drive, MS 3B1
Arlington, VA 22201
USA
E-mail: earle@gmu.edu

* We thank the National Science Foundation for support (Grant 1262269 to George Mason University) and participants in presentations at the Southern Economic Association Annual Meetings, the Comparative Analysis of Enterprise Data Conference in Nuremberg, George Mason University, Central European University, the Census Bureau, the Hungarian National Bank, the Small Business Administration, the Kauffman-Brandeis Entrepreneurial Finance and Innovation Conference, Wesleyan University, the Consumer Financial Protection Bureau, Emek Basker, and Javier Miranda for helpful comments on preliminary results. We thank Mee Jung Kim and Yana Morgulis for excellent research assistance. We also thank the SBA for providing the list of loans we use in the analysis. Any opinions and conclusions expressed herein are those of the authors and do not necessarily reflect the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information on individual firms is disclosed.

“Considering the practical difficulties of increasing investment incentives through tax reduction, particularly in a period of heavy defense expenditures, or of moderating the investment preference of individuals, some institutional solution to the problem of providing small business with long-term capital must be sought. In view of the special requirements of small business financial aid and the problems encountered in financing small business, it is doubtful if any of our existing financial institutions are equipped to perform this service.”

Schmidt (*Journal of Finance*, 1951)

1. Introduction

The “special requirements and problems in financing small business” have been recognized since at least the early 1950s, when the Small Business Administration (SBA) was founded. SBA loan programs have since then become one of the most significant government interventions affecting firm-level access to finance in the U.S. and have recently reached “all-time records in the Agency’s history, with over \$30 billion in lending support to 60,000 small businesses in its top two lending programs — 7(a) and 504” during fiscal year 2011.¹ Yet little is known about the outcomes of these programs.

In this paper, we estimate the firm-level impact of access to SBA loans on employment growth, which the SBA describes as its number one “strategic goal.”² Theoretically, the employment effects of SBA loans are ambiguous. On the one hand, easier access to capital may enable expansion, a scale effect. But it may also induce capital-labor substitution, and it would reduce employment if capital and labor are gross substitutes. Moreover, even if the scale effect dominates, so that the factors are gross complements, the employment rise may be attenuated if the program crowds out other sources of capital, and the aggregate effect may be reduced if there are general equilibrium displacement effects (negative spillovers onto competing firms).³ Empirically analyzing the programs is also difficult for several reasons: many factors influence employment and growth, loan receipt may be subject to selection bias (positive or negative), and appropriate firm-level micro-data have usually been unavailable.

Perhaps as a result of these problems – and despite the prominence of SBA programs and the high hopes in their power to stimulate business growth – there have been few attempts to measure their impacts using appropriate data and econometric methods. Unlike evaluations of job training programs, for example, research on SBA loan effects has had to rely on small samples, short time series, or aggregated data that do not permit the use of recent developments in econometrics (e.g., Imbens and Wooldridge, 2009). Most studies consist of simple comparisons before and after the interventions, with little use of comparison groups of non-recipients. The most common unit of observation in SBA studies is a geographic area such as the county or metropolitan statistical area, with outcomes measured as overall employment or

¹ SBA programs have received strong support both from congress and all recent presidential administrations, and small businesses are frequently cited as “...the places where most new jobs begin” (e.g., Whitehouse.gov, President’s Weekly Address, February 6, 2010). For the budget figures, see SBA (2011, p. 1).

² See SBA (2012, p. 7). The first goal is “growing businesses and creating jobs,” while the other two (which would be still more difficult to investigate) are “building an SBA that meets the needs of today’s and tomorrow’s small businesses” and “serving as the voice for small business.”

³ Hurst and Pugsley (2011) have recently criticized SBA programs on the basis that typical (median) small firms neither grow nor report wanting to grow, and thus that the emphasis on small businesses is misplaced. Our paper does not analyze the full range of assistance for small businesses, but we do show that recipients of SBA loans differ from the median in tending to grow prior to loan receipt, which is also an important issue for our identification strategy.

per-capita income in the local area; Craig et al. (2009) review these studies. Many factors affect county-level employment and income, of course, which makes it difficult to disentangle the effects of a program that is small relative to the local economy. The SBA itself reports a “performance indicator” – the number of “jobs supported,” reported in recent years at over 0.5 million, based on summing up borrowers’ estimates on their loan applications of the number of jobs they anticipate creating or retaining as a result of the loan.⁴

Our research aims to contribute to estimating these effects by using much better data than were heretofore available and by applying econometric methods developed for estimating causal effects with such data. We link administrative data on every SBA 7(a) and 504 program loan to long-panel data on the universe of employers in the U.S. economy, and we use the linked data to implement a longitudinal matching estimator (e.g., Heckman et al., 1997, 1998). The annual panels in our data run from 1987 to 2012 and permit us to select control firms based on age, industry, year, and several years of employment history, to control for fixed effects for time and treatment-control groups, and to measure the evolution of employment before and after the loans were awarded.

To account for unobservables in the selection process by which firms become SBA loan recipients, we also develop an instrumental variables strategy based on some institutional features of the SBA program as well as empirical regularities about small business credit markets. The SBA program works through partial (50-85 percent) loan guarantees provided to private lenders, some of which have a special status that reduces their administrative costs (preferred lender (PLP) status in the Preferred Lenders Program).⁵ As we show, these PLPs account for a large share of all SBA loans, particularly the larger loans, implying that participation in SBA loan provision is to some extent a function of corporate bank (firm-level) policy, such that some banks participate a great deal in SBA lending, while others do little or none. Even within the PLPs, the average size of SBA loans varies substantially, which may also reflect differences in corporate lending policy. Furthermore, the PLP banks are unevenly dispersed across the U.S. and have changed over time.

These are key facts for SBA loan provision, as they imply variation in the ease and likelihood of obtaining an SBA loan, especially when combined with previous observations on the continuing importance of the local credit market for small business lending (e.g., Brevoort and Hannan 2006). Indeed, we find that only 3.2 percent of SBA loans are in counties where there was no presence of branches of PLP lenders within the previous year, implying that SBA borrowers also tend to rely on local lenders. At the same time, SBA loans are a tiny share of the total banking business, accounting for only about one percent of all small business loans which themselves are only a fraction of all banking business, suggesting that locational decisions of banks are unlikely to be based on SBA business.

Based on this reasoning, we construct instrumental variables that reflect the local presence of a branch of a PLP bank weighted by the importance of SBA lending for the parent bank, measured outside the local lending market. In alternative specifications, “presence” is sometimes measured as a dummy variable for any PLP branch, sometimes as the number of PLP banks with local branches, and sometimes as the local share of bank branches affiliated with PLP

⁴ The figure is 583,737 for Fiscal Year 2010 (SBA 2011, p. 4). This number includes jobs supported not only by the 7(a) and 504 loan programs, but also the much smaller microloan and surety bond guarantee (SBG) programs.

⁵ Strictly speaking, PLP is the term used in the 7(a) loan program. The analogous preferred lender status for 504 loans is called the Preferred Certified Lenders Program (PCLP). Hereafter we use the term PLP to refer to both PLP and PCLP status.

lenders. The “importance” of the SBA lending to a bank is measured alternatively by the degree of consistency of PLP loan provision across a bank’s counties and by the average size of SBA loans offered by the parent outside the local area. Our identifying assumption for the instrumental variables is that banks do not enter new geographic (county) markets based on the potential for SBA loan business. When the instrument is defined in terms of the share of local branches which are part of PLP banks, rather than the presence or number of PLP banks, the identifying assumption is weaker, only requiring that PLP banks not choose their locations disproportionately in order to pursue SBA business. We also add controls for the local volume of small business lending and for the growth in the loan recipient’s county-industry cell during the same period for which we estimate an employment growth effect. The estimates imply an increase of 3-4 jobs in the first three years after loan receipt for each million dollars of loans. Provided our identification assumptions are valid, this suggests that SBA loan recipients did not simply replace conventional loans with SBA loans, but that they were genuinely credit constrained prior to receiving the SBA loans, and they used the loans to grow.

Finally, we examine several extensions to these basic results. We compare estimated effects for the two largest SBA programs, the 7(a) and 504 programs and find similar effects. We estimate the degree to which the estimated effects vary over the business and credit cycles, finding smaller effects when credit is more plentiful. We provide some assessment of the degree to which displacement of non-treated firms may be a concern, possibly an important factor underlying the estimates, distinguishing the non-tradables industries and comparing local controls with controls from other counties in which no SBA loans were given in the same industry. The results suggest there may be modest displacement effects, though they are imprecisely estimated. We also provide some robustness analysis concerning firm exit, and for the loans we study that were issued between 1992-2007 we calculate estimates of total job creation (1.0 – 2.1 million) and the cost per job (\$8,200 - \$18,000) measured five years after the loan year. As we discuss, this estimated cost per job is substantially lower than that for many other government job creation programs.

The paper builds on previous research on small business, finance, and government programs in several ways. It is important to clarify that much of the recent small business controversy in the U.S. has not concerned government interventions directly, but rather the empirical relationship between business size and employment growth. Such a description of empirical regularities underlies Birch’s (1987) claim that small businesses are responsible for most job creation and has been widely cited as the basis for government programs supporting this sector, although the underlying methods have been questioned by Davis, Haltiwanger, and Schuh (1996).⁶ But the best way to measure the size-growth relationship is a different issue from the impact of government programs on small business growth and performance, which is the question relevant for decisionmakers and the one we address in this paper.

The research in this paper is relevant to the broader theoretical and empirical literature on finance and growth (Levine 2005; Clementi and Hopenhayn 2006). As emphasized in Beck’s (2009) review of the econometric research, a standard identification problem is determining the direction of causality between growth and finance, and despite a long list of empirical studies, the degree to which financial development promotes economic growth remains controversial. Many studies use aggregate country-level data (e.g., King and Levine 1993 and Demetriades and Hussein 1996). Those using firm-level data frequently employ country-level measures of

⁶ See also Neumark, Wall, and Zhang (2011), Hurst and Pugsley (2011), and Haltiwanger, Jarmin, and Miranda (2013).

financial development, because of the difficulty of measuring financial constraints at the firm level; see the controversy over the approach of analyzing the relationship between investment and cash flow (Hubbard 1998).

Most closely related to our paper are some recent studies that have advanced this literature by employing firm-level data and innovative identification strategies. Banerjee and Duflo (2014) use changes in firm size eligibility for directed credit in India to identify the effects on firm growth, which they estimate to be strongly positive. Two French studies, Lelarge, Sraer, and Thesmar (2010) and Bach (2014), take advantage of changes in industry sector eligibility for a French loan guarantee program to measure the impact of the program on debt, employment, capital growth, financial expenses, and bankruptcy. The former paper finds that the program is associated with positive growth effects, but also a higher probability of bankruptcy; the latter finds that the program has positive effects on credit growth with no evidence of substitution between subsidized and unsubsidized finance, and no increase in default risk. Our paper also studies firm-level data and financial access through a government program, but rather than using eligibility rules for the identification strategy, we exploit geographic variation in access to the program, that is, in the geographic distribution of branches of PLP lenders.⁷

It is important to bear in mind a number of caveats to our estimates. As will be discussed further below, incompleteness in linking the databases imply measurement error in the treatment variable. We are unable to control directly for firm growth potential or for access to other sources of finance. Our instrumental variable approach assumes that banks do not choose the local markets in which they operate branches based on the local availability of projects suitable for SBA loans, and that the bank-level decision on how much to use the SBA guarantees outside a particular local market is not a function of the profitability of SBA lending within that market.

An additional caveat concerns external validity. In general, estimates of the average-treatment-effect-on-the-treated can be extrapolated to non-treated sub-populations only under strong assumptions, but our analysis of the SBA case suggests a particular reason for caution. We find that SBA loan recipients, unlike many firms, tend to grow before they receive the loan. This represents a pre-program trend of treated relative to non-treated firms that our matching procedure addresses in order to identify control firms with similar growth (and other characteristics). Our analysis thus compares treated firms with controls that have similar growth histories, not with the many non-treated small firms that never grow. Therefore, our results should not be interpreted as implying similar responsiveness to the program among all small firms, including the many that do not grow and do not want to grow (Hurst and Pugsley 2011), but they may be relevant for the sub-group of small firms that show growth for a significant period prior to receiving an SBA loan.

Section 2 describes the SBA programs we analyze. Section 3 describes the data, including the matched control samples. Section 4 explains our identification strategy. Section 5 provides results for alternative estimators of the employment effects of \$1 million of loans, and also contains several extensions, including analyses of robustness and displacement and

⁷ Whereas the Indian and French studies have implications for the effects of expanding or contracting a loan program along the firm size or sectoral dimensions, our focus on geographic variation in the ease of access is, we believe, a more relevant dimension for the SBA programs. Few SBA loan recipients are close to the SBA firm size thresholds, and nearly all legal, for-profit sectors are eligible for loans, so there is little scope for expanding the loan programs along those dimensions. In contrast, as we document in this paper, local access to PLP lenders varies widely across counties and over time.

estimates of the total job creation and cost per job associated with the SBA loan programs. Section 6 is a short summary.

2. SBA Loan Programs

The SBA has several small business loan guarantee programs. In this paper, we focus on the largest two categories of programs, 7(a) and 504, and this section briefly describes their current characteristics.⁸ Most 7(a) loans (those not part of a special subprogram) have a maximum amount of \$5 million, with a maximum 85 percent SBA guarantee for loans up to \$150,000, and a 75 percent maximum guarantee for higher amounts. Usually loans for working capital and machinery (not to exceed the life of equipment) have a maturity of up to 10 years, while loans for purchase of real estate can have a term up to 25 years. The SBA sets maximum loan interest rates, which decrease with loan amount and increase with maturity. To qualify, a business must be for-profit; meet SBA size standards;⁹ show good character, management expertise, and a feasible business plan; not have funds available from other sources; and be an eligible type of business.¹⁰ The SBA itself makes the final credit decisions for most of these loans.

Some 7(a) programs are more streamlined. Unlike with other 7(a) loans, in the Preferred Lender Program (PLP) the SBA delegates the final credit decision and most servicing and liquidation authority to PLP lenders. The SBA's role is to check loan eligibility criteria. The SBA selects lenders for PLP status based on their past record with the SBA, including proficiency in processing and servicing SBA-guaranteed loans. In payment default cases, the PLP lender agrees to liquidate all business assets before asking the SBA to honor its guaranty. In the 7(a) Certified Lender Program (CLP), the SBA promises a loan decision within three working days on applications handled by CLP lenders. Rather than ordering an independently conducted analysis, the SBA conducts a credit review, relying on the credit knowledge of the lender's loan officers. Lenders with a good performance history may receive CLP status.

A final large category of 7(a) is the express loan program. These have a \$350,000 maximum loan amount and 50 percent maximum SBA guaranty. Interest rates can be higher than on other 7(a) loans, but the SBA promises a decision on approval within 36 hours. PLPs also have an advantage here, as they may make eligibility determinations on their own.

The 504 Loan Program offers loan guarantees up to \$5.5 million, depending on the type of business. Typically a lender covers 50 percent of the project costs without an SBA guarantee, a Certified Development Company (CDC) certified by the SBA provides up to 40 percent of the financing (100 percent guaranteed by an SBA-guaranteed debenture), and the borrower contributes at least 10 percent (the borrower is sometimes required to contribute up to 20

⁸ SBA (2015) is the primary source for our description, and it contains further details. Glennon and Nigro (2005) and de Andrade and Lucas (2009) also describe the 7(a) program; they focus on defaults and on costs to borrowers, respectively, and they do not analyze the effects of the loans on employment or other outcomes.

⁹ The size standards vary by industry. In some industries the criterion is the average number of employees, with a cut-off ranging from 50 to 1,500. In other industries it is average annual receipts, ranging from \$750,000 to \$35.5 million. For many types of financial institutions, the cut-off is \$175 million in assets. Businesses not meeting these size standards, but which have tangible net worth of no more than \$15 million and average net income after Federal income taxes for the two full fiscal years before the date of application is not more than \$5 million. Also, size standards are increased by 25 percent when the applicant agrees to use all the financial assistance within a labor surplus area.

¹⁰ This includes engaging in business in the United States; possessing reasonable owner equity to invest; and using alternative financial resources, including personal assets, before seeking financial assistance.

percent). CDCs are nonprofit corporations promoting community economic development via disbursement of 504 loans. Proceeds may be used for fixed assets or to refinance debt in connection with an expansion of the business via new or renovated assets. Eligibility requirements for 504 loans are similar to those listed for 7(a) loans above.

For 7(a) loans, lenders must pay a guaranty fee that increases with maturity and guaranteed amount. For both programs they must sign the “Credit Elsewhere Requirement,” which states “Without the participation of SBA to the extent applied for, we would not be willing to make this loan, and in our opinion the financial assistance applied for is not otherwise available on reasonable terms.” This requirement, also called the “Credit Elsewhere Test,” must be accompanied by a detailed explanation why the loan would be unavailable on conventional terms.¹¹ Both the fee and the requirement create costs of using SBA loan guarantees. In addition, there are administrative costs to the lender, including the specific bureaucratic formulae for loan application and the monitoring that SBA carries out on any lender participating in the program. Probably because of scale economies in these costs, SBA loans tend to be concentrated in a relatively small number of lenders.¹² The concentration is especially great among PLP lenders.

To illustrate our points about SBA lenders and PLP lenders relative to all banks, Table 1 contains summary statistics for our database linking the SBA lenders to firms and establishments in the Census Bureau’s Longitudinal Business Database (LBD), as described in the next section. Statistics are shown for all banks (here defined as operating in the commercial banking sector or issuing an SBA loan), divided between non-SBA lenders, SBA but non-PLP lenders, and PLP lenders: there are 7,276, 1,485, and 258 of each of these, respectively. Measured as the number of branches, non-SBA lenders tend to be smaller than non-PLP SBA lenders, and both are smaller than PLP lenders; the same order is preserved in the number of counties in which each bank type is represented: 2.1, 4.2, and 36.0, on average. PLP lenders give SBA loans to borrowers in 10.4 counties of their counties with branches each year, on average, compared with only 1.2 for non-PLP SBA lenders. A PLP bank gives an average of 123.1 SBA loans each year, while the mean for non-PLPs is 6.5, and the PLP banks give larger loans. The mean PLP loan is \$460,822 while the mean SBA loan at non-PLP banks is \$286,288.

Further information on loan size is contained in Figure 1, with distributions of both all SBA and PLP loans for non-PLP and PLP lenders. Loan size varies substantially, implying that estimation should rely on this variation rather than a simple binary treatment dummy. PLP loans clearly tend to be larger, dominating all size ranges from \$150,000 on up. Access to a PLP lender therefore seems to raise both the probability and the likely size of an SBA loan.

These are interesting facts about SBA loans and lenders that have not before been documented, to our knowledge. The measures of numbers of branches and county distribution are not possible with the SBA database alone, but require linking to a database like the LBD. The concentration of SBA lending among banks suggests that there are likely to be sizable fixed costs to be involved with SBA, such that many banks do not participate in the SBA programs at all, while others do much more fully. The various dimensions of this concentration, in conjunction with other facts about small business lending, are useful in developing instrumental variables for our identification strategy below.

¹¹ Examples of acceptable factors are that the business needs a larger loan or longer maturity than the lender’s policy permits, or the collateral does not meet the lender’s policy requirements.

¹² de Andrade and Lucas (2009) also note that relatively few banks originate large numbers of 7(a) loans in 2006.

3. Data

We use a confidential database on all 7(a) and 504 loans guaranteed by the SBA from the fourth quarter of 1990 through the third quarter of 2009 to identify loan recipients, amounts, and time of receipt. We set the loan year to be on a fiscal year basis (October of the previous calendar year through September of the current calendar year), using the date the SBA approved the loan. We convert the year from calendar year to fiscal year so that the loan year is roughly centered on the Census Bureau's Longitudinal Business Database (LBD – described below) employment measure, which is the number of employees in the pay period including March 12. As shown in Appendix Table A1, our regression samples include loans starting in fiscal year 1992, as our instrumental variables (which also use information from the confidential loan database) are lagged one year relative to the firm's loan receipt. Loans to firms in U.S. territories are excluded, because of uneven coverage of other data sources. Cancelled loans are excluded, because the cancellation may be at the initiative of the borrower. Some borrowers receive multiple SBA loans in the same year, and we aggregate those amounts for our loan amount variable.¹³ To focus on the effects of the first treatment, we drop loans received in subsequent years.

We match the confidential SBA 7(a) and 504 data as well as publicly available 7(a), 504, and disaster loan data covering loans since the inception of these programs to the Census Bureau's employer and non-employer business registers.¹⁴ Since both the confidential 7(a) and 504 data and the Census Bureau data contain Internal Revenue Service Employer Identification Numbers (EINs) and Social Security Numbers (SSNs), we first link by EIN and SSN.¹⁵ For confidential 7(a) and 504 records that cannot be linked by EIN or SSN, and for the publicly available data without EINs or SSNs, we probabilistically link records by different combinations of business name, street address, and zip code. Appendix Table A1 shows that 86.9 percent of the confidential loan records are linked to a business register. Of these, 7.9 percent are linked only to a non-employer business register (i.e., they are self-employed and have no payroll employment). We exclude firms receiving a disaster loan before their first 7(a) or 504 loan, as well as firms receiving a 7(a) or 504 loan prior to 1992. Firms require an industry code, state (for those with 19 or fewer employees), and employment in the year prior to loan receipt to be included in the matching process with LBD control firms, as described in the next section. Of these, we could not find any control firms meeting our matching criteria (discussed in the next section) for 19 percent of them. About 57,000 treated firms do not have employment in each of the next three years following loan receipt, which is necessary for the dependent variable in the main regression sample. About 6,400 additional treated firms cannot be included in the main regression sample, because none of their matched controls has employment in each of the next three years after the treated firm's loan receipt.

¹³ Our loan amount variable is the amount disbursed, converted to real 2010 prices using the annual average Consumer Price Index. We use the total amount from loan financing, not just the amount guaranteed by the SBA. For 504 loans, we impute the total loan amount based on the guaranteed amount specified in the database, using the 504 program guidelines. The SBA-guaranteed portion is 40 percent, the equity share is 10 percent plus an additional 5 percent if a new business and/or an additional 5 percent if special use property, and the residual is a non-guaranteed bank loan. We are unable to observe if the project is for special use property; our imputations assume there is none. The database includes a third party loan amount, but it contains many implausibly high values.

¹⁴ The SBA has a separate disaster loan program, and we have names and addresses for the recipients from 1953 through March 31, 2011. We have chosen to focus the analysis on 7(a) and 504 loans in the confidential database, because the match rate to Census data is much higher due to the presence of EINs and SSNs.

¹⁵ About three-fourths of the linked records are linked via EIN or SSN.

To shed light on differences in the sample of SBA loan applications we study in this paper from samples we have had to exclude for various reasons, Appendix Table A2 provides the descriptive statistics for five different samples: those not linked to any business register, those in a business register but lacking one or more variables necessary for matching to control firms, those where an unsuccessful attempt to find controls is made, the regression sample where zeroes are imputed for post-exit years for firms exiting prior to three years after treatment, and the main regression sample including only those firms with positive employment in the first three years after treatment. Loan recipients not linked to any business register on average have much smaller loans and fewer employees, and a higher percentage are sole proprietorships, minority-owned, and female-owned than is the case in the other groups. Finding at least one matched control firm in the LBD for the treated firm is easiest in the service sector and hardest in manufacturing. The regression samples contain larger firms with bigger loans on average compared to any of the other groups.

The LBD consists of longitudinally linked employer business registers (Jarmin and Miranda 2002) tracking all firms and establishments with payroll employment in the U.S. non-farm business sector on an annual basis in 1976-2012. The SBA loan match to employer business registers allows us to link the SBA data to the entire LBD. Besides employment, the LBD contains annual payroll, establishment age (calculated based on the first year the establishment appears in the dataset), state, county, zip code, industry code, and firm id. The industry code is a four-digit SIC code through the year 2001 and a six-digit NAICS code in 2002-2012.

We aggregate the LBD to the firm level, assigning each firm the location of its largest establishment by employment and its modal industry code. The firm birth year is defined as the earliest birth year among establishments belonging to the firm when it first appears in the LBD, and the firm exit year is the latest exit year among establishments belonging to the firm in the last year the firm appears in the LBD. Our firm employment measure aggregates establishment employment, adjusted to focus on organic job creation. Employment in $t-1$, the year prior to the treatment year (defined for control firms as the matched treated firm's treatment year), is the base year (unadjusted) for treated firms and their matched controls. If a merger or acquisition occurs before the base year, the employment of the acquired establishments as of the year of the merger is included in the firm's employment in all years prior to the merger, as if the establishments were always together. If a divestiture occurs before the base year, the employment of the divested establishments is not included in the firm's employment prior to divestment, as if the establishments were never together. If a merger, acquisition, or divestiture occurs after the base year, employment of divested establishments as of the year prior to divestment is included in all subsequent years, while that of the acquired establishments is not included at all.¹⁶

To construct our instrumental variables, described in the next section, we link SBA lenders, as listed in the confidential SBA loan database, to the Census Bureau business registers and the LBD via name and address matching (no lender EINs or SSNs are included).¹⁷ We first

¹⁶ For acquisitions prior to the base year and divestitures after the base year, we use a single employment value applied to all pre-acquisition years and post-divestment years, respectively, to avoid including employment changes occurring under other firms' ownership.

¹⁷ In cases where an SBA lender was later purchased by another bank, SBA has replaced the original lender's name and address with the name and address of the acquirer. This introduces some measurement error in the instrumental variables.

perform a computer matching procedure like the one we use for loan recipients, then do a clerical match for the lenders remaining unmatched and which have issued at least five SBA loans. Appendix Table A3 shows that 74.5 percent of lenders and 84.9 percent of PLP lenders are matched to the LBD. The shares of all SBA loans issued by matched banks are higher – 91.9 percent of all SBA loans, 94.1 percent of PLP lenders’ SBA loans, and 90.1 percent of PLP lenders’ PLP loans. The LBD contains the locations of each of the establishments (branches) of the matched lenders. In some cases, the SBA database lists multiple lenders that we find belong to the same firm in the LBD, in which case we consolidate those SBA lenders’ loan activity.

4. Empirical Strategy

4.1 The Identification Problem

Ideally, we would like to estimate the causal effect of financial access on firm-level growth. The difficulties in such an estimation include the definitions of financial access and of growth and the fact that finance may take various forms, some of them unobserved. Most daunting is the problem that observed receipt of finance is a function of not only the availability of sources of finance but also of characteristics of the borrower including ability to repay and to benefit from more capital, which are again typically unobserved. In other words, we observe an equilibrium outcome determined jointly by supply and demand, both of them partially driven by unobservables.

Our approach in this paper is to focus upon the firm-level employment growth effects of a governmental intervention, the SBA loan programs. We choose employment growth as the outcome variable because it is the cleanest firm size measure in our data, and because the SBA itself states “growing business and creating jobs” as its top strategic goal. We focus on the SBA intervention because it is interesting in its own right and also because it provides the possibility of estimating the marginal value of this form of finance beyond that from other sources, including conventional loans and informal sources such as friends and family. We do not observe these other sources, so our estimates are not of the effects of expanding or contracting financial access more generally, which could be much more difficult to identify, but only of the effects of the programs. Given some identifying assumptions, however, the estimates can be interpreted as lower bounds on the effects of financial access for SBA loan recipients.

Although most of our estimation focuses on the impact of additional loan dollars, it is convenient to use the standard binary treatment notation for describing the identification problem. Let $TREAT_{it} \in \{0,1\}$ indicate whether firm i receives an SBA loan in year t , and let y_{it+s}^1 be employment at time $t+s$, $s \geq 0$, following loan receipt. The employment of the firm if it had not received a loan is y_{it+s}^0 . The loan’s causal effect on employment growth for firm i at time $t+s$ is defined as $y_{it+s}^1 - y_{it+s}^0$, and the average effect of treatment on the treated as

$$E\{y_{it+s}^1 - y_{it+s}^0 | TREAT_{it} = 1\} = E\{y_{it+s}^1 | TREAT_{it} = 1\} - E\{y_{it+s}^0 | TREAT_{it} = 1\}.$$

The standard evaluation problem is that y_{it+s}^0 is unobserved for recipients; thus $E\{y_{it+s}^0 | TREAT_{it} = 1\}$, the average employment outcome of loan recipients had they not received a loan, must be estimated. In one form or another, identification relies on some way of constructing this counter-factual average using the average employment of never-treated firms, $E\{y_{it+s}^0 | TREAT_{it} = 0 \forall t\}$. In doing so, a number of issues arise from the nature of the data and selection of the sample, from the observable characteristics of firms that tend to receive treatment, from possible self-selection into the SBA loan programs based on unobservables, and from unobservable characteristics of non-treated firms. Each of these creates potential biases that we consider in turn.

Beginning with the nature of the data and sample, our inability to link all SBA recipients in the SBA data to the LBD could lead to biases in either direction. The constructed control group may well contain unlinked treated firms whose treatment is unobserved. Indeed, because the controls are relatively high-growth firms (as a result of the matching to treated firms), the incidence of unobserved treatment among them is likely to be higher than in the whole population. We cannot estimate the magnitude of this problem, but if the true SBA effect is positive, then misallocating loan recipients to the controls implies a downward bias. Other potential biases stem from restrictions on the sample necessary for convincing identification. Start-ups are excluded because the lack of employment history reduces our ability to pair them with otherwise similar firms, but if start-ups have a stronger (weaker) employment response to SBA loans than do existing firms, then this exclusion again implies a negative (positive) bias. Finally, firms with no employees prior to loan receipt are excluded; if such firms have a stronger (weaker) employment response, the effect will be to bias our estimates downward (upward).

A second set of issues arises from observable characteristics of the treated and non-treated firms. Table 2 contains basic characteristics of these two groups, calculated from the LBD. Treated firms are smaller on average than non-treated firms (14 versus 20 employees), but they are larger at the median (7 versus 4 employees). Treated firms are younger (mean age of 8 versus 11 years, and median 6 versus 9). They are more likely to operate in certain industries, including manufacturing and trade.

Perhaps most interesting is that firms receiving an SBA loan tend to have grown much faster prior to the loan than typical non-recipient firms in the U.S. economy. Table 2 shows mean employment growth from 3 years to 1 year prior to the loan of 0.116 for treated firms versus 0.023 for all other firms measured in all years in the LBD. Figure 2 shows the distribution of growth rates over this same period: while the modal growth is zero for both treated and non-treated firms, the treated firms are much more likely to have grown substantially prior to treatment. This fact, which our research has uncovered for the first time, is not inconsistent with the regularity that most small firms exhibit low growth rates and many report little interest in growing (e.g., Hurst and Pugsley 2011), but it suggests that a naïve comparison of post- minus pre-loan employment levels of treated versus non-treated firms is likely to yield an upward-biased estimate. We use matching and regression procedures, described below, to account for such differences in observables as well as unobservables that are time-invariant.

The final set of issues concerns possible time-varying unobservables, both for the treated and control firms. An important case would be an idiosyncratic demand, productivity, or cost shock received by treated firms – but not by controls – precisely during the treatment year. The shock could be positive, for instance an idea or a project to raise demand or productivity, or negative, when demand falls or costs rise for some unobserved reason. In either case, a firm experiencing the shock might be motivated to seek a loan, in the first case for growth and in the second for survival. The former implies positive selection into SBA loan treatment, while the latter implies negative selection. In seeking to estimate a lower bound for loan effects, we are most concerned about the first case, but the second is also possible, so that *a priori* the direction of the estimation bias is indeterminate. In more formal discussion of the identification assumptions and potential biases below, we use the notation $S_{it} = 1$ for firm i receiving a shock to demand, cost, or productivity in year t ; we normalize the shock so that the expected effect on employment is positive. Defining variables as regression-adjusted to account for aggregate time effects, time-invariant unobservables, and other factors, treated firms receiving the shock will have average employment $E\{y_{it+s} | TREAT_{it} = 1, S_{it} = 1, X\}$, and treated firms not receiving the

positive shock will have $E\{y_{it+s}|TREAT_{it} = 1, S_{it} = 0, X\}$, where t is again the treatment year, s is the number of years since treatment, and X is the set of matching variables. The average among all treated firms is then a weighted average of these two groups, where the weights are proportions of firms in each group: $E\{S_{it}|TREAT_{it} = 1\}$ and $1-E\{S_{it}|TREAT_{it} = 1\}$, respectively.

A second type of unobservable is finance from a non-SBA source, such as informal credit or a conventional loan from a bank. Our data contain no firm-level information on any sources of capital aside from the SBA loan, so we cannot measure these, but it is useful to consider how they might affect the results. In general, if SBA loan recipients have more (less) additional financing in the SBA loan receipt year than controls, this might bias upward (downward) estimates of the effect of additional dollars in the SBA loan because they would then be reflecting the impact of all the additional credit, not just the amount of the SBA loan. However, there are several reasons to doubt that compared to controls, treated firms would be receiving larger amounts of fresh finance from non-SBA sources at the same time that they receive the SBA loan. To start with, “pecking-order theory” suggests that because of imperfect information and agency problems, firms should exhaust their internal and personal sources of finance, including informal borrowing from family and friends, before turning to the market (Myers and Majluf 1984), so SBA borrowers may be less likely to be freshly drawing upon these sources compared with controls.¹⁸ Moreover, SBA loans are heavily if not fully collateralized, and the SBA loan itself lowers the firm’s remaining debt capacity, so treated firms are *ceteris paribus* less likely than controls to receive other loans in addition. Because the implicit subsidy created by the guarantee lowers the interest cost, the entrepreneur should prefer a larger SBA loan to adding another conventional loan, even assuming such a loan is available.

Finally, the presence of significant financing from conventional loans received contemporaneously with an SBA loan would raise questions about whether the firm meets the “Credit Elsewhere Requirement” discussed above, possibly leaving the lender open to charges of fraud and abuse, resulting in potential fines and exclusion from SBA program participation. This consideration could be especially important for the PLP banks we focus on in our IV strategy, since they would have the most to lose.

For these reasons, we focus on the possibility that control firms receive non-SBA finance. A comparison of loans less than \$1 million from the SBA database with those from Community Reinvestment Act (CRA) disclosure data on business loan originations under \$1 million for 1996-2009 (FFIEC 2015) implies that the number of the former account for 0.83 percent of the latter and 4.48 percent of the volume.¹⁹ This suggests that many control firms are receiving such non-SBA finance. If we use $K_{it} = 1$ to denote those firms receiving finance through a non-SBA source in year t , we may distinguish four types of controls according to whether they receive the shock and whether they get non-SBA finance, with the following average employment in year s after the loan: $E\left\{y_{it+s} \mid \begin{matrix} TREAT_{it} = 0, \\ S_{it} = 0, K_{it} = 0 \end{matrix} \right\}$ for controls receiving neither a shock nor non-SBA finance, $\left\{y_{it+s} \mid \begin{matrix} TREAT_{it} = 0, \\ S_{it} = 1, K_{it} = 0 \end{matrix} \right\}$ for those receiving the shock but no other

¹⁸ Also, to be eligible for an SBA loan, applicants should have already exhausted these sources.

¹⁹ FFIEC (2015) contains CRA data only from 1996, and we calculate these ratios through 2009 because this is the last treatment year we study. Only about 0.06 percent of SBA loans exceed \$1 million during this period. An alternative calculation uses CRA loans to businesses with less than \$1 million in revenue as the denominator and all SBA loans in the numerator; the ratio in this case is 2.6 percent, which is still very low, and no doubt many SBA recipients have revenue above this threshold.

finance, $E\left\{y_{it+s} \mid \begin{matrix} TREAT_{it} = 0, \\ S_{it} = 0, K_{it} = 1 \end{matrix}\right\}$ for those not getting the shock but receiving non-SBA finance, and $E\left\{y_{it+s} \mid \begin{matrix} TREAT_{it} = 0, \\ S_{it} = 1, K_{it} = 1 \end{matrix}\right\}$ for those getting both the shock and other finance. The average among all controls is a weighted average of these, with weights determined by the proportions with $S_{it} = 1$ and $K_{it} = 1$ among the controls. If S and K are defined so that $E\{y_{it+s}|S_{it} = 1\} > E\{y_{it+s}|S_{it} = 0\}$ and $E\{y_{it+s}|K_{it} = 1\} > E\{y_{it+s}|K_{it} = 0\}$, implying shocks and increased finance lead to higher employment, then the estimator $E\{y_{it+s}|TREAT_{it} = 1\} - E\{y_{it+s}|TREAT_{it} = 0\}$ will be increasing in the share of treated firms receiving positive shocks and decreasing in the share of controls receiving shocks and in the share of non-SBA finance.

This analysis clarifies both some potential sources of bias and the interpretation of our estimates. Assuming, conditional on our matching procedures and regression adjustments, that treatment is independent of the shock, so that $E\{TREAT_{it}|S_{it} = 1, X\} = E\{TREAT_{it}|S_{it} = 0, X\}$ and if in addition, there is no non-SBA finance ($K_{it} = 0$), then our estimate can be interpreted as the causal effect of finance on growth. Maintaining the conditional independence of treatment, but permitting non-SBA finance, implies that our estimate is still the causal effect of the program. If we assume either that K_{it} is received only by control firms or, alternatively, that more is received by controls than by treated firms (with the same marginal effect on employment growth), then our estimate will be a lower bound of the causal effect of finance on growth.

Relaxing the conditional independence assumption, selection bias reflects differences between the counterfactual outcome for treated firms and the observed outcome for the controls: $B(X) = E\{y_{it+s}^0|TREAT_{it} = 1, X\} - E\{y_{it+s}|TREAT_{it} = 0, X\}$. If $E\{TREAT_{it} * S_{it}|X\} > 0$, so that if treated firms would have performed better than non-treated even if the former had not been treated, then $B(X) > 0$. The magnitude of $B(X)$ in this case depends on the strength of the correlation between treatment and the shock. Our final step, therefore, will be to construct instrumental variables for SBA loan receipt that are functions only of SBA loan supply and therefore independent of the shock. We also add variables to control for industry-county-specific growth and the availability of conventional finance in the locality. Before turning to the instrumental variables, we first discuss the matched sample and the regression specification.

4.2 Matching and Regression Procedures

To take into account observable and fixed unobservable characteristics that may affect both the growth of the firm and the size and probability of an SBA loan, we estimate regression functions on a matched sample of treated and non-treated firms. Our matching procedures include sample restrictions, exact matching on several firm characteristics, and propensity score matching on several other variables, including a four-year history of employment. As discussed in Section 3 above, we limit our treated sample to firms in the LBD that are at least one year old when receiving their first SBA loan, those receiving their first SBA 7(a) or 504 loan in 1992-2009, those not receiving an SBA disaster loan prior to their first 7(a) or 504 loan, and with non-missing and non-outlier employment in the LBD in the year before loan receipt.²⁰ We restrict the non-treated sample so that a firm eligible to be a candidate control firm for a particular treated firm must have non-missing employment in the year prior to the treated firm's loan

²⁰ We define the following cases as outliers: an employment increase or decrease of more than ten times between the first and second year of life or the second-to-last and last year of life; or an employment increase (decrease) of more than five times followed in the next year by an employment decrease (increase) of more than five times.

receipt (which also means it isn't a new start-up in the year of loan receipt) and no employment outliers in the LBD; and it can never have received an SBA 7(a), 504, or disaster loan at any time between 1953-2012.

The exact matching requirements for control firms include industry, age-group, size-group, and year. Specifically, for a firm to be a candidate control firm for a particular treated firm it must be in the same four-digit industry (this is the four-digit SIC code through 2001 and the first four digits of the NAICS code in 2002-2009) in the treated firm's loan receipt year. It must be in the same firm age category (1-2 years old, 3-5 years old, 6-10 years old, and 11 or more years old) in the treated firm's loan receipt year. It must be in the same firm employment category (1 employee, 2-4 employees, 5-9 employees, 10-19 employees, 20-49 employees, 50-99 employees, and 100 or more employees) in the year before the treated firm's loan receipt year. Among firms with 19 or fewer employees in the previous year, we also require the candidate control firm to be located in the same state (firms with 1-19 employees are much more numerous than ones with more than 19 employees, so we can afford to impose more restrictions on this group). In addition, we impose a restriction that the ratio of the treated firm's employment in the previous year to the control firm's previous year employment be greater than 0.9 and less than 1.1. This means that among firms with ten or fewer employees, a majority of the sample, employment must match exactly.

The exact matching creates a set of controls essentially identical in pre-treatment size, industry, and age. But there are other variables that we would like to match on, especially employment history and wage for which it is difficult to design matching thresholds for each variable separately, so we reduce this dimensionality problem with propensity score matching. After dropping treated firms with no candidate controls based on the sample restrictions and exact matching conditions, we estimate separate probit regressions for each age-size group using the sample of treated firms and their candidate controls. There are 28 age-size groups, based on the four age groups and seven size groups defined above. The probit regresses a dummy for SBA loan receipt on the following variables: the log of employment in the year prior to the treated firm's loan receipt; the square of the log of employment in the year prior to the treated firm's loan receipt; the cube of the log of employment in the year prior to the treated firm's loan receipt; the log of employment one year before minus log employment two years prior to the treated firm's loan receipt; the square of the log of employment one year before minus log employment two years prior to the treated firm's loan receipt; the cube of the log of employment one year before minus log employment two years prior to the treated firm's loan receipt; the log of employment two years before minus log employment three years prior to the treated firm's loan receipt; the square of the log of employment two years before minus log employment three years prior to the treated firm's loan receipt; the cube of the log of employment two years before minus log employment three years prior to the treated firm's loan receipt; the log of employment three years before minus log employment four years prior to the treated firm's loan receipt; the square of the log of employment three years before minus log employment four years prior to the treated firm's loan receipt; the cube of the log of employment three years before minus log employment four years prior to the treated firm's loan receipt; the analogous variables for log revenues and log assets as for employment above; the log of payroll/number of employees in the year prior to the treated firm's loan receipt; firm age; firm age squared; a multi-unit firm dummy; and year dummies. For the lagged employment growth variables, all revenue and assets variables, and for the log of payroll/number of employees in the year prior to the treated firm's loan receipt, we also impute zeroes in place of missing values and include dummies for such

cases. Conditioning on four years of lagged employment, revenues, and assets is intended to create a control group with very similar histories to the treated firms, and it follows Heckman et al.'s (1997, 1998, 1999) recommendations for evaluating the outcomes of labor market training programs.

The treated firm observations in the probit regressions are each assigned a weight of $\frac{(N-R)}{R}$, where N is the total number of firms in the regression and R is the number of treated firms in the regression. The non-treated firms are assigned a weight of 1. This equalizes the total weight of the treated firm and non-treated firm groups. The purpose of this weighting is to produce propensity scores that span a wider range, centered around 0.5 rather than near zero.

We limit the treated and non-treated firms in the final sample to those within a common support, meaning that no propensity score of a treated (non-treated) firm that we use is higher than the highest non-treated (treated) firm propensity score, and no propensity score of a treated (non-treated) firm that we use is lower than the lowest non-treated (treated) firm propensity score. A non-treated firm is included as a control for a particular treated firm if the ratio of the treated to the non-treated firm's propensity score is at least 0.9 and not more than 1.1. Treated firms with no controls meeting all these criteria are not included in the employment regression analysis. Non-treated firms appear in the employment regressions as many times as they have treated firms to which they are matched (i.e., this is matching with replacement). Kernel weights are applied to the controls.²¹ In the employment regressions, each control is assigned a final weight of their kernel weight divided by the sum of the kernel weights for all controls for a particular treated firm, and the treated firm is given a weight of 1. As a result, the treated firm and all its control firms together receive equal weight.

Evidence for the success of the matching procedures in almost entirely eliminating the observable differences between the treated and non-treated firms can be found in Appendix Table A4, particularly in the calculations of percentage bias reduction for each variable. Treated firm employment and average wage are substantially larger than for non-treated firms prior to propensity score matching, and treated firms experience more employment growth in the four years prior to treatment. After matching, these differences are negligible. The standardized difference measures confirm this: employment, employment growth, revenue, age, and multi-unit biases are reduced by over 90 percent. None of the standardized differences (labelled final % biases in Table A4) are close to being large after matching.²²

One way to estimate of the loan effect would be to simply compare the average outcomes for treated versus control firms. Another would compare outcomes for the data after differencing from the pre- to post-loan periods to eliminate unobserved, time-invariant differences; Heckman et al. (1997, 1998, 1999) recommend combining matching with difference-in-difference (DID) methods. For a number of reasons, we implement DID methods using the matched sample to estimate regressions. The regressions not only to account for time-invariant unobservables in a natural way, but also time-varying observables and unobservables, and they permit us to estimate the effect of the size of the SBA loan, rather than the binary treatment. As we show in Figure 1, loan size varies quite considerably, and the analysis should allow for this variation.

Our analysis focuses on the first SBA loan, as subsequent loan receipt (approximately 20-25 percent of the loan sample is subsequent loans) may be influenced by the outcome of the first

²¹ The kernel weight is $1 - \left(\frac{\text{abs}\left(\frac{\text{propensity score}_{tr} - 1}{\text{propensity score}_{ntr}}\right)}{0.1} \right)^2$, where tr refers to treated and ntr to non-treated firms.

²² Rosenbaum and Rubin (1985) consider a value of 20 to be large.

one. Also, given our long time series, we find it useful to constrain the time frame around which we calculate employment growth to focus on the short- and medium-term effects of the loan. This puts all of the loan cohorts on an equal footing, so that each counts equally rather than having longer time series for the early cohorts and shorter time series for the later ones. The basic form of the regression therefore uses the change in employment as the dependent variable, as follows:

$$\Delta E_{ijt} = \boldsymbol{\rho}_t + \alpha_j + X_{ijt}\boldsymbol{\beta} + \theta_i\delta + u_{ijt},$$

where ΔE is the change in the number of employees over some period, i indexes firms from 1 to I , j indexes from 1 to J the treated firms to which the firm is a control (for treated firms $i = j$), and t indexes the loan years from 1 to T . $\boldsymbol{\rho}_t$ is a vector of year dummies for loan cohorts, α_j is a fixed effect for each group of treated firms and its matched controls (the “treatment-control-group”), X_{ijt} is a set of other variables including firm age and age squared, and in some specifications proxies for demand conditions and finance availability; u_{ijt} is an idiosyncratic error. θ_i is the amount of the SBA loan (which equals 0 for non-treated firms) received in year t , and δ is the loan effect of interest. In some specifications, we permit δ to vary by type of loan or other variables, as explained in Section 5.2 on extensions below.

The dependent variable is defined in one specification as the change in the average employment from three years before to three years after the loan: $\Delta E_{ijt} = E_{ij,post} - E_{ij,pre}$, with $E_{ij,post} = (E_{ijt+1} + E_{ijt+2} + E_{ijt+3})/3$, and $E_{ij,pre,t} = (E_{ijt-1} + E_{ijt-2} + E_{ijt-3})/3$.²³ In other specifications, we define a base year as the year prior to the loan, $t-1$, and compute employment differences for each year from five before to five after the loan: $\Delta E_{ijt} = E_{ijt+s} - E_{ijt-1}$, $s = -5, -4, \dots, 4, 5$. The pre-loan coefficients are useful for diagnosing some forms of selection (they permit a “pre-program test” in the sense of Heckman and Hotz 1989, or a “pseudo-outcome” test in the sense of Imbens and Wooldridge 2009). The post-loan coefficients permit measurement of short- versus long-term effects.

4.3 Instrumental Variables

The matching and regression procedures account for observable differences and time-invariant unobservable differences between the treated and control firms. For the estimates from these methods to be interpreted as causal, the identifying assumption is that there is no systematically different shock to the treated versus control firms’ demand, productivity, costs, or access to capital from other sources. As discussed in Section 4.1 above, however, there are plausible reasons to expect that such shocks may occur, and there are several factors so that the direction of the bias is *a priori* unclear. The greatest concern for our attempt to estimate a lower-bound effect would be that treatment results from a positive shock, such as a new idea for business growth, and that the observed post-loan growth results from this shock rather than the access to capital provided by the SBA programs.

To account for such unobservables in the selection process by which firms receive SBA loans of varying sizes, we therefore develop an instrumental variables strategy based on some institutional features of the SBA program as well as on empirical regularities about small business credit markets. As discussed in Section 2, the SBA program works through partial (50-85 percent) loan guarantees provided to private lenders, some of which are PLPs with a special status that reduces their administrative costs. As we showed, these PLPs account for a large share of all SBA loans, particularly the larger loans. The concentration of loans suggests that

²³ In cases of missing values, we average employment during the available years $t-3$, $t-2$ and $t-1$.

participation in SBA loan provision is to some extent a corporate bank (firm-level) decision, such that some banks participate a great deal in SBA lending, while others do little or none. Even within the PLPs, the average number and average size of SBA loans varies substantially, which may reflect differences in corporate lending policy among these banks. Furthermore, the PLP lenders are unevenly dispersed across the U.S. and have changed over time.

Another key fact is that SBA borrowers, like other small businesses, seem to rely heavily on local lenders. The importance of the local credit market for small business lending is well-documented in previous research.²⁴ While our SBA data contain headquarters rather than lending branch location, so that distance cannot be computed directly, indirect evidence of local lending for SBA loans comes from examining the incidence of SBA loans in counties in which no PLP lenders, nor any SBA lenders, are represented. The share of SBA loans in counties where there was no PLP branch in the previous year is 3.2 percent, and 0.3 percent for no SBA branch. The corresponding shares for PLP loans are 0.02 percent and 0.0003 percent respectively. Together, these facts imply geographic variation in the ease and likelihood for a small business to obtain an SBA loan and in the likely loan size. Our analysis defines the financial institutions in a county as the local credit market.

At the same time, SBA loans are a tiny share of the total banking business, even for lenders who do the most SBA lending, and it seems unlikely that banks choose to enter and exit local credit markets as a result of this line of business. An exact calculation of the share of SBA loans in all loans seems to be impossible with existing data sources, but we can combine several of them to produce a rough estimate. For the top ten SBA commercial bank lenders in 2012, for instance, we have combined information from the SBA database with CRA bank disclosure reports to calculate for all new loans under \$1 million the ratio of the total amount of SBA loans to all CRA loans, both measured at origination. This ratio is 0.050, implying that the total volume of SBA loans under \$1 million is 5 percent of the total volume of all loans under \$1 million. To assess the role of loans of this size in all bank lending, we must turn to FDIC call reports. For the same top ten SBA commercial bank lenders in 2012, we can calculate the ratio of the volume of loans in the categories of commercial and industrial loans and those secured by nonfarm nonresidential properties that were originally under \$1 million to the volume of all loans, in both cases on a stock basis, so that this refers to loans outstanding.²⁵ The ratio is 0.052, implying that small loans (under \$1 million in this measure) to nonfarm businesses only account for about 5 percent of all bank loans. Although these two ratios are on different bases (flow versus stock), if we were to multiply them, they would suggest that SBA loans in 2012 accounted for only 0.25 percent of all bank loans, even for these largest SBA lenders. Banks also have other lines of business – deposit collection, payment processing, asset management, financial advising, and investment banking – that are not accounted for here, but may be important determinants of branch locations.

Finally, in any given year, PLP lenders actually issue PLP loans in only a fraction of the counties in which they operate: Table 1 shows that PLP banks have branches in an average of 36

²⁴ See, e.g., Amel and Brevoort (2005), and Agarwal and Hauswald (2010), who report median distance of 2.4 to 4 miles from borrower to lending branch.

²⁵ The amount of outstanding loans that were less than \$1 million when originated is the sum of the June 30, 2012 FDIC call report variables for amount currently outstanding of loans secured by nonfarm residential properties and commercial and industrial loans (RCON5565, RCON5567, RCON5569, RCON5571, RCON5573, and RCON5575). Total loans outstanding is the June 30, 2012 FDIC call report variable total loans and leases, net of unearned income (RCON2122).

counties during 1991-2009, while the mean number of counties in which borrowers received PLP loans from one of these banks is 5.4. For all these reasons, banks' choices of counties in which to locate their branches seem quite unlikely to be related to the possibilities for SBA loan business.

Based on this reasoning, we construct instrumental variables that reflect the local presence of a branch of a PLP bank weighted by the importance of SBA lending for the parent bank, measured outside the local lending market. The "importance" is measured by the degree of consistency of PLP loan provision and by the average size of SBA loans offered by the parent outside the local area. Besides considering various instrumental variables, we also examine robustness with controls for the local business and financial environment, including the availability of conventional loans in the immediate pre-loan year and the growth in the loan recipient's county-industry cell during the same period for which we estimate an employment growth effect. We present results for five different instrumental variables that rely on different assumptions for the process by which small businesses search for loans and are potentially matched with an SBA lender. The identifying assumption for all the instrumental variables is that banks do not enter new geographic (county) markets based on the potential for SBA loan business. In the simplest case, the variable is simply an indicator for the presence in county c of at least one branch of at least one bank b that made at least one PLP loan in a different county $-c$ in year $t-1$:

$$\text{IndicatorIV}_{ct} = \mathbb{I}(1 | D_{cbt-1}P_{-cbt-1} = 1 \text{ for any } b),$$

where $D_{cbt-1} = 1$ if bank b has at least one branch in this county c in year $t-1$ and $P_{-cbt-1} = 1$ if bank b has at least one PLP loan in some other county $-c$ in year $t-1$ and t is the loan year. The assumption here, based on our analysis of the lender data in Section 2, is that borrowers are much more likely to get an SBA loan and to receive a bigger SBA loan if they are located in a county with the presence of a lender that offered a PLP loan in the previous year. We do not count PLP loans issued to firms in the same county as the borrower to avoid potential endogeneity of PLP loan availability in the same county.

A second variable weights the presence of different banks in a county by the strength of the corporate policy favoring SBA loans, here measured by the share of other counties in which the bank made PLP loans in the previous year:

$$\text{SumSharesIV}_{ct} = \sum_b \left[D_{cbt-1} \times \left[\frac{\sum_{-c} D_{-cbt-1} P_{-cbt-1}}{\sum_{-c} D_{-cbt-1}} \right] \right]$$

For instance, in a county with three banks present, two of which made PLP loans the previous year, one in half its counties and the other in one quarter, $\text{SumSharesIV}_{ct} = 0.75$. The assumption is that a borrower is more likely to get an SBA loan, and to get a bigger one, if there are more local branches of PLP banks for which the PLP loan business is important, in the sense of being given to borrowers in a larger share of the counties in which they operate.

We have no information on how small businesses search across lenders for a loan, and how they may be eventually matched with an SBA lender, but one possibility is that it is not the number of PLP lenders with local branches that matters, but rather the share of all local branches they account for. Also, since the SumSharesIV_{ct} is additive in the number of PLP lenders in the county, counties with more banks are likely to have larger values. The number of banks in a county could be correlated with unobserved factors contributing to small business growth. The

next variable addresses these possibilities by weighting each locally represented PLP bank by the bank's share of the county's bank branches:

$$\text{BranchShareCounty*SumSharesIV}_{ct} = \sum_b \left[[E_{cbt-1}/E_{ct-1}] \times D_{cbt-1} \times \left[\sum_{-c} D_{-cbt-1} P_{-cbt-1} / \sum_{-c} D_{-cbt-1} \right] \right]$$

where E_{cbt-1} is the number of establishments (branches) of bank b in county c in time $t-1$ and E_{ct-1} is the total number of banking establishments in county c in time $t-1$.

So far these variables do not account for differences across PLP banks in the sizes of SBA loans that they offer, but Figure 1 of Section 2 shows substantial variation across banks in average loan size. The final two instrumental variables take into account this variation by multiplying each bank-county term by the average loan size of bank b . The first of these is a variant of the SumSharesIV_{ct} :

$$\text{AvgLoanAmount*SumSharesIV}_{ct} = \sum_b \left[\left[\sum_{-c} A_{-cbt-1} / \sum_{-c} L_{-cbt-1} \right] \times D_{cbt-1} \times \left[\sum_{-c} D_{-cbt-1} P_{-cbt-1} / \sum_{-c} D_{-cbt-1} \right] \right]$$

where A_{-cbt-1} is the total SBA loan amount issued by bank b to firms located in other county $-c$ in time $t-1$ and L_{-cbt-1} is the number of SBA loans issued by bank b to firms located in other county $-c$ in time $t-1$.

The last instrument is a variant of the $\text{BranchShareCounty*SumSharesIV}_{ct}$ including the average loan amount A_{-cbt-1}/L_{-cbt-1} as an additional weighting factor:

$$\begin{aligned} \text{AvgLoanAmount*BranchShareCounty*SumSharesIV}_{ct} = \\ \sum_b \left[\left[\sum_{-c} A_{-cbt-1} / \sum_{-c} L_{-cbt-1} \right] \times [E_{cbt-1}/E_{ct-1}] \times D_{cbt-1} \right. \\ \left. \times \left[\sum_{-c} D_{-cbt-1} P_{-cbt-1} / \sum_{-c} D_{-cbt-1} \right] \right] \end{aligned}$$

This variable is positive if the county has at least one branch of a bank offering a PLP loan outside the county in the previous year, and it is increasing in the average loan amount offered outside the county and the share of branches in the county affiliated with a PLP lender.

Table 3 contains summary statistics for the instrumental variables and their components. All show substantial variation over the sample, with standard deviations about twice the mean. The correlations of these variables with the treatment variable loan amount are shown in the first-stage results in the next section.

5. Results

We report estimates in two sub-sections. Section 5.1 contains the results from estimating the average treatment effect of an SBA loan on treated firms for a variety of different specifications. We briefly consider a binary treatment specification before moving on to specifications that use the variation in loan amount, which as we showed in Section 2 differ widely among SBA loan recipients. To assess robustness, we provide results for unmatched and matched samples, for specifications with and without matched group fixed effects, and for estimation by OLS and with the five instrumental variables discussed in the previous section.

Section 5.2 then presents results for several extensions to this basic analysis, including a comparison of the two the SBA 7(a) and 504 programs, estimation of the degree to which the implied loan effects vary over the business and credit cycles, and preliminary analyses of the extent to which displacement of non-treated firms and firm exit might be driving the results and of the overall employment and dollar cost per job created under the SBA programs.

5.1 Estimating the Average Effect of SBA Loans

In order to shed more light on the nature of the data and the identification problem, it is useful before presenting the results from our preferred estimators to take a step back and consider some alternative empirical strategies, which include some approaches that are more traditional when measuring program impacts. These methods include use of a binary measure for treatment (a dummy variable for loan receipt rather than a continuous variable for loan amount), and using some simpler versions of matching without treatment group fixed effects, instrumental variables, or the regression controls we consider later. We assess the plausibility of the alternative specifications by studying the estimated pre-loan dynamics, a type of Heckman-Hotz “pre-program test,” and we report the short- and long-run dynamics of the estimated effects.

The top panel of Figure 3 contains first results from a binary treatment specification using a control sample drawn from the entire LBD with essentially no matching (we line up the control firms only by calendar year, to be able to define the controls, and by pre-loan year size so that the size distributions for the treated and control samples are similar). As discussed in Section 4.2, we define t as the loan year, and use the year prior to the loan, $t-1$, as the base year, computing employment differences for each year from five before to five after the loan, so that $\Delta E_{ijt} = E_{ijt+s} - E_{ijt-1}$ ($s = -5, -4, \dots, 4, 5$). The regressors include only the binary treatment, firm age and age-squared, and calendar year effects. The graph shows steady growth of SBA-recipient firms relative to non-recipients from $t-5$ to $t+5$, with all estimates highly statistically significant (dotted lines show 99 percent confidence intervals). A naïve interpretation might compare $t-1$ to $t+5$ and infer 3.3 additional jobs added as a result of the loan. But the steady growth in pre-loan employment suggests the selection bias in this analysis that we already pointed out when presenting the summary statistics in Table 2 and Figure 2: the pre-loan growth of SBA loan recipients tends to be much higher than the average growth of U.S. businesses of comparable size.

The bottom panel of Figure 3 takes the additional step of employing the matching procedures described in Section 4.2, including both the full set of exact matching criteria and propensity score matching to restrict attention to treated and control firms that are similar for these variables, including measures of growth back to $t-4$. The implied impact of SBA loan receipt from $t-1$ to $t+5$ is slightly smaller at 2.5 (and still highly statistically significant) than with the previous analysis of the unmatched sample, but the matching appears effective at eliminating growth differences prior to $t-1$. However, neither of these estimates uses information on loan size, which we have shown (in Figure 1, above) varies substantially across SBA loans.

Table 4 and Figure 4 contain results where the treatment variable is defined as SBA loan amount, in real (2010) millions of dollars. The table contains results with the dependent variable specified as the change in the average employment from three years before to three years after the loan: $\Delta E_{ijt} = E_{ij,post} - E_{ij,pre}$, with $E_{ij,post} = (E_{ijt+1,i} + E_{ijt+2,i} + E_{ijt+3,i})/3$, and $E_{ij,pre} = (E_{ijt-1} + E_{ijt-2} + E_{ijt-3})/3$. For the same unmatched sample (except by calendar year and pre-loan firm size) used in the top panel of Figure 3, the relative employment growth of treated versus non-treated firms

is estimated as 4.1. When we also match on the other exact criteria (age category, industry, and state (for firms with 19 or fewer employees in $t-1$)), the estimate is 4.8, while when we in addition employ propensity score matching to control for pre-loan growth, the estimate falls to 3.7. This suggests that not controlling for pre-loan growth produces an upward bias in the estimates. The top panel of Figure 4 contains the full dynamics based on estimation for each year relative to $t-1$ ($\Delta E_{ijt} = E_{ijt+s} - E_{ijt-1}$ for $s = -5, -4, \dots, 4, 5$) for the matched sample. By contrast with the bottom panel of Figure 3, which uses the same sample but the binary treatment dummy, these loan amount results imply that faster growth in the pre-loan period is associated with a larger loan amount, which again suggests that the analysis has failed to eliminate the selection of loan amount on previous growth (even if it seems to have eliminated selection into binary treatment).

We next add treatment-control group fixed effects (FE), the α_j from Section 4.2, to reduce heterogeneity and carry out estimation only within these groups. The result in the bottom panel of Figure 4 show a post-loan rise implying an average of 4.2 extra employees by $t+5$, but prior to the loan year there are only tiny, statistically insignificant differences associated with loan amount. The success of the FEs in eliminating this manifestation of selection bias leads us to include them in the remaining specifications. But this result does not completely allay all concerns about selection bias. As discussed in Section 4.1, there could still be shocks to demand, productivity, costs, or ideas that affect SBA loan amounts but are not reflected in pre-loan growth differences. To handle such shocks, we now turn to results using the instrumental variables based on SBA loan access developed in Section 4.3. After examining the estimated dynamics from $t-5$ to $t+5$ for each of these variables, we also present results for specifications controlling for local industry growth and local availability of conventional loans.

Figure 5 contains results for dynamics for the five IV specifications developed in Section 4.3. As made clear from the 99 percent confidence intervals, the results are less precisely estimated than the OLS specifications, but the estimates for the post period are always statistically significantly different from zero. The range of point estimates implies employment effects per million dollars of loans by $t+5$ from 5.5 to 9, which are all higher than the OLS (with FE) estimate of 4.2 in Figure 4.

This suggests that any selection bias in the OLS is negative rather than positive, i.e., that firms receiving loans tend to have unobserved prospects that are relatively worse than control firms. Note also that in nearly all cases, the post-loan dynamics show a steady upward trend in relative employment of treated versus control firms through $t+5$, suggesting that the loan produces longer term growth rather than just a temporary fillip followed by a return to previous trend.

The pre-loan dynamics differ quite a bit across the IV specifications in Figure 5, however, with some cases of statistically significant coefficients. The most persuasive dynamics in showing no statistically significant coefficients in the pre-loan period and no evidence of a pre-existing trend is the *BranchShareCounty*SumSharesIV_{ct}*. For this instrumental variable we therefore report a full set of estimates in the text, while the results for the others can be found in the appendix.

Continuing to examine robustness further, we also include controls for local growth and for conventional financial access in both the OLS and IV specifications. These controls address the possibilities that SBA loan recipients are systematically located in counties with higher growth potential and where banks (including PLP lenders) think small business lending is more profitable. For ease of presentation of all the coefficients, we define the dependent variable here

as the difference in 3-year averages ($\Delta E_{ijt} = E_{ij,post} - E_{ij,pre}$). Our measure of local growth is the employment growth rate in the 4-digit industry, county, and time period; thus, these estimates can be interpreted as reflecting the impact of SBA loan amount on the deviation of the treated firm's growth from the industry-county-year period average. If SBA loan amount and firm growth are both correlated with this local growth measure, this would suggest a possible omitted variable bias. Similarly, if SBA loan amount, firm growth, and local presence of PLP lenders are correlated with local financial access more generally, there may again be an omitted variable problem. Our measures of local access to conventional loans involve two different county-level variables reported in the Community Reinvestment Act (CRA): the volume of CRA loans under \$1 million and the volume of CRA loans to firms with revenue under \$1 million. Neither of these is a perfect measure, insofar as some SBA loans are over \$1 million and some SBA recipients have revenue over \$1 million, but these variables are the best available for measuring access to conventional finance at the local level that might be correlated with SBA loan availability. We scale these variables by the number of firms in the LBD with positive employment in the county-year, so they represent alternative measures of the average CRA loan per business.

The CRA data are available only for a shorter time period, from 1996 onwards, so to keep a consistent sample we impute the 1996 values in the years 1991-1995.²⁶ None of the estimates is affected to any considerable extent by the controls. Indeed, the addition of these controls raises the SBA loan amount coefficient slightly. These results suggest that the estimates of the employment growth effect of SBA loans are not driven by local differences in labor demand or credit supply.

This pattern is very similar for the other four instrumental variables, so Table 6 contains results for basic specifications without the local growth and finance controls to save space. By comparison with the OLS estimate of 3.0 (in the first column of Table 5), the IV estimates range from 2.8 to 4.2 additional employees per million dollars of SBA loan in the first three post-loan years. Thus, the IV estimates bracket the OLS and are quantitatively similar. They are also similar to the results from less complex methods such as the simple regression on the matched sample, which yielded a coefficient of 3.7, and not very far from the largest estimate of 4.8 based on exact but not propensity score matching, both of which we show in Table 4. Taking the average of the coefficients for $t+1$, $t+2$, and $t+3$ from Figures 4 and 5 also yields similar results for the OLS and IV estimates. The trajectories are somewhat different in the last two years as we saw, however, resulting in larger estimated effects by $t+5$, ranging from 4.2 to 9, with most of them close to 6.

Note that the F-statistics, reported towards the bottom of Tables 5 and 6, imply strong statistical relationships between each of the instrumental variables and the endogenous SBA loan amount variable. The F's range from 600 to 2,484, easily clearing any conventional threshold for significance. The 1st-stage coefficients on the instrumental variables are also substantial, again implying a meaningful relationship with SBA loan amount. For example, a one-standard-deviation increase in $BranchShareCounty*SumSharesIV_{ct}$ is associated with a \$16,000 larger loan, and a county with at least one PLP lender is associated with a \$77,000 higher loan than one with none. Also, the F-statistics are only slightly smaller when including the additional controls

²⁶ As a robustness check to investigate whether results are sensitive to these imputations, we have estimated the regressions two additional ways. In both sets we include only the treatment years 1997-2009. In the first set of regressions, the actual CRA numbers are used in all years, while in the second set 1996-2000 values are imputed using 2001 values. The results are very similar.

for county growth and CRA lending. This is notable, because if PLP lenders were making location decisions based on the potential for small business lending, one would expect the inclusion of these controls to significantly reduce the explanatory power of the instrumental variables for local PLP lender presence.

5.2 Extensions

In this section we explore the heterogeneity of the treatment effect by broad loan type categories, the business cycle, and local conventional loan availability. We also examine the robustness of the results to different control samples and to accounting for exit, and we provide estimates of the extent of displacement effects, the magnitude of overall job creation, and the cost to the government per job that are implied by our estimated program impacts.

The 7(a) and 504 loan programs are similar in providing guarantees, but as noted above they also have differences. While the results above are based on pooled samples and estimated using a single coefficient for both programs, Table 7 displays the results from separate regressions for each program. The estimates, particularly those with the IV, are similar to those in Table 5, but they imply slightly higher impacts per million dollars for the 504 program. The difference between the loan amount coefficients in the IV regressions is not statistically significant, however.

The extent to which SBA loan programs foster job creation may be of particular interest during times of high unemployment, when they might function as cyclical stabilizers. With this motivation, Table 8 contains results when we permit the loan amount coefficient to vary with two alternative measures of the business cycle: a dummy for years when the national unemployment rate is below the median during the treatment years 1992-2009, and a similar dummy for below-median state unemployment. The results are inconsistent across OLS and IV methods. The OLS results imply a counter-cyclical effect, insofar as the loans are estimated to have a smaller effect when unemployment is low and a larger effect when it is high; the difference is not large, however. Except for the $IndicatorIV_{ct}$ with national unemployment, all the IV interactions with the low unemployment dummies are positive, implying a pro-cyclical effect, and they are statistically significant for the $SumSharesIV_{cb}$, $AvgLoanAmount * SumSharesIV_{cb}$, and $AvgLoanAmount * BranchShareCounty * SumSharesIV_{ct}$ specifications. One interpretation of the difference between the OLS and IV results is that the OLS reflects cyclically varying selection bias such that firms with relatively good unobserved prospects tend to receive SBA loans during recessions and those with worse prospects tend to receive SBA loans during expansions.

Table 9 shows results when we permit the loan amount effect to vary with local credit conditions. We use the two CRA measures of the volume of loans per firm in the county-year that were employed in Table 5, for all loans under \$1 million, and for all loans to businesses with revenue less than \$1 million. The estimation results suggest smaller employment effects of SBA loans when local credit is more plentiful, intuitively suggesting that the marginal benefit of the intervention is declining when credit is looser.

An important concern about the estimates of employment changes at firms receiving loans is the implicit assumption the program has no effect on non-treated firms used as controls in the analysis.²⁷ Because only a tiny fraction of firms in the U.S. receive SBA-backed loans,

²⁷ The program evaluation literature sometimes refers to this as the “stable unit treatment value assumption” (SUTVA) (Imbens and Wooldridge 2009).

this assumption is plausible.²⁸ But it is nevertheless possible that even if treated firms grow as a result of loan receipt, the program creates general equilibrium effects, or spillovers on other firms. Spillovers may be positive if the loan enables innovation that is somehow copied or imitated by other firms, or if suppliers or customers benefit together with the loan recipient. They could also be negative if they are displacement effects that reduce employment at non-treated firms that compete with the treated in product and labor markets. In either case, the total job creation – including these indirect effects as well as the direct effect – would differ from the direct effect we have estimated. Estimating such general equilibrium effects is intrinsically difficult, and it is largely ignored in the program evaluation literature (e.g., Heckman et al., 1999). Positive spillovers would imply that our estimates of the direct effect are lower than the total effect, and therefore we focus attention here on the possibility of negative displacement. If these result from product market competition, where loan receipt gives the beneficiary an advantage over its competitors, then negative effects should be larger for nearby firms in industries with a narrow geographic scope.

To assess this possibility, we distinguish a non-tradable sector and focus on treated firms for which we have two control groups: the first is matched controls within the same county as the treated firm, and the second consists of controls in counties where there are no firms in the same industry receiving SBA loans that year.²⁹ Results for all industries and for the non-tradable from dividing the sample by are shown in Table 10. For all sectors together, we find virtually identical coefficients with the two control groups. But the difference is much greater when we restrict attention to the non-tradable sector, where the employment growth associated with \$1 million of loans is 2.9 relative to within-county controls versus 2.6 relative controls from non-SBA county-industries (though the difference is not statistically significant). These differences are inconsistent with correlated demand shocks among firms in closer geographic proximity. The pattern could result from displacement, however, as nearby controls grow more slowly as a result of the loan. But the difference is not large, implying only a small role for displacement – on the order of 10 percent of the employment effect.

The analysis so far assumes no differences in survival rates between treated firms and controls, although the SBA frequently refers to business survival as a performance measure, and access to loans may well affect survival. The direction of the effect is not certain, however, because while more finance may help a business through hard times, increased leverage and possible over-extension may create greater vulnerability. Nor is the measurement of survival unambiguous, and any disappearance from the database is classified as an exit. Though great effort has been made to link establishments across time in the LBD, we cannot always distinguish bankruptcy and other genuine shutdowns from buy-outs or reorganizations that lead

²⁸ The annual number of disbursed SBA loans ranges from 22,586 in 1992 to 96,458 in 2007, compared to 4,689,051 and 5,520,998 firms with positive employment in the LBD in those years.

²⁹ Recall that all controls are matched on 4-digit industry; in the second control group they are included only when they are in counties where no firm in that industry received an SBA loan that year, so they should not be influenced by a displacement effect. Following Mian and Sufi (forthcoming), the non-tradable sector is defined to include the following retail trade industries: grocery stores; specialty food stores; beer, wine, and liquor stores; health and personal care stores; gasoline stations; clothing stores; shoe stores; jewelry, luggage, and leather goods stores; sporting goods, hobby, and musical instrument stores; book, periodical, and music stores; department stores; other general merchandise stores; florists; office supplies, stationery, and gift stores; used merchandise stores; other miscellaneous store retailers; full-service restaurants; limited-service eating places; special food services; drinking places (alcoholic beverages); automobile and other motor vehicle dealers; automobile parts accessories and tire stores; furniture stores; home furnishing stores; and electronics and appliance stores in the non-tradable sector.

to a change in the identifying code in the LBD. As some of these outcomes represent business failure, others reflect success, and some level of exit is a normal feature of a dynamic economy, the analysis of exit is thus also not as clear normatively as our analysis of employment effects.

With these qualifications in mind, we are nonetheless interested to ascertain the degree to which our results might be driven by exit effects. For the matched sample before imposing the constraint that the firm must have positive employment in $t+1$, $t+2$, and $t+3$, exit rates by $t+3$ are 24.2 percent for treated firms and 23.3 percent for controls (kernel weighted, as in the regressions, for greatest comparability). A crude comparison therefore suggests only a slight difference in survival rate associated with treatment, but this does not take into account the timing of exit or the size of firms exiting. Assuming exit represents job loss, then if exit is more common among loan recipients, our earlier results are overstated in ignoring the employment decline associated with exit. On the other hand, if SBA-backed loans raise survival, our earlier results could be understated. To distinguish these alternatives, we impute a zero value for employment following exit and re-estimate. The loan amount coefficients, shown in Table 11, are very similar, only slightly smaller than those without the imputations in Table 5, and the differences are not statistically significant, so we conclude that different patterns of exit are unlikely to play an important role in our results. Further analysis of exit effects, including the characteristics of survivors and exitors, could be of considerable interest, but we leave it for future research.

Finally, we can provide rough calculations of the overall job creation and cost per created job that the estimates suggest are attributable to SBA loans. To estimate employment effects through $t+5$ over the entire period of our sample, we may apply the coefficients in Figures 4 and 5 to the total volume of loans over the period we study. For fiscal years 1992-2007, the total volume is \$230,358 million of 7(a) and 504 loans, in 2010 dollars. The results, in Figure 6, show for $t+5$ a range of 1.2 to 2.1 million jobs created by SBA loans estimated with the five IV specifications, and a distinctly lower 1.0 million when estimated with OLS. For these same loans, the applicants predicted they would create or retain 5.6 million jobs as a result of the loan.³⁰ Thus, our estimates are significantly less than those anticipated by applicants (and those reported by the SBA), but they are not an entirely different order of magnitude, ranging from 21 to 37 percent, depending on the particular IV specification. If we subtract 10 percent due to displacement (an upper bound in the sense it applies to the non-tradable sector), net job creation would be about 0.9 to 1.9 million over this period.

It bears emphasis that our study is not a cost-benefit analysis. It does not estimate the full benefits of the program, which would include producer surplus of borrowers, lenders, and workers; possible consumer surplus (if loans help firms to produce at lower cost, resulting in lower prices); possible positive spillovers into other sectors; and any external effects of increased employment for society or the government budget. Our estimates do permit us, however, to calculate a rough range on the cost per job created by the SBA programs. The main costs are default and administration, while SBA receives revenue from guaranty fees. Our own calculations from the SBA loan database yield \$18.934 billion (in 2010 dollars) in loan default chargeoffs on all 7(a) and 504 loans issued in 1992-2007, the years in our sample for the effects

³⁰ The SBA's estimate of the number of jobs supported by the programs is based on the sum of the predicted jobs created and retained from the loan applications, but the exact methodology is not explained and has apparently changed over time. SBA (2011) reports a "change in methodology" and removal of cancelled loans from the calculations (p. 11); SBA (2012) reports "a filter applied to reduce outliers" (p. 4). We also exclude cancelled loans, but we do not know the SBA procedure regarding outliers.

through $t+5$.³¹ SBA (2010) reports \$138,601 million in administrative expenses for all its business loan programs in 2009. This cost is not available for earlier years, but if we assume that the administrative expenses were the same in real terms every year, the total administrative expenses in 1992-2007 are \$2.262 billion in 2010 dollars.³² When we apply the SBA guaranty fees that have been in place since at least August 1, 2008 (the effective date of the earliest available Standard Operating Procedure manual available on SBA's website) to the 1992-2007 loans, the total amount of fees is \$3.643 billion in 2010 dollars. Adding losses from defaults and administrative expenses and subtracting fees yields a total cost estimate of \$17.553 billion. Using the results from the OLS and IV specifications, Figure 7 shows the implied cost of \$8,200 - \$18,000 per job created within five years after loan receipt.

6. Conclusion

Our estimates of the effects of the SBA loan programs in this paper are based on an unusual linking of administrative and census data and an application of econometric methods originally designed for evaluating job training interventions. We exploit the large size and completeness of the data to combine matching and regression methods. The first step is to match exactly on firm age, industry, year, and pre-loan size, and the second is to carry out kernel-based matching on propensity scores estimated as a function of four years of employment history and other variables. We use the matched sample to estimate program effects with treatment-control group fixed effects regressions and with instrumental variables based on arguably exogenous variation in the supply of SBA loans.

The estimation results imply that \$1 million of loans raises loan recipient employment by 3-4 jobs on average during the first three post-loan years. This basic result is robust across OLS and IV specifications and to controlling for detailed industry-county growth and county-level credit conditions. Estimating the effects farther out, at five years after loan receipt, the implied job creation ranges more widely, at 4.2 estimated by OLS, and from 5.5 to 9, depending on the particular IV specification. The uniformly positive impact of increased capital access on employment suggests these two factors of production are gross complements rather than substitutes.

Extending these results to consider several dimensions of heterogeneity, our analysis finds little difference between the estimated employment effects of the 7(a) and 504 programs, and for most IV specifications the weight of the evidence suggests a pro-cyclical effect. A clearer finding is that employment effects of SBA loans tend to be smaller where local credit is more plentiful, suggesting that the marginal benefit of the intervention is declining with general credit availability conditions. We also provide a preliminary investigation of possible displacement of competing firms by focusing on non-tradable industries, comparing estimated effects using controls in the same county with estimates using controls from counties in which there was no SBA loan in the same industry and county that year; the results imply a displacement effect on the order of 10 percent of the total. While our basic analysis considers only surviving firms (through the post-loans in a particular estimation), survival and exit are clearly important issues, so, as a robustness check, we estimate similar equations with zeroes

³¹ The chargeoff amount is the gross balance of the defaulted loan minus the gross amount that could be recovered through liquidation. These are defaults through May 31, 2015. Some of the loans still outstanding as of that date may later default.

³² This is likely an overestimate, since the programs expanded significantly during the period, and some of the expenses were for other business loan programs besides 7(a) and 504 (e.g., the SBIC program and direct loans).

imputed for firms after exit; the results differ little, suggesting that our basic results are not driven by exit.

Finally, our rough estimates of the total job creation from the SBA 7(a) and 504 loans issued in 1992-2007 range from 1.0 to 2.1 million, with a cost per job created ranging from \$8,200 - \$18,000. This range is far to the left of the usual estimated cost per job of government programs. Neumark (2013), for instance, reports the cost from fiscal stimulus under the American Recovery and Reinvestment Act (based on estimates from Congressional Budget Office 2010) to lie in the range of \$158,000 to \$407,000, and the cost per job from a hiring credit under the New Jobs Tax Credit (based on Bartik and Erickcek 2010) at \$37,500-\$75,000.³³ A full assessment of costs should account for possible distortions, such as those highlighted by Hurst and Pugsley (2011), and opportunity costs of the funds, but these figures suggest that alleviating financial constraints through SBA loan programs may be a relatively low-cost tool to generate employment.

If these estimates are correct, they imply that credit constraints impeded growth of small businesses prior to receiving the SBA loans. The estimates pertain to the effect of the program, not directly to credit access more generally. As we have shown, however, the SBA loans are only a small fraction of all small business loans in the U.S. economy (less than 1 percent in number and less than 5 percent in volume). We have argued that non-treated firms in the control group, which is constructed to match the pre-loan growth of treated firms, are more likely to receive these non-SBA loans. In this case, our estimates of the SBA loan impact can also be interpreted as lower bound estimates of the effect of financial access by similar credit-constrained firms more generally.

³³ Other estimates based on stimulus spending under the ARRA include Feyrer and Sacerdote (2012), and Wilson (2012), who report of the costs per job of \$107,000-\$400,000 and \$125,000, respectively. Chodorow-Reich et al. (2012) report an increase of 3.8 job years for additional \$100,000 in Medicaid spending under the ARRA through July 2010, implying a cost of \$26,000 per job year. If we restrict the horizon to the first post-loan year, for comparability, our estimates imply a cost of \$10,100-\$20,800 per job year; of course we also show sustained employment increases beyond the first post-loan year.

References

- Agarwal, Sumit, and Robert Hauswald, "Distance and Private Information in Lending," *Review of Financial Studies*, Vol. 23(7), 2757-2788, 2010.
- Amel, Dean F., and Kenneth P. Brevoort, "The Perceived Size of Small Business Banking Markets." *Journal of Competition Law and Economics*, Vol. 1(4), 771-784, 2005.
- de Andrade, Flavio, and Deborah Lucas, "Why Do Guaranteed SBA Loans Cost Borrowers So Much?" Working Paper, Northwestern University, 2009.
- Bach, Laurent, "Are Small Businesses Worthy of Financial Aid? Evidence from a French Targeted Credit Program." *Review of Finance*, Vol. 18, 877-919, 2014.
- Banerjee, Abhijit V., and Esther Duflo, "Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program." *The Review of Economic Studies*, Vol. 81(2), 572-607, 2014.
- Bartik, Timothy J., and George Erickcek, "The Employment and Fiscal Effects of Michigan's MEGA Tax Credit Program." Upjohn Institute Working Paper No. 10-164. 2010.
- Beck, Thorsten, "The Econometrics of Finance and Growth." *Palgrave Handbook of Econometrics* (Terence C. Mills and Kerry Patterson, eds.), Palgrave Macmillan, 2009.
- Brevoort, Kenneth P., and Timothy H. Hannan, "Commercial Lending and Distance: Evidence from Community Reinvestment Act Data," *Journal of Money, Credit, and Banking*, Vol. 38(8), 1991-2012, 2006.
- Birch, David L., *Job Creation in America: How Our Smallest Companies Put the Most People to Work*. New York: Free Press, 1987.
- Chodorow-Reich, Gabriel, Laura Feiveson, Zachary Liscow, and William Gui Woolston, "Does State Fiscal Relief during Recessions Increase Employment? Evidence from the American Recovery and Reinvestment Act," *American Economic Journal: Economic Policy*, 4(3), 118-145, 2012.
- Clementi, Gian Luca, and Hugo A. Hopenhayn, "A Theory of Financing Constraints and Firm Dynamics." *Quarterly Journal of Economics*, Vol. 121(1), 229-265, 2006.
- Congressional Budget Office (CBO), "Estimated Impact of the American Recovery and Reinvestment Act on Employment and Economic Output from April 2010 through June 2010." CBO, 2010.
- Craig, Ben R., William E. Jackson III, and James B. Thomson, "The Economic Impact of the Small Business Administration's Intervention in the Small Firm Credit Market: A Review of the Research Literature." *Journal of Small Business Management*, Vol. 47(2), 221-231, 2009.
- Davis, Steven, John Haltiwanger, and Scott Schuh, *Job Creation and Destruction*. Cambridge, Mass: The MIT Press, 1996.
- Demetriades, Panicos O., and Khaled A. Hussein, "Does financial development cause economic growth? Time-series evidence from 16 countries." *Journal of Development Economics*, Vol. 51(2), 387-411, December 1996.

- Federal Financial Institutions Examination Council (FFIEC), *A Guide to CRA Data Collection and Reporting*, 2015. Downloadable from http://www.ffiec.gov/cra/pdf/2015_CRA_Guide.pdf, accessed July 30, 2015.
- Feyrer, James, and Bruce Sacerdote, "Did the Stimulus Stimulate? Real Time Estimates of the Effects of the American Recovery and Reinvestment Act," mimeo, June 21, 2012.
- Glennon, Dennis, and Peter Nigro, "An Analysis of SBA Loan Defaults by Maturity Structure." *Journal of Financial Services Research*, Vol. 28(1/2/3), 77-111, 2005.
- Haltiwanger, John, Ron Jarmin, and Javier Miranda, "Who Creates Jobs? Small vs. Large vs. Young." *Review of Economics and Statistics*, Vol. 95(2), 347-361, May 2013.
- Heckman, James, and Joseph V. Hotz, "Choosing among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training." *Journal of the American Statistical Association*, Vol. 84(408), 862-74, December 1989.
- Heckman, James J., Hidehiko Ichimura, and Petra E. Todd, "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *Review of Economic Studies*, Vol. 64(4), 605-654, 1997.
- Heckman, James J., Hidehiko Ichimura, and Petra E. Todd, "Matching as an Econometric Evaluation Estimator." *Review of Economic Studies*, Vol. 65(2), 261-294, 1998.
- Heckman, James J., Robert J. LaLonde, and Jeffrey A. Smith, "The Economics and Econometrics of Active Labor Market Programs." In *Handbook of Labor Economics*, Vol. 3A. Edited by Orley Ashenfelter and David Card. Elsevier Science B.V., Amsterdam, 1999.
- Hubbard, R. Glenn, "Capital-Market Imperfections and Investment." *Journal of Economic Literature*, Vol XXXVI, 193-225, March 1998.
- Hurst, Erik, and Benjamin Wild Pugsley, "What Do Small Firms Do?" *Brookings Papers on Economic Activity*, 73-118, Fall 2011.
- Imbens, Guido W., and Jeffrey M. Wooldridge, "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature*, Vol. 47(1), 5-86, 2009.
- Jarmin, Ronald S., and Javier Miranda, "The Longitudinal Business Database." CES Working Paper 02-17, 2002.
- King, R.G. and Ross Levine, "Finance and Growth: Schumpeter Might Be Right." *Quarterly Journal of Economics*, Vol. 108, 717-737, 1993.
- Lelarge, Claire, David Sraer, and David Thesmar, "Entrepreneurship and Credit Constraints: Evidence from a French Loan Guarantee Program." In *International Differences in Entrepreneurship*. Edited by Josh Lerner and Antoinette Schoar. University of Chicago Press, Chicago, 2010.
- Levine, Ross, "Finance and Growth: Theory and Evidence." In *Handbook of Economic Growth*, Philippe Aghion and Steven Durlauf (eds.). Netherlands: Elsevier Science, 2005.
- Mian, Atif R., and Amir Sufi, "What Explains the 2007-2009 Drop in Employment?" *Econometrica*, forthcoming.

Myers, Stewart C., and Nicholas Majluf, "Corporate Financing and Investment Decisions when Firms have Information that Investors do not have." *Journal of Financial Economics*, Vol. 13(2), 187-221, June 1984.

Neumark, David, "Spurring Job Creation in Response to Severe Recessions: Reconsidering Hiring Credits." *Journal of Policy Analysis and Management*, Vol. 32(1), 142-171, 2013.

Neumark, David, Brandon Wall, and Junfu Zhang, "Do Small Businesses Create More Jobs? New Evidence for the United States from the National Establishment Time Series." *Review of Economics and Statistics* 93(1), 16-29, 2011.

Rosenbaum, P., and D. B. Rubin, "Constructing a Control Group Using a Multivariate Matched Sampling Method that Incorporates the Propensity Score." *The American Statistician*, Vol. 39, pp. 33-38, 1985.

Small Business Administration (SBA), *Agency Financial Report, Fiscal Year 2010*, 2010. Downloadable from https://www.sba.gov/sites/default/files/aboutsbaarticle/afr_2010_final.pdf, accessed on July 30, 2015.

Small Business Administration (SBA), *Agency Financial Report, Fiscal Year 2012*, 2012. Downloadable from <https://www.sba.gov/sites/default/files/files/FY%202012%20AFR%2011%2015%202012%20.pdf>, accessed on July 29, 2015.

SBA, *Summary of Performance and Financial Information, FY 2011*, 2011. Downloadable from <https://www.sba.gov/sites/default/files/files/FY%202011%20Summary%20of%20Performance%20and%20Financial%20Information.pdf>, accessed on July 29, 2015.

SBA, *SOP 50 10 5(H), Lender and Development Company Loan Programs*, effective May 1, 2015. Downloadable from https://www.sba.gov/sites/default/files/sops/SOP_50_10_5_H_FINAL_FINAL_CLEAN_5-1-15.pdf, accessed July 29, 2015.

Whitehouse.gov, "Weekly Address: President Obama Calls for New Steps to Support America's Small Businesses," February 6, 2010.

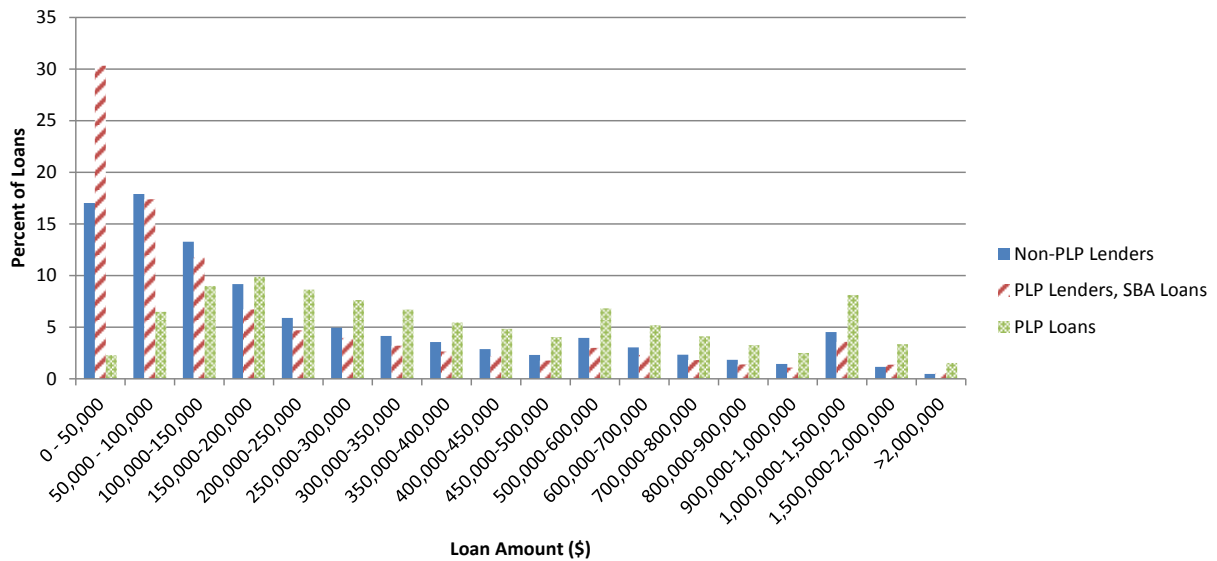
Wilson, Daniel J., "Fiscal Spending Jobs Multipliers: Evidence from the 2009 American Recovery and Reinvestment Act." *American Economic Journal: Economic Policy*, 4(3), 251-282, 2012.

Table 1: Bank Descriptive Statistics

	Mean for Non-SBA Banks	Mean Across Non-PLP SBA Banks	Mean Across PLP Banks
Number Of Banks	7,276	1,485	258
Number Of Branches Per Bank	6.4	16.6	255.9
Number Of Counties In Which The Bank Has At Least One Branch	2.1	4.2	36.0
Of Which, Number of Counties In Which At Least One Firm Has An SBA Loan From This Bank	N.A.	1.2	10.4
Of Which, Number of Counties In Which At Least One Firm Has A PLP Loan From This Bank	N.A.	N.A.	5.4
Number of SBA Loans	N.A.	6.5	123.1
Number of PLP Loans	N.A.	N.A.	36.9
Bank's Mean SBA Loan Amount	N.A.	253,130	328,330
Bank's Mean PLP Loan Amount	N.A.	N.A.	411,213
Mean SBA Loan Amount		286,288	297,010
Mean PLP Loan Amount		N.A.	460,822

Banks are defined as firms that have at least one establishment in the commercial banking sector or which issued at least one SBA loan in the year. Only establishments with positive employment in the year are included. For banks firms that did not issue any SBA loans in the year, the establishment counts include only those in the commercial banking sector. The numbers are means across 1991-2009 annual averages. Loan amounts are in 2010 dollars.

Figure 1: SBA and PLP Loan Size Distributions, Non-PLP vs. PLP Lenders



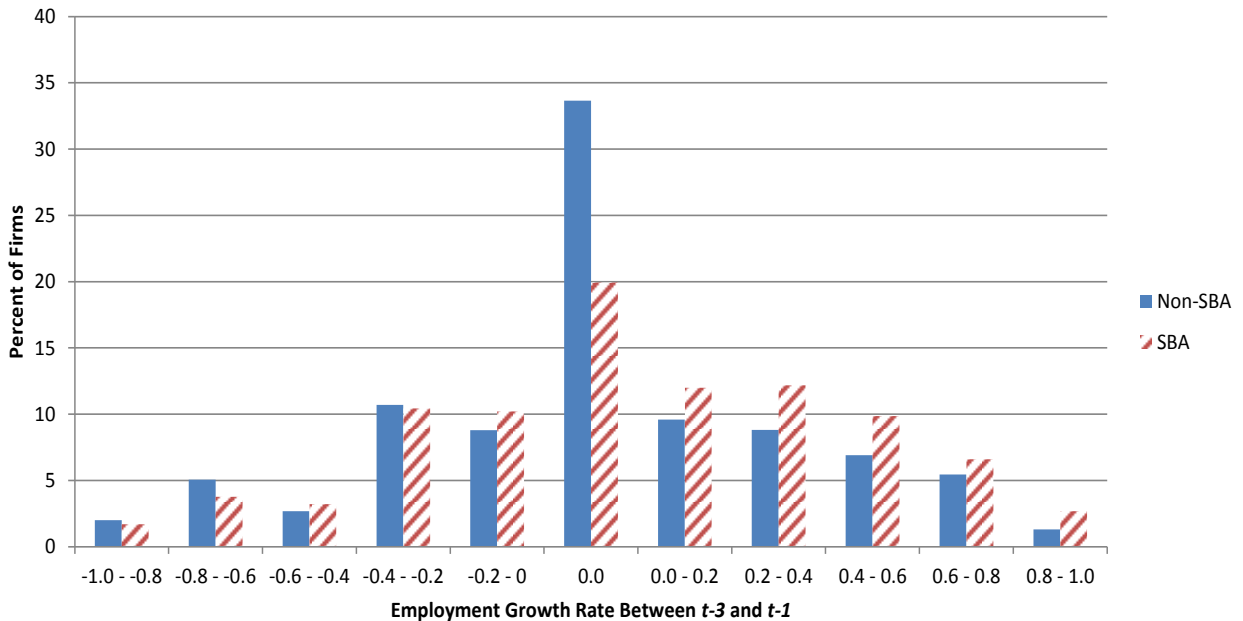
These are the distributions for all SBA loans in 1991-2009.

Table 2: Treated and Non-Treated Firm Descriptive Statistics

	All Non-Treated Firms	All Treated Firms
Employment _{t-1} Mean	19.85	14.33
Employment _{t-1} Standard Deviation	678.82	36.78
Employment _{t-1} Median	4	7
Mean Emp. Growth Between $t-3$ and $t-1$	0.023	0.116
Age Mean	10.91	8.04
Age Standard Deviation	8.34	7.21
Age Median	9	6
Percent by Sector:		
Construction	11.61	10.44
Manufacturing	5.57	10.33
Wholesale Trade	6.74	9.67
Retail Trade	16.07	20.70
Finance/Insurance/Real Estate	8.74	3.72
Services	45.83	40.17
Other	5.44	4.97

Non-treated firms are included in each year they appear in the LBD between 1992 and 2009, while treated firms are included only in the treatment year. Year t is the treatment year for treated firms. The variables come from the LBD.

Figure 2: Employment Growth Distribution Between $t-3$ and $t-1$, SBA Loan Recipients vs. All Non-Treated Firms



Non-treated firms are included in each year they appear in the LBD between 1992 and 2009, while treated firms are included only in the treatment year. Year t is the treatment year for treated firms. The variables come from the LBD.

Table 3: Instrumental Variables and Components Descriptive Statistics, 1991-2008

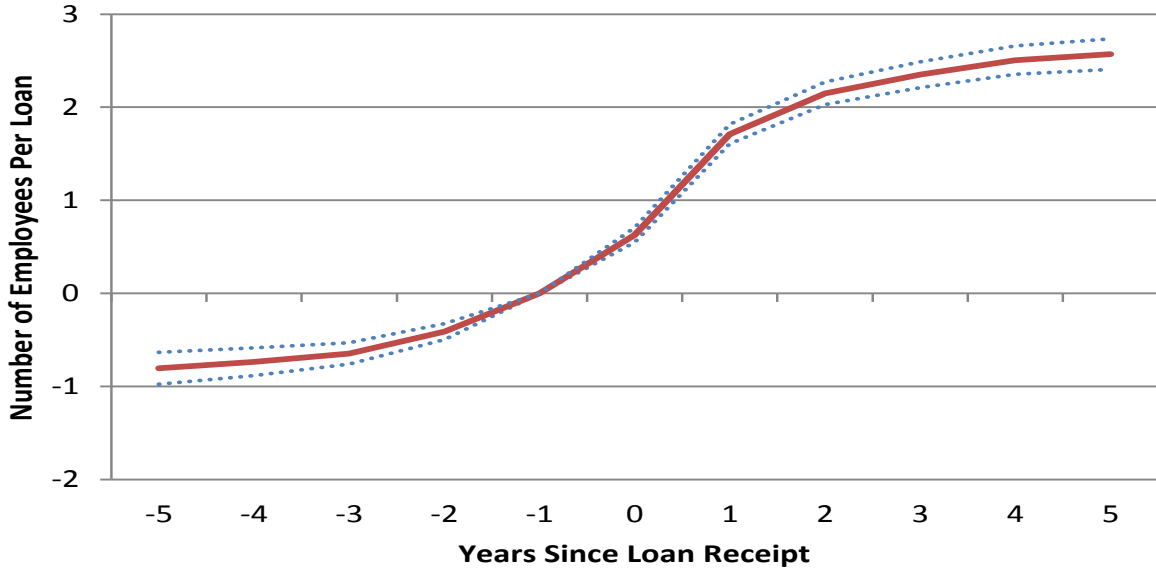
	Mean	Standard Deviation	50 th %
PLP Bank's Counties with PLP Loan as Share of Other Counties with Branch (Bank-County-Years)	0.146	0.171	0.10
PLP Bank's Share of County's Branches, Conditional on >0 (Bank-County-Years)	0.101	0.132	0.05
PLP Bank Average Loan Amount in Other Counties (2010 \$Millions), (Bank-County-Years)	0.365	0.224	0.30
<i>IndicatorIV_{ct}</i> (County-Years)	0.581		
<i>SumSharesIV_{ct}</i> (County-Years)	0.310	0.658	0.07
<i>BranchShareCounty*SumSharesIV_{ct}</i> (County-Years)	0.032	0.059	0.007
<i>AvgLoanAmount*SumSharesIV_{ct}</i> (County-Years)	0.108	0.238	0.02
<i>AvgLoanAmount*BranchShareCounty*SumSharesIV_{ct}</i> (County-Years)	0.010	0.020	0.00

Table 4: Treatment Effect Without and With Propensity Score Matching

	Sample Matching Only on Employment and Year, No Propensity Score Matching	Sample with Exact Matching, No Propensity Score Matching	Sample with Exact Matching and Propensity Score Matching
Loan Amount	4.075 (0.101)	4.804 (0.226)	3.669 (0.103)
Number of Observations	49,496,600	15,393,000	4,738,900
Number of SBA Firms	211,700	187,800	161,600

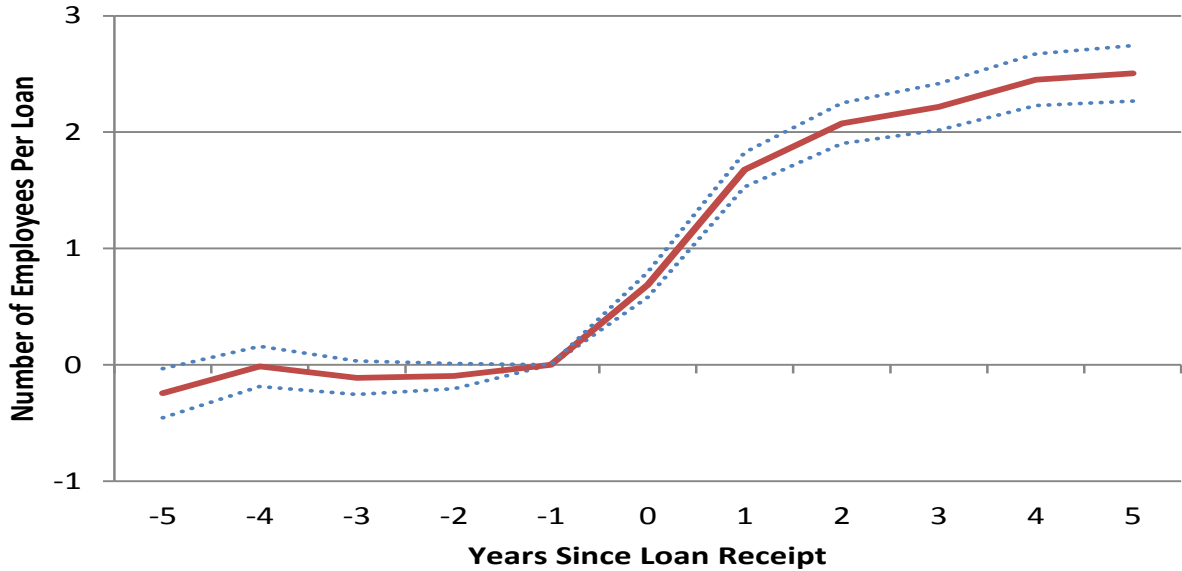
For each sample treated firms receive a weight of one. For the column one sample, each control firm is assigned a weight of the ratio of treated to control firms in the control firm's $t-1$ employment category-year cell (we use 41 employment categories). For the column two sample, each control firm appears once for each treated firm to which it is exact matched by employment (within +/- 10 percent), four-digit industry, treatment year age category, and year. The control firm receives a weight of the ratio of one over the number of control firms exact matched to the treated firm. The controls in the column three sample receive kernel weights from propensity score matching. The dependent variable is average employment in $t+1$ through $t+3$ minus average employment in $t-3$ through $t-1$, and loan amount (in millions of 2010 dollars) and year of loan receipt dummies are the independent variables. Standard errors, reported in parentheses, are cluster-adjusted by firm. The number of observations and SBA firms are rounded to the nearest 100 for disclosure avoidance.

Figure 3: Dynamics for Number of Employees Per Loan
 OLS, Matching on Employment and Year Only



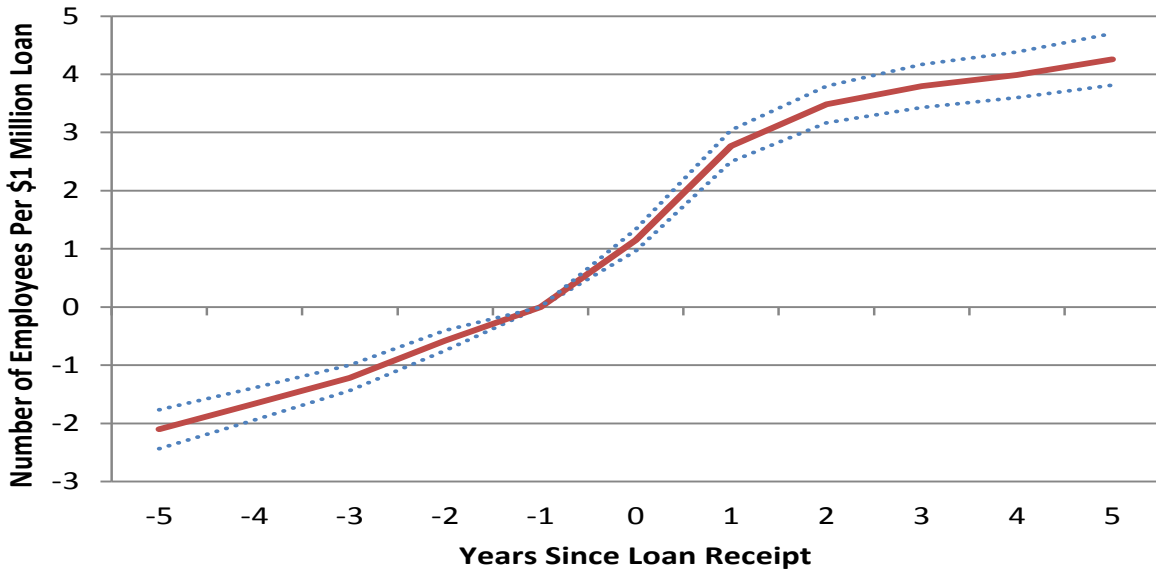
These are loan receipt dummy coefficients from OLS regressions with a dependent variable of the firm's employment in the respective year minus employment in year $t-1$. Age, age squared, and loan receipt year dummies are included as controls. Treated firms receive a weight of one. Each control firm is assigned a weight of the ratio of treated to control firms in the control firm's $t-1$ employment category-year cell (we use 41 employment categories). The sample is the same in the regressions in years t , $t+1$, $t+2$, $t+3$, $t+4$, and $t+5$. To be in the samples for the pre-treatment years, firms must have positive employment in the respective year, as well as in $t-1$ through $t+5$. The dotted lines are the bounds of the 99 percent confidence interval, based on standard errors clustered by firm.

OLS, Matched Sample

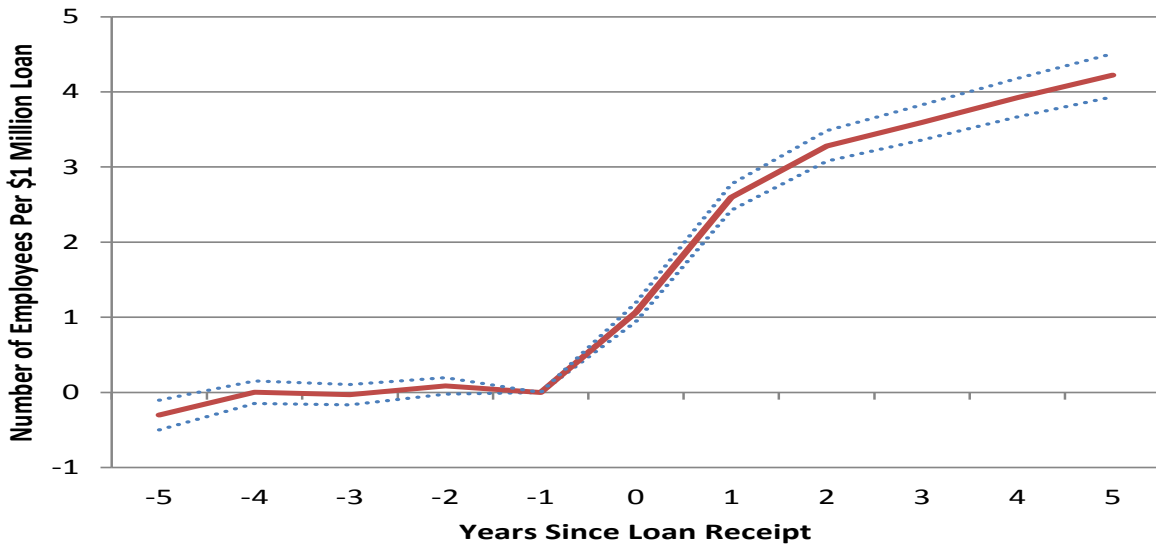


These are loan receipt dummy coefficients from OLS regressions with a dependent variable of the firm's employment in the respective year minus employment in year $t-1$. Age, age squared, and loan receipt year dummies are included as controls. The sample is after exact matching and propensity score matching, and kernel weights from the propensity score matching are used. The sample is the same in the regressions in years t , $t+1$, $t+2$, $t+3$, $t+4$, and $t+5$. To be in the samples for the pre-treatment years, treated firms and at least one control must have positive employment in the respective year, as well as in $t-1$ through $t+5$. The dotted lines are the bounds of the 99 percent confidence interval, based on standard errors clustered by firm.

Figure 4: Dynamics for Number of Employees Per \$1 Million Loan, OLS
 OLS, Matched Sample

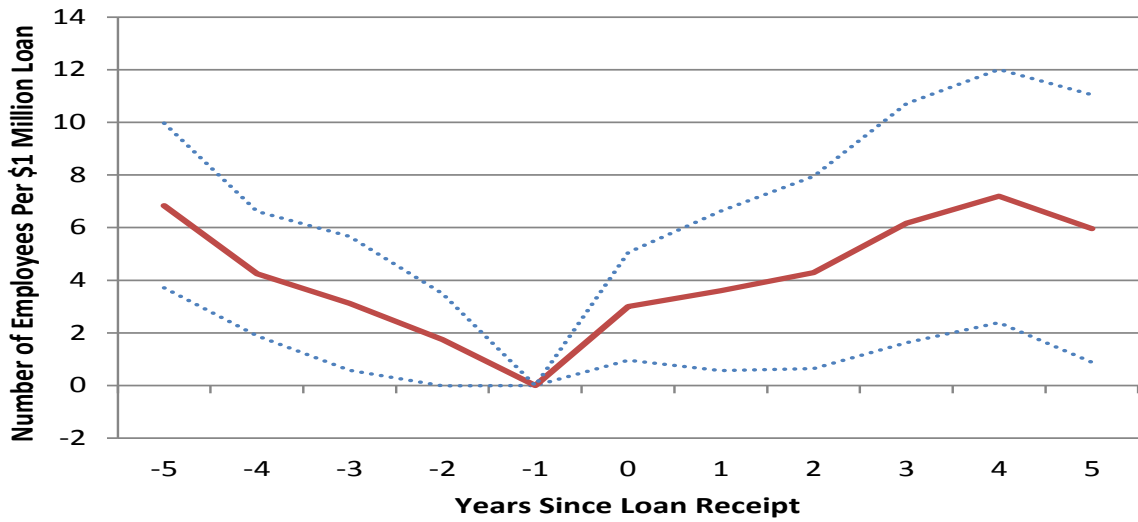


OLS with Treated Firm-Control Fixed Effects, Matched Sample

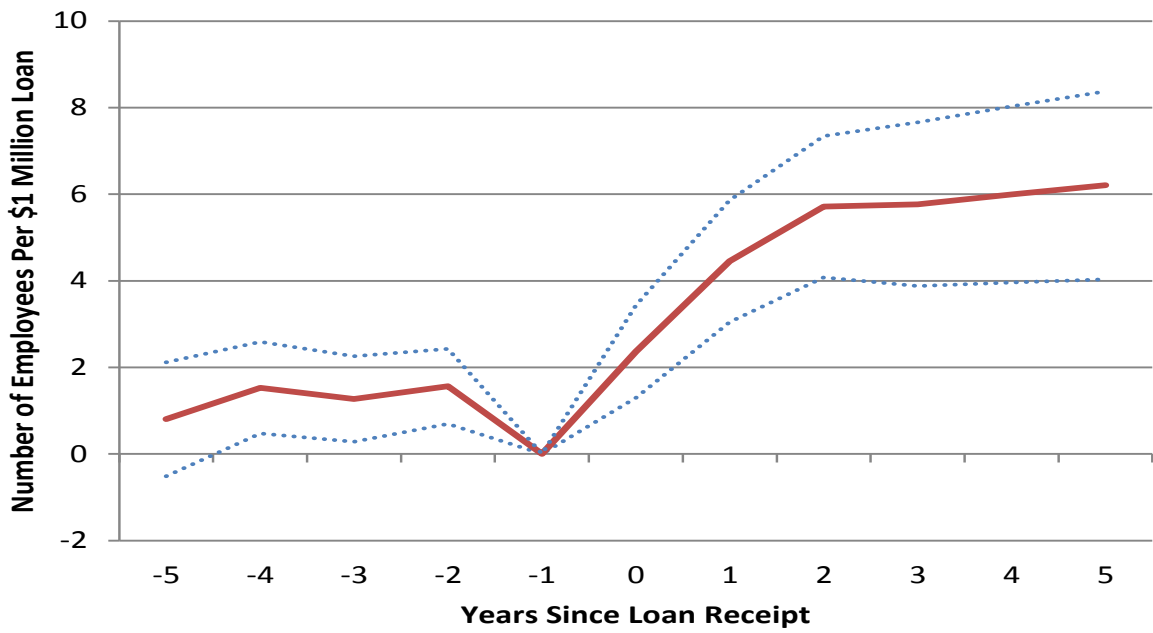


These are loan amount coefficients from OLS regressions with a dependent variable of the firm's employment in the respective year minus employment in year $t-1$. Loan amount is in millions of 2010 dollars. Age, age squared, and treatment year dummies are included in the regressions without fixed effects, and treatment year age, age squared, and treated firm-control fixed effects are included in the treated firm-control fixed effects regressions. The sample is the same in the regressions in years t , $t+1$, $t+2$, $t+3$, $t+4$, and $t+5$. To be in the samples for the pre-treatment years, treated firms and at least one control must have positive employment in the respective year, as well as in $t-1$ through $t+5$. The dotted lines are the bounds of the 99 percent confidence interval, based on standard errors clustered by firm.

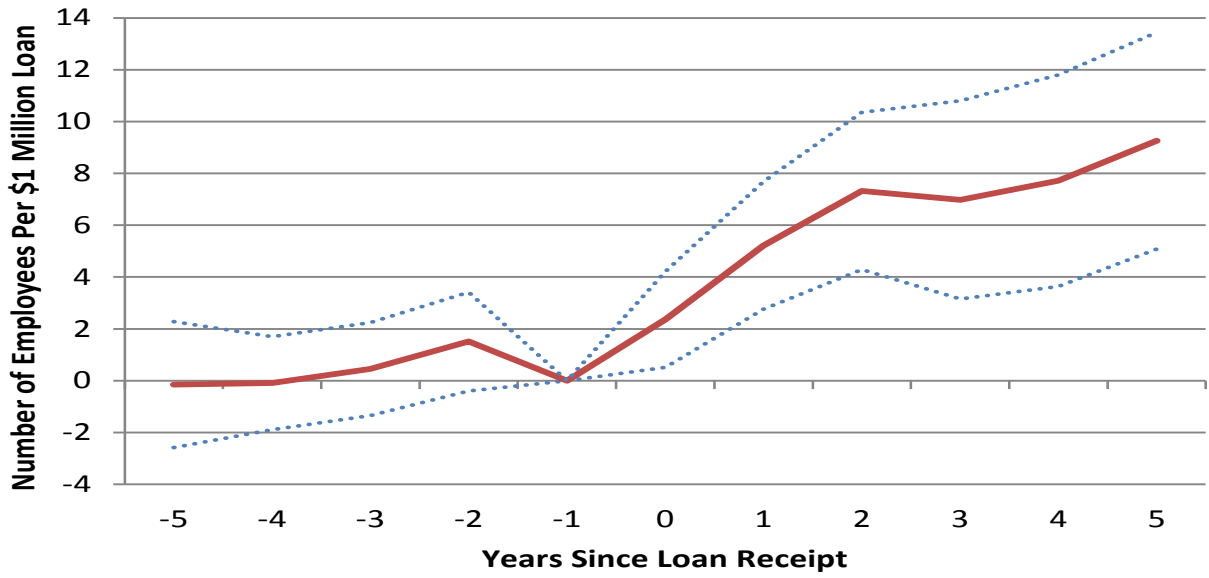
Figure 5: Dynamics for Number of Employees Per \$1 Million Loan, IV
IndicatorIV_{ct}



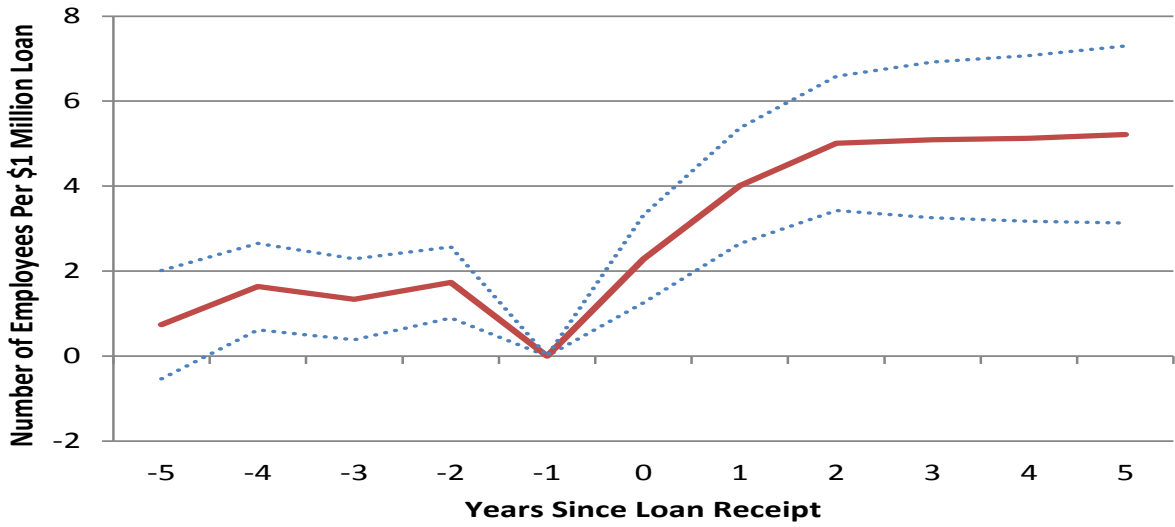
SumSharesIV_{ct}



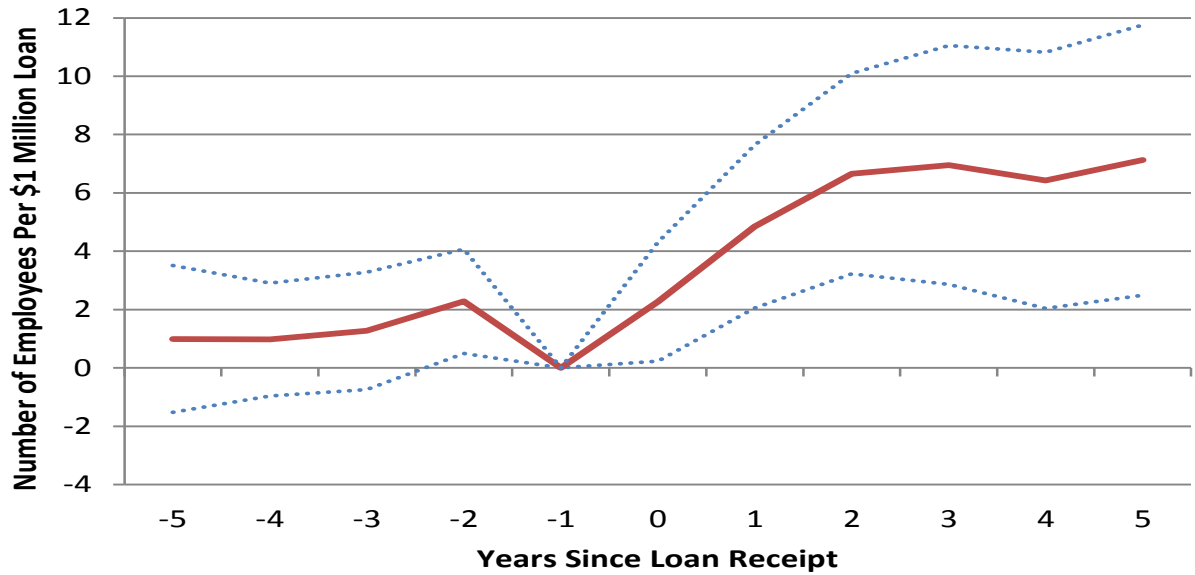
*BranchShareCounty*SumSharesIV_{ct}*



*AvgLoanAmount*SumSharesIV_{ct}*



$$\text{AvgLoanAmount} * \text{BranchShareCounty} * \text{SumSharesIV}_{ct}$$



These are loan amount coefficients from IV regressions with a dependent variable of the firm's employment in the respective year minus employment in year $t-1$. Loan amount is in millions of 2010 dollars. Treatment year age, age squared, and treated firm-control fixed effects are also included in the regressions. The sample is the same in the regressions in years t , $t+1$, $t+2$, $t+3$, $t+4$, and $t+5$. To be in the samples for the pre-treatment years, treated firms and at least one control must have positive employment in the respective year, as well as in $t-1$ through $t+5$. The dotted lines are the bounds of the 99 percent confidence interval, based on robust standard errors. 1st Stage F Statistics: *IndicatorIV_{ct}*: 533.77 ($t-5$), 624.67 ($t-4$), 691.38 ($t-3$), 681.89 ($t-2$), 754.98 ($t...t+5$). *SumSharesIV_{ct}*: 1761.93 ($t-5$) 1853.01 ($t-4$), 1907.42 ($t-3$), 1991.06 ($t-2$), 2154.04 ($t...t+5$). *BranchShareCounty*SumSharesIV_{ct}*: 512.76 ($t-5$), 583.08 ($t-4$), 605.20 ($t-3$), 599.50 ($t-2$), 639.92 ($t...t+5$). *AvgLoanAmount*SumSharesIV_{ct}*: 1794.83 ($t-5$), 1908.46 ($t-4$), 1952.42 ($t-3$), 2037.80 ($t-2$), 2180.85 ($t...t+5$). *AvgLoanAmount*BranchShareCounty*SumSharesIV_{ct}*: 393.74 ($t-5$), 425.37 ($t-4$), 429.06 ($t-3$), 449.41 ($t-2$), 450.99 ($t...t+5$).

Table 5: Base Specifications

	OLS		
Loan Amount	2.999 (0.067)	3.003 (0.067)	3.000 (0.067)
County-Industry Employment Growth		0.243 (0.088)	0.243 (0.088)
CRA Loans <=\$1m Count/LBD Firm Ratio		-0.211 (0.081)	
CRA Loans to Firms <=\$1m Rev. Count/LBD Firm Ratio			-0.416 (0.162)
	<i>BranchShareCounty*SumSharesIV_{ct}</i>		
Loan Amount	4.159 (1.054)	4.620 (1.152)	4.286 (1.062)
County-Industry Employment Growth		0.240 (0.087)	
CRA Loans <=\$1m Count/LBD Firm Ratio		-0.343 (0.115)	
CRA Loans to Firms <=\$1m Rev. Count/LBD Firm Ratio			-0.498 (0.167)
1 st Stage <i>BranchShareCounty*</i>	0.272 (0.009)	0.246 (0.009)	0.269 (0.009)
<i>SumSharesIV_{ct}</i>			
1 st Stage F Statistic	882.84	736.57	869.67
N	4,738,900	4,717,300	4,717,300
Number of SBA Firms	161,600	161,100	161,100

The number of observations and number of SBA firms are rounded to the nearest 100 for disclosure avoidance. The regressions are kernel-weighted. The dependent variable is average employment in $t+1$ through $t+3$ minus average employment in $t-3$ through $t-1$. Loan amount is in millions of 2010 dollars. Treatment year age, age squared, and treated firm-control fixed effects are also included in the regressions. Robust standard errors are reported in parentheses; in the OLS regressions they are cluster-adjusted by firm. County-industry employment growth, measured at the county-four-digit industry level (SIC industries for treatment years 1992-1998 and NAICS industries for treatment years 1999-2009), is the sum of county-industry employment in years $t+1$, $t+2$, and $t+3$ minus county-industry employment in $t-3$, $t-2$, and $t-1$, divided by $\frac{1}{2}$ the sum of county-industry employment in years $t-3$, $t-2$, $t-1$, $t+1$, $t+2$, and $t+3$. It excludes the employment of the firm in the regression to which the variable refers. The number of Community Reinvestment Act (CRA) loans in the respective category in the county in $t-1$ is divided by the county's number of firms with positive employment in the LBD in $t-1$. Each CRA/LBD variable is demeaned over the sample, using the kernel weights. Robust standard errors are reported in parentheses; in the OLS regressions they are cluster-adjusted by firm.

Table 6: Alternative Instrumental Variables

Loan Amount	3.911 (1.270)	3.531 (0.575)	2.828 (0.585)	2.764 (1.164)
1 st Stage <i>IndicatorIV_{ct}</i>	0.0769 (0.0026)			
1 st Stage <i>SumSharesIV_{ct}</i>		0.0204 (0.0004)		
1 st Stage <i>AvgLoanAmount*SumSharesIV_{ct}</i>			0.0516 (0.0010)	
1 st Stage <i>AvgLoanAmount*BranchShareCounty* SumSharesIV_{ct}</i>				0.561 (0.023)
1 st Stage F Statistic	857.69	2462.88	2484.16	599.87

The number of observations is 4,738,900, and the number of SBA firms is 161,600, rounded to the nearest 100 for disclosure avoidance. The regressions are kernel-weighted. The dependent variable is average employment in $t+1$ through $t+3$ minus average employment in $t-3$ through $t-1$. Loan amount is in millions of 2010 dollars. Treatment year age, age squared, and treated firm-control fixed effects are also included in the regressions. Robust standard errors are reported in parentheses.

Table 7: 7(a) vs. 504 Loans

	7(a)	504
OLS	2.526 (0.098)	3.219 (0.087)
IV	4.012 (2.074)	4.261 (0.976)
1 st Stage <i>BranchShareCounty*SumSharesIV_{ct}</i>	0.146 (0.006)	0.783 (0.039)
1 st Stage F Statistic	602.66	395.30
N	3,856,900	882,000
Number of SBA Firms	132,500	29,100

The regressions are kernel-weighted. The dependent variable is average employment in $t+1$ through $t+3$ minus average employment in $t-3$ through $t-1$. Loan amount is in millions of 2010 dollars. Treatment year age, age squared, and treated firm-control fixed effects are also included in the regressions. Robust standard errors are reported in parentheses; in the OLS regressions they are cluster-adjusted by firm. The number of observations and SBA firms are rounded to the nearest 100 for disclosure avoidance.

Table 8: Interactions with Business Cycle Indicators

	OLS	
Loan Amount	3.017 (0.067)	3.012 (0.067)
Below-Median State Unemployment in Loan Year	0.722 (0.196)	
Loan Amount*Below-Median State Unemployment	-0.867 (0.135)	
Loan Amount*Below-Median National Unemployment		-0.723 (0.134)
	<i>BranchShareCounty*</i>	
	<i>SumSharesIV_{ct}</i>	
Loan Amount	4.101 (1.062)	4.117 (1.046)
Below-Median State Unemployment in Loan Year	-0.633 (0.698)	
Loan Amount*Below-Median State Unemployment	3.503 (2.247)	
Loan Amount*Below-Median National Unemployment		1.570 (2.091)

The number of observations is 4,738,900, and the number of SBA firms is 161,600, rounded to the nearest 100 for disclosure avoidance. The regressions are kernel-weighted. The dependent variable is average employment in $t+1$ through $t+3$ minus average employment in $t-3$ through $t-1$. Loan amount is in millions of 2010 dollars. Treatment year age, age squared, and treated firm-control fixed effects are also included in the regressions. Robust standard errors are reported in parentheses; in the OLS regressions they are cluster-adjusted by firm. Median state and national unemployment rates are calculated over the 1992-2009 period. The uninteracted national unemployment rate is collinear with the treated firm-control fixed effects and is thus omitted. Each of the business cycle indicators is demeaned over the sample, using the kernel weights.

Table 9: Interaction with County-Level CRA Loans Per Firm

	OLS		<i>BranchShareCounty*</i> <i>SumSharesIV_{ct}</i>	
	CRA Loans ≤\$1 million	CRA Loans to Businesses with ≤\$1 million Revenues	CRA Loans ≤\$1 million	CRA Loans to Businesses with ≤\$1 million Revenues
Loan Amount	3.187 (0.068)	3.163 (0.069)	5.568 (1.241)	5.460 (1.261)
CRA Count/LBD Firm Ratio	0.078 (0.081)	0.221 (0.161)	0.617 (0.249)	1.510 (0.551)
Loan Amount*CRA Count/LBD Firm Ratio	-1.047 (0.069)	-2.671 (0.178)	-3.689 (1.020)	-8.827 (2.620)

The number of observations is 4,717,300, and the number of SBA firms is 161,100, rounded for disclosure avoidance. The regressions are kernel-weighted. The dependent variable is average employment in $t+1$ through $t+3$ minus average employment in $t-3$ through $t-1$. Loan amount is in millions of 2010 dollars. Treatment year age, age squared, and treated firm-control fixed effects are also included in the regressions. The number of Community Reinvestment Act (CRA) loans in the respective category in the county in $t-1$ is divided by the county's number of firms with positive employment in the LBD in $t-1$. Each CRA/LBD variable is demeaned over the sample, using the kernel weights. Robust standard errors are reported in parentheses; in the OLS regressions they are cluster-adjusted by firm.

Table 10: Controls in Same Counties vs. Controls in County-Industries
Without SBA Loans in Year t

	All Sectors	Non-Tradable Sector
Controls in Same Counties		
Loan Amount	2.348 (0.140)	2.871 (0.318)
N	253,900	39,500
Number of SBA Firms	54,000	9,500
Controls in County-Industries Without SBA Loans in Year t		
Loan Amount	2.360 (0.073)	2.604 (0.194)
N	2,327,900	594,500
Number of SBA Firms	54,000	9,500

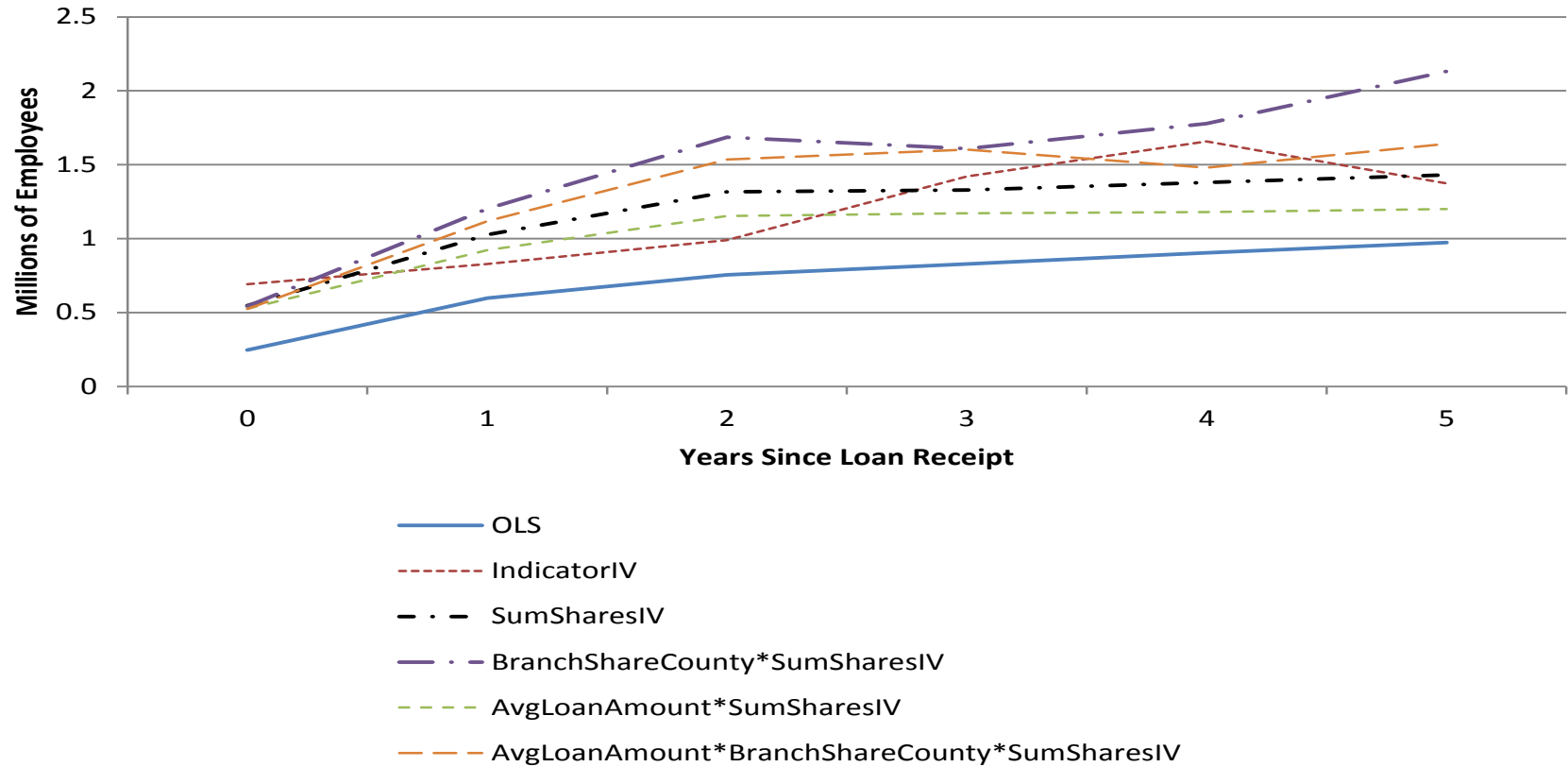
These are kernel-weighted OLS regressions. The dependent variable is average employment in $t+1$ through $t+3$ minus average employment in $t-3$ through $t-1$. Loan amount is in millions of 2010 dollars. Treatment year age, age squared, and treated firm-control fixed effects are also included in the regressions. Standard errors, reported in parentheses, are cluster-adjusted by firm. The number of observations and SBA firms are rounded to the nearest 100 for disclosure avoidance. The same set of SBA firms is included in both the regressions with controls in the same counties and with controls in county-industries with no SBA loans in year t .

Table 11: Employment Growth With Exit Zeroes

	OLS	<i>BranchShare County* SumSharesIV_{ct}</i>
Loan Amount	2.917 (0.066)	3.992 (1.107)
1 st Stage <i>BranchShareCounty* SumSharesIV_{ct}</i>		0.223 (0.007)
1 st Stage F Statistic		1004.19

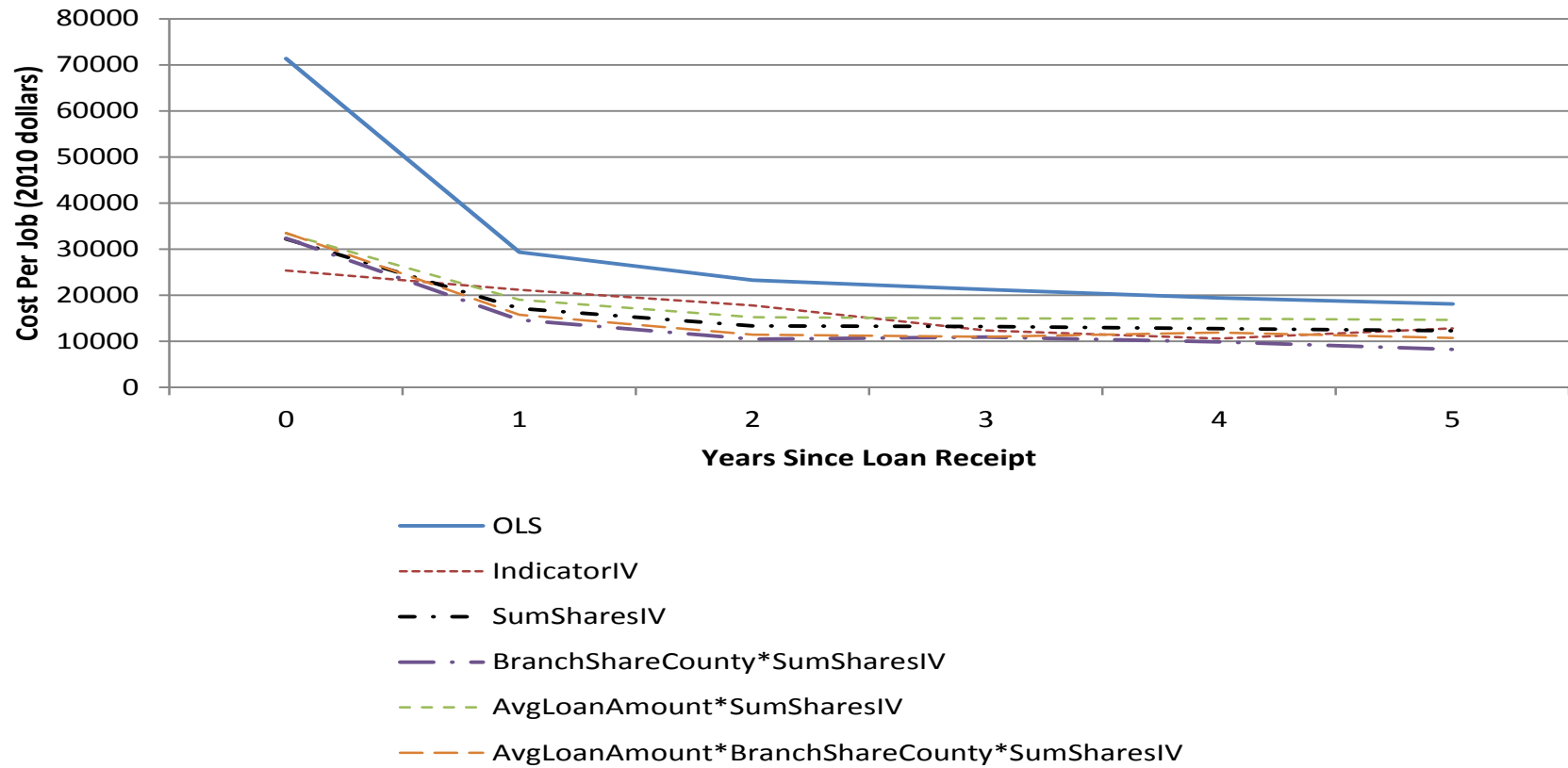
The number of observations is 8,085,100, and the number of SBA firms is 225,100, rounded to the nearest 100 for disclosure avoidance. The regressions are kernel-weighted. The dependent variable is average employment in $t+1$ through $t+3$ minus average employment in $t-3$ through $t-1$. For employment in years $t+1$ through $t+3$, if a firm has exited from the LBD prior to that year, zero employment is imputed for the year. Loan amount is in millions of 2010 dollars. Treatment year age, age squared, and treated firm-control fixed effects are also included in the regressions. Robust standard errors are reported in parentheses; in the OLS regressions they are cluster-adjusted by firm. For the dependent variable, zero employment is imputed for each year after the firm exits, up to three years after year t .

Figure 6: Total Number of Jobs Created in 1992-2007



The total amount of loans disbursed in the 50 states plus the District of Columbia in 2010 dollars in fiscal years 1992-2007 is \$230,358,424,195. This number is multiplied by the coefficients from the OLS regressions with treated firm-control fixed effects in Figure 4 and the IV regressions in Figure 5 for the years t through $t+5$.

Figure 7: Cost Per Job Created in 1992-2007



These numbers are calculated by dividing the estimated total cost to the SBA of loans in 1992-2007, \$17.553 billion, by the numbers of jobs created in Figure 6.

APPENDIX

Table A1: Path from Full SBA Loan Dataset to Treated Firms in Final Matched Regression Sample

	Number
Total SBA Loans in 1992-2009	1,120,600
Except U.S. Territories	1,104,600
Except cancelled loans	962,100
After consolidating loans to the same borrower in the same year	930,200
Except loans not matched to any business register	808,700
Except loans matched to non-employer business register	744,800
Except loans matched to a business register but not matched to the LBD firm data	662,700
Except SBA 7A/504 loans in years after the first loan year in the 1992-2009 period	493,800
Except firms with first SBA loan before 1992 or a SBA disaster loan at any time	484,000
Except firms with missing exact matching variables (employment in year before loan receipt, industry, or state (sample for control matching process))	355,700
Except firms without matched controls for three-year employment growth with exit zeroes	303,200
Except firms receiving SBA loan at start-up (sample for three-year employment growth with exit zeroes analysis)	225,100
Except firms missing three-year employment growth without exit	168,000
Except firms without matched controls for three-year employment growth without exit (sample for three-year employment growth without exit analysis)	161,600

Numbers are rounded to the nearest one hundred for disclosure avoidance.

Table A2: Sample Comparisons Among SBA Loan Recipients

	Not in Regression Samples			Regression Sample with Exit Zeroes	Regression Sample Without Exit
	Not Linked to Any Business Register	Linked to Business Register, Lacks Variables in LBD for Matching to Control Firm	Has Matching to Control Firm Variables, No Control Firms Found		
Total Number	121,400	453,000	52,500	225,100	161,600
Percent Matched by EIN/SSN	0.00	64.05	76.58	75.64	78.61
Mean Employment	4.96	10.43	12.95	14.77	15.73
Mean Loan Amount	177,741	278,819	332,302	367,987	399,065
Median Loan Amount	52,000	110,000	150,000	155,000	185,000
Percent Sole Proprietorships	57.14	23.70	16.92	19.59	16.88
Percent Partnerships	4.92	4.56	3.22	3.42	3.16
Percent Minority	33.79	26.90	21.86	24.55	23.20
Percent Majority	24.97	21.28	18.36	17.91	17.22
Female					
Percent Veteran	9.62	10.05	10.78	10.15	10.21
Percent By Sector:					
Construction	6.80	6.80	9.22	10.34	10.55
Manufacturing	6.23	11.28	15.09	10.01	10.54
Wholesale Trade	5.14	6.53	8.85	7.37	7.67
Retail Trade	23.36	22.61	19.67	19.94	18.40
Finance/Insurance/Real Estate	3.76	2.84	3.21	3.55	3.52
Services	44.67	43.64	37.56	44.49	45.10
Other/Unknown	10.04	6.29	6.39	4.30	4.21

The variables come from the SBA loan recipient database, incorporating data from the loan applications. Loan amounts are in constant 2010 dollars. Median loan amount is rounded to the nearest \$1,000 for disclosure avoidance.

Table A3: Path from Full SBA Bank Dataset to Sample Linked to LBD

	All SBA Banks		Banks Ever With PLP Status		
	Number of Banks	Number of SBA Loans	Number of Banks	Number of SBA Loans	Number of PLP Loans
Number in SBA data in Oct. 1990-Sept. 2008	6,075	938,845	942	802,620	192,352
Matched to Business Register	4,645	874,948	830	764,704	175,788
Matched to LBD	4,526	863,240	800	755,268	173,311

PLP includes both the preferred lender program (PLP) for 7(a) loans and premier certified lender program (PCLP) for 504 loans.

Table A4: Bias Before and After Propensity Score Matching

	All Non-Treated	All Treated	Control Sample	Treated Sample	Final % Bias	% Bias Reduction
Log Emp t-1	1.761	1.902	2.004	2.003	-0.143	98.69
Log Emp t-1 – t-2	0.017	0.071	0.072	0.075	0.647	94.83
Log Emp t-2 – t-3	0.024	0.072	0.069	0.072	0.705	93.83
Log Emp t-3 – t-4	0.027	0.070	0.066	0.069	0.653	93.81
Log Rev t-1	6.396	6.611	6.723	6.704	-1.268	91.26
Log Rev t-1 – t-2	0.031	0.091	0.093	0.111	3.901	70.06
Log Rev t-2 – t-3	0.041	0.095	0.099	0.109	2.319	81.38
Log Rev t-3 – t-4	0.052	0.098	0.094	0.103	2.044	82.31
Log Assets t-1	4.867	5.107	5.229	5.110	-6.519	50.33
Log Assets t-1 – t-2	0.017	0.116	0.095	0.121	4.049	74.15
Log Assets t-2 – t-3	0.030	0.073	0.077	0.088	1.787	75.09
Log Assets t-3 – t-4	0.038	0.076	0.074	0.084	1.776	72.75
Log Wage	3.042	3.048	3.088	3.094	0.839	-11.31
Age	10.706	8.324	8.926	8.836	-1.075	96.22
Multi-Unit	0.044	0.051	0.052	0.051	-0.281	91.82

The % bias is the standardized difference, which for a given variable, say age, is $SDIFF(age) = \frac{100 \frac{1}{N} \sum_{i \in A} [age_i - \sum_{j \in C} g(p_i, p_j) age_j]}{\sqrt{\frac{Var_{i \in A}(age) + Var_{j \in C}(age)}{2}}}$. The group before propensity score matching is treated and control

firms satisfying exact matches on employment, treatment year age category, industry, year, and state (if it has 19 or fewer employees in the prior year). Firms in the group after propensity score matching satisfy the propensity score bandwidth criterion, the common support criterion, and are in the regression sample.