# Microfinance beyond self-employment: Evidence for firms in Bulgaria

Erhardt, Eva

University of Mainz

March 2017

# MICROFINANCE BEYOND SELF-EMPLOYMENT: EVIDENCE FOR FIRMS IN BULGARIA

Eva Erhardt

Johannes Gutenberg University Mainz, Germany

Accepted manuscript as of March 2017

See http://dx.doi.org/10.1016/j.labeco.2017.04.009 for published journal article

**Abstract**

This paper provides new evidence on the impact of microfinance on job creation beyond self-employment. We examine wage-employment effects for a typical program in Eastern Europe with average loan sizes that are considerably above what has been studied so far. We apply propensity score matching extended by a difference-in-differences estimator to panel data from an individual-lending program to firms in Bulgaria. Our results indicate that microcredit has very positive effects on job creation. Participating firms have on average 2.5 (or 33 percent) more employees two years after receiving a microcredit than matched non-participants. This strong effect seems to be related to a certain loan size threshold necessary for positive impacts to unfold. Effects are largest for the smallest firms, supporting findings from other studies that small firms are more constrained by credit than large firms. Investigating dynamic effects for up to six years after treatment, we furthermore show that effects are long lasting.

Keywords: microfinance, wage employment, small firms, impact evaluation, Bulgaria

JEL Classification Number: G21, J23, D21, C21, P34

Contact details: Eva Erhardt, Chair of Applied Statistics and Econometrics, Johannes Gutenberg University Mainz, Jakob-Welder-Weg 4, 55128 Mainz, Germany;
Email: erhardt-eva@t-online.de

## 1 INTRODUCTION

Microfinance is deemed to be a strategy for creating jobs. The European Union (EU, 2010), for example, has embraced microfinance as a measure to increase employment levels in Europe until 2020. The rationale on which the promise to create jobs rests is typically as follows: Poor entrepreneurial individuals could earn high marginal returns through business activity but are credit constrained. Access to microcredit, defined as small loans to underserved entrepreneurs and their micro-enterprises, would then help realize growth opportunities by starting or expanding businesses, thus spurring employment (see Banerjee and Duflo, 2014; De Mel et al., 2008; Fafchamps and Schündeln, 2013; Karlan and Morduch, 2010; McKenzie and Woodruff, 2008).

Given that job creation is a major concern for policymakers worldwide, it is one of the reasons why microfinance has become increasingly popular since its emergence in the mid-1970s. Microfinance programs that provide microcredit next to other financial services are now widespread in low- and middle-income countries and have recently emerged in high-income countries as well. In total, about 200 million people worldwide are considered to be clients of some 3,600 microfinance institutions (Maes and Reed, 2012) and vast amounts of public and private funds are committed to microfinance programs; for example, USD 34 billion was committed in 2015 (CGAP, 2016).

The empirical evidence on the impact of microcredit on employment remains nevertheless astonishingly scarce. Among the large number of microfinance evaluations, most of the studied programs have other objectives than job creation, such as income stabilization, consumption or empowerment of women. Moreover, studies that do assess programs aiming at employment creation are usually limited to the impact of microcredit on self-employment (e.g. Angelucci et al., 2015; Attanasio et al., 2015; Augsburg et al., 2015; Chemin, 2008; Coleman, 1999 and 2006; Crépon et al., 2015; Duvendack and Palmer-Jones, 2012; Montgomery, 2005; Pitt and Khandker, 1998; Roodman and Morduch, 2014; Setboonsarng and Parpiev, 2008). While enabling self-employment is an important outcome, we believe that the acclaimed substantial contribution to job creation and economic growth is only achieved if microcredit succeeds in increasing wage-employment.

As far as the impact on wage-employment in the literature is concerned, results are still very mixed. Interestingly, among the few existing studies those focusing on the smallest loan amounts (around USD 200 in market exchange rates or USD 750 in purchasing power parity adjusted exchange rates)[1] find negative effects (Banerjee, Duflo, et al., 2015; Karlan and Zinman, 2011). Those studies on the other hand which evaluate the increase in wage-employment for programs with loan sizes about twice as large (around USD 500 in market exchange rates and USD 1,500 PPP-adjusted) find moderately positive effects (Dunn and Arbuckle, 2001; Tedeschi and Karlan, 2010; Gubert and Roubaud, 2011). The evidentiary base is still far too small for being conclusive. Nevertheless, these first results point to the fact that a certain loan size threshold might be necessary for wage-employment effects to unfold. This is also suggested by Banerjee (2013) who argues that one of the main reasons why most impact studies find a limited impact of microcredit could be that loans are too small for productive purposes. As even the largest loans in experiments amount to less than USD 1,000 (market exchange rates), the effects from larger microloans remain largely unknown.

Against this background, it seems very worthwhile to investigate wage-employment effects of a microcredit program with larger loan sizes than what has been studied so far. This not only provides an opportunity for testing if there indeed exits a necessary loan size threshold, but also if larger loans lead to larger positive effects. For this purpose, the Eastern European region provides an interesting case study with average loan sizes of approximately USD 2,500 compared to a global average of USD 700 at market exchange rates (MIX-Market, 2013). This difference only partially reflects the more advanced economic development of Eastern Europe compared to other regions. Larger loan sizes can be additionally explained by the widespread provision of microfinance in Eastern Europe by financially self-sustaining banks which tend to lend in greater volumes than not-for profit institutions (Cull et al., 2009; Hartarska et al., 2006). The provision of larger loans is furthermore enabled by the fact that microfinance institutions in the region typically lend to individual borrowers with already existing micro-businesses rather than applying the traditional group-lending approach for poor individuals or households who do not necessarily have previous entrepreneurial experience and a track record of successful lending and repayment (Armendáriz de Aghion and Morduch, 2000).

It is all the more surprising that the assessment of microfinance programs in Eastern Europe has been largely neglected so far. With regard to employment, the only existing study by Augsburg et al. (2015) investigates self-employment effects in Bosnia and Herzegovina. The lack of evidence for Eastern Europe is moreover in sharp contrast to the fact that in regional terms, the former Soviet Republics of Eastern Europe and Central Asia continue to receive the largest share of worldwide commitments to microfinance (CGAP, 2016).

[1] To compare loan sizes across countries we use purchasing power parity (PPP) adjusted exchange rates which are calculated by applying a conversion factor to market exchange rates. The PPP adjusted rates indicate how many USDs are needed to buy a USD's worth of goods in a domestic market as compared to the United States. If not stated otherwise we use rates for 2006 according to http://data.worldbank.org/indicator/PA.NUS.PPPC.RF (Accessed: 13 March 2017). Any conversions from EUR to USD are based on exchange rates for 31 December 2006 as indicated on https://www.oanda.com/lang/de/currency/converter/ (Accessed: 13 March 2017).

This paper therefore helps close two research gaps simultaneously. It first adds to the scarce literature on impact assessments of microfinance beyond self-employment by assessing wage-employment effects from microloans which are larger in size than what has been studied so far. By drawing on empirical evidence from a program in Bulgaria with typical features for the region, the study also contributes to the very limited evidence on microfinance in Eastern Europe. In fact, this study is to the best of our knowledge the first assessment of wage-employment effects from microcredit in an economically more developed context with relatively large loan sizes.

The unique combination of program data with a country-wide firm-level database enables us to apply propensity score matching methods extended by a difference-in-differences estimator. Our main sample consists of a balanced panel dataset from 2003-2010 of 974 participating firms that received a program loan in 2004 and 60,032 non-participating firms used for matching. The context of the program was chosen in a manner that ensures a largely limited supply of credit to micro and small firms other than from the assessed microfinance program. We test the robustness of our results regarding variations within the matching process. We also assess the sensitivity of results to a potentially remaining selection on unobservables into the program. Our results prove to be robust, and we find strong positive and significant effects of microcredit with respect to the number of employees in participating firms. In fact, participants have on average 2.5 (or 33 percent) more employees two years after receiving a microcredit than matched non-participants. This positive effect also holds if we include further variables in our propensity score estimations and longer time trends before treatment for subsamples. Heterogeneous effects for different firm size classes are largest for the smallest firms, which supports findings from other studies that small firms are more constrained by credit than large firms. With respect to loan size, we test whether larger loans lead to larger effects. As we find a very similar impact of 25-30 percent on wage-employment irrespective of the loan amount, this rather points to a certain loan size threshold for positive effects to unfold and that the loans assessed in this study are above that threshold. Observing dynamic effects for up to six years after receiving a microcredit, we furthermore show that effects are long-lasting. Our results also suggest considering an evaluation period of minimum two or more years after treatment for at least two reasons: First, an immediate expansion of employees directly after treatment is followed by a subsequent contraction in firm size before a sustainable new level of firm size is attained two years after treatment. Second, effects for different loan purposes, i.e., working capital or fixed assets, also occur at different points in time and seem to warrant longer evaluation periods.

The remainder of this paper is organized as follows: Section 2 provides a literature review on the impact of microfinance on job creation, an overview of access to finance for small firms in Bulgaria and a brief description of the program. Section 3 describes the data used in this study and presents descriptive results. Section 4 illustrates the identification strategy. The main results of our estimations are then discussed in Section 5, which also contains an analysis of dynamic and heterogeneous effects. Finally, we present sensitivity tests of our results in Section 6 before concluding in Section 7.

# 2 MICROFINANCE, SMALL FIRMS AND JOB CREATION

## 2.1 PREVIOUS IMPACT ASSESSMENTS ON MICROFINANCE AND EMPLOYMENT

Impact assessments of microfinance are numerous by now (for an overview see for example the meta-analyses by Duvendack et al., 2011; Pande et al., 2012 or van Rooyen et al., 2012). With regard to results, Banerjee, Karlan, et al. (2015) conclude in the introductory chapter to six randomized evaluations that there is “a consistent pattern of modestly positive, but not transformative effects of microfinance.” This finding now seems to have emerged after very high initial expectations on the part of proponents of microcredit (including for example the award of the Nobel Prize for Peace in 2006 to Mohammad Yunus and the Grameen Bank as pioneers of microfinance or the declaration of 2005 as the Year of Microcredit by the United Nations) were followed by widespread skepticism on the part of researchers about its positive impact (see, for example, Armendáriz de Aghion and Morduch, 2010).

As far as effects on employment are concerned, the evidence is still more limited for two reasons. First, the vast majority of evaluated programs are not designed to create employment. Rather, they aim at increasing income and consumption, poverty reduction or various other outcomes such as school attendance of children or empowerment of women. This is particularly the case for so-called group lending approaches from economically less developed regions which dominate the body of evidence and require borrowers, often women with no previous business experience, to form small groups for joint lending and repaying. Second, those studies which do assess programs aiming at employment creation are usually limited to the impact on self-employment. Only a handful of evaluations investigate the impact of microfinance on wage-employment. Table 1 summarizes the existing literature on employment effects of microfinance by distinguishing between effects on self- and wage-employment. Self-employment in this context includes the creation of a micro-enterprise. Moving from non-employer to employer by growing an enterprise in terms of employees on the other hand is regarded as wage-employment effect.

*(Table 1 about here)*

While some studies find positive effects on self-employment (Abou-Ali et al., 2010; Banerjee, Duflo, et al., 2015; Dunn and Arbuckle, 2001; Tedeschi and Karlan, 2010), others show negative ones (Gubert and Roubaud, 2011; Karlan and Zinman, 2011), or mixed results depending on the gender of borrowers (Chemin, 2008; Montgomery, 2005; Pitt and Khandker, 1998; Roodman and Morduch, 2014), individual versus group lending (Attanasio et al., 2015) and informal versus formal sector (Bruhn and Love, 2014) and yet a last group of studies does not detect any significant impact at all (Angelucci et al., 2015; Coleman, 1999 and 2006; Crépon et al., 2015; Duvendack and Palmer-Jones, 2012; Setboonsarng and Parpiev, 2008). The heterogeneity of estimated effects for self-employment is also reflected in a synthesis study on the impact of improved access to finance (including microfinance) on employment by Grimm and Paffhausen (2015). They conclude that only 20 out of 54 impact estimates from studies under consideration showed a positive effect on employment, in 2 cases a negative effect was found and 32 estimates were statistically not significant. Overall, the studied programs were more effective in creating self-employment than wage-employment. While enabling self-employment is an important outcome, we believe that the acclaimed substantial contribution to job creation is only achieved if microcredit succeeds in increasing wage employment.

*(Table 2 about here)*

Similar to self-employment, the impact on wage-employment is mixed and rather inconclusive. Table 2 provides more details on the few existing studies assessing wage-employment effects including average loan sizes.[2] The evidentiary base is still too small for being conclusive, but it is very interesting to see that those studies which focus on the smallest loan amounts (around USD 200 in market exchange rates and USD 750 PPP-adjusted exchange rates) find negative effects (Banerjee, Duflo, et al., 2015; Karlan and Zinman, 2011). Those studies on the other hand which evaluate the increase in wage-employment for programs with loan sizes about twice as large (around USD 500 in market exchange rates and USD 1,500 PPP-adjusted) find moderately positive effects (Dunn and Arbuckle, 2001; Tedeschi and Karlan, 2010; Gubert and Roubaud, 2011).

More in detail, the average loan size for borrowers businesses in the Indian program evaluated by Banerjee, Duflo, et al. (2015) was USD 200 (or USD 777 PPP-adjusted). The study found a negative, but insignificant impact on wage-employment for borrowers with existing businesses before treatment

[2] While Abou-Ali et al. (2010) and Bruhn and Love (2014) also assessed wage-employment effects, their outcome measure is limited to aggregate employment rates for the population in program areas and municipalities. Information on wage-employment effects in borrowers businesses is not available.

(0.05 fewer employees than for controls or about -20% compared to an initial firm size of 0.3 employees). For borrowers with new businesses the impact was significantly negative (0.2 fewer employees than for controls or again about -20% assuming an initial firm size of 0 employees for newly created businesses). Because new businesses in treatment areas also had lower profits than in control areas, the authors argued that microcredit might have a negative selection effect. It might draw individuals into entrepreneurship who actually have less propensity to become successful entrepreneurs than existing entrepreneurs. Karlan and Zinman (2011) investigate effects for a program in the Philippines with a very similar average loan size of USD 220 (or USD 726 PPP-adjusted). Again, they found negative effects. The number of paid employees in businesses of the treatment group decreased by 0.27 (-39%) relative to controls.

In contrast, Dunn and Arbuckle (2001) who evaluated loans twice as large (USD 586 or USD 1,368 PPP-adjusted) for a program in Peru found moderatly positive effects of microcredit on wage-emplyoment. The total number of days worked per month for non-houshold members in the borrowers' top three micro-businesses increased by 2.49 days per month or 0.13 full-time employees (4%) assuming 20 days of work per month. Tedeschi and Karlan (2010) re-estimated the impact for the same dataset by additionally accounting for program attrition between the baseline and follow-up survey. Their results show that the increase in wage-employment was over-estimated by Dunn and Arbuckle (2001), but is still positive with an increase of 1.18 days per month, 0.06 employees or 3%. Finally, Gubert and Roubaud (2011) study the largest loan sizes so far in terms of PPP with an average amount of USD 443 (1,616 PPP-adjusted). The effects on wage-employment in borrowers businesses were clearly positive (+9% employees), but insignificant. Results for other measures of firm growth such as sales and profits were however significantly positive.

## 2.2 Credit Constraints for small Firms in Bulgaria

The impact of microcredit on wage-employment is supported by economic theory arguing that entrepreneurial individuals are credit constrained; otherwise, they would engage in profitable business activities. Access to microcredit is supposed to help overcome these credit constraints, making it possible for individuals to realize growth opportunities by starting or expanding businesses, which creates employment. To assess impacts, it is therefore necessary to identify an adequate group of credit-constrained firms that do not participate in the program and compare them to participants. The purpose of this section is to discuss whether such a control group can be found in our research context.

One might argue that, ideally, neither participants nor non-participants should have previous access to finance before the start of the program and that, moreover, non-participants should remain without access during the evaluation period. In reality, it is very unlikely to find firms that are totally excluded from the financial sector, and to some extent, it is not even a necessary condition. Most businesses have at least access to informal sources of finance such as moneylenders or family and friends. Research shows, however, that informal financial institutions co-exist with microfinance and play a complementary role in the formal financial sector, offering small, unsecured, short-term loans (Ayyagari et al., 2010; Guirkinger, 2008; Madestam, 2014). Microfinance's promise to create wage-employment by enabling investments in profitable business opportunities therefore seems independent of the existence of an informal financial sector. As far as access to loans from the formal financial sector is concerned, we nonetheless need to take great care in discussing its availability to the control group.

The data for our empirical analysis contain information on take-up[3] of formal loans from the program lender ProCredit Bank Bulgaria before and during the evaluation. Accordingly, we are able to exclude any firms from the control group of non-participants that received loans from ProCredit Bank. Since ProCredit Bank was the largest provider of microfinance in Bulgaria, we already control for the most

[3] Information on access would be the better variable, but is not available. In section 5.4.4, we however test the matching of participants with non-participants from districts without program presence.

important source of formal finance for small firms. Furthermore, we chose a context in terms of country and time period where access to formal finance was still very limited in general. However, since the data do not, unfortunately, provide information on access to loans from other banks or microfinance institutions, we next provide an overview on the general availability of bank credit during the period under review.

It is worth noting that the availability of alternative credit to non-participants is of course an issue that also other impact assessments of microcredit face. In the case of Banerjee, Duflo, et al. (2015), for example, almost no microcredit was available at baseline (although about two-thirds of sample households had at least one loan outstanding from informal financial sources), but during treatment other microcredit programs simultaneously started operations. The authors suggested that one can still interpret differences between treated and controls as being a result of the microcredit program as long as borrowing rates are at least higher among treated than controls. We have good cause to assume that this is the case for our group of participants compared to non-participants. Furthermore, Wydick (2016) demonstrates that experimental studies focusing on marginal borrowers are likely to understate the impacts yet realised by inframarginal borrowers – those having taken microfinance loans prior to implementation of an experiment. Provided that loans have a positive impact on employment, we can furthermore argue that if a certain share of control firms benefits from alternative loans, we would rather underestimate the positive effect of the program loans on participant employment levels.

### 2.2.1 THE DEMAND FOR CREDIT

To answer how credit-constrained our non-participants were, we first address the demand side of finance. The target group of the program under review is micro-enterprises with less than 10 employees. More precisely, a vast majority of the 78 percent of microcredit clients in our sample were micro-enterprises at the start of the program. Further, 20 percent were small enterprises (with less than 50 employees), and 2 percent were medium-sized enterprises (with up to 250 employees) defined according to EU (2005). As many statistical sources treat micro-, small- and medium-enterprises (MSMEs) as one group of firms in contrast to large enterprises, we use the term ‘small firms’ when referring to the group of firms with up to 250 employees.

By 2003 – the starting point of the period under review - existed around 240,000 micro-, small- and medium-sized enterprises in Bulgaria.[4] Small firms were considered the engine of the Bulgarian economy: They constituted 99.7 percent of all enterprises (90 percent were micro-enterprises), generated 79 percent of employment, accounted for 75 percent of turnover and 61 percent of the added value of private enterprises (Ministry of Economy, 2004). Bulgaria is not an exceptional case in that regard. de Kok et al. (2011) showed that small firms also accounted for 99.8 percent of non-financial enterprises in the European Union, with the overwhelming majority of 92 percent being micro-enterprises. Micro-, small- and medium-sized enterprises employed two-thirds of the formal EU workforce and 85 percent of total net employment growth was created by small firms from 2002-2010. The general economic context for small firms in Bulgaria had improved by 2003 after more than a decade of transition to a market economy. Bulgaria had a population of 7.8 million, GDP grew at 5.4 percent, unemployment amounted to 13.7 percent and it was classified as a lower middle income country. In 2007, Bulgaria became a member of the European Union. Nevertheless, its GDP per capita remains among the lowest in Europe.[5]

Despite the vital importance of small firms for the Bulgarian economy, they had difficulties accessing formal finance. A survey by the Bulgarian SME Promotion Agency in 2002 showed, for example, that

[4] https://infostat.nsi.bg/infostat/pages/reports/result.jsf?x_2=1376 (Accessed: 24 February 2017).

[5] Economic indicators are based on data from Worldbank, available at: http://data.worldbank.org/country/bulgaria (Accessed: 16 November 2015). GDP per capita was USD 2,697 in 2003 and USD 7,499 in 2013.

67 percent of small- and medium-sized firms had no access to bank financing (Simeonova-Ganeva et al., 2011). The Worldbank (2006) confirmed that in 2002 only 16 percent of interviewed Bulgarian firms used formal borrowing from the financial sector for new investments (21 percent in 2005). Not surprisingly, access to finance was reported as a major constraint to firm growth. A survey of managers from small- and medium-sized enterprises in 2003 pointed in the same direction: The shortage of financial resources for investment and working capital was considered a major barrier to growth (BSMEPA, 2004). In a study by Pissarides et al. (2003) on principal constraints of small and medium enterprises in Bulgaria and Russia, top managers also identified the lack of external financing as a particularly serious constraint. In addition, Budina et al. (2000) showed that investment decisions of medium and large Bulgarian firms in the period 1993-1995 were constrained by liquidity and the smaller the firm, the larger were their constraints.

Evidence from other countries adds further support to the fact that firms are often credit constrained and smaller firms are more constrained than larger firms. Ayyagari et al. (2008) showed that of ten obstacles in the business environment reported by firms, lack of finance had the largest direct effect on firm growth. As for firm size, the authors found smaller firms to be more constrained by financing than larger firms. Beck et al. (2005) also found that the significance of growth constraints varies considerably with firm size. It was consistently the smallest firms that are most constrained. For small firms, the financing obstacle had almost twice the negative effect on annual growth than it does for large firms. In a related study, Beck et al. (2006) focused specifically on the determinants of financing obstacles and found that younger, smaller and domestic firms report higher obstacles.

As for Eastern Europe specifically, Hartarska and Nadolnyak (2008) compared in Bosnia and Herzegovina the investment sensitivity to internal funds of micro-firms in municipalities with a significant presence of microfinance institutions to those in municipalities without any (or limited) presence. Their results indicated that microfinance alleviated the financing constraints of micro-enterprises. In a similar vein, the findings of J. D. Brown et al. (2005) also suggest that the availability of loans is an important factor for the growth of start-up firms in Romania.

### 2.2.2 The Supply of Credit

With regard to the supply of formal financing in 2003, the Bulgarian banking system had recovered from several crises since the transition to a market economy. The banking sector consisted of 29 banks and 6 branches of foreign banks. Foreign ownership dominated and all but 2 banks were private. While the banking sector was still relatively small with a ratio of domestic credit to the private sector (as percentage of GDP) of 26 percent in 2003 compared to Western countries such as Germany with 110 percent or the U.S. with 177 percent,[6] bank credit experienced rapid real growth of 40 percent as in most other countries in Eastern Europe. The risk-averse behavior of banks during the early transition period, gradually gave way to increased lending aided by economic recovery and privatization of state banks (Duenwald et al., 2007).

Notwithstanding the overall expansion of bank credit, financing for small firms remained limited as the following evidence suggests: Duval and Goodwin-Groen (2005) estimate that in 2001 banks had collectively extended only around 16,000 business loans under EUR 100,000 to the total of Bulgarian enterprises (about 240,000). Anecdotal evidence from the program points into the same direction. As described in ProCredit Bank Bulgaria's annual report 2004 'when the bank was founded in October 2001, traditional banks in Bulgaria were still extending loans to small and medium-sized enterprises only in exceptional cases. The conventional commercial banks often focus their lending operations on corporate finance and consumer lending, but tend to neglect small businesses as a potential clientele. Their main reasons for not lending to micro, small and medium-sized enterprises are the perceived

[6] http://data.worldbank.org/country/bulgaria (Accessed: 16 November 2015).

inadequacy of MSMEs' accounting methods, their inability to provide sufficient collateral and the high administrative costs incurred in small business lending.' In terms of market share of Bulgarian banks in the segment of small business lending, a survey by BSMEPA (2004) covering 8 out of 29 banks reveals: out of a total of 13,771 loans extended to small- and medium-sized firms by the surveyed banks in 2003, ProCredit Bank Bulgaria is the largest provider in terms of number of loans (5,560 loans) followed by United Bulgarian Bank (4,537 loans).

Next to banks, the microfinance sector constitutes another source of formal finance for small firms. According to data from MIX-Market[7] the gross loan portfolio of microfinance institutions in Bulgaria in 2004 was worth about USD 200 million and there existed about 37,000 active microfinance borrowers. The microfinance sector was largely dominated by ProCredit Bank Bulgaria, making up for more than 90 percent of the loan portfolio and 77 percent of active borrowers.

The evidence therefore points to the fact that small firms were largely constrained by access to formal finance and that it is possible to identify an adequate control group in this research context. Only about one third of micro, small and medium-sized firms in Bulgaria indicated any access to formal finance in 2002 before start of the program (Simeonova-Ganeva et al., 2011). Access to loans for micro- and small firms is assumed to be even much lower (e.g. Budina et al., 2000; Ayyagari et al., 2008; Beck et al., 2005). In addition, ProCredit Bank Bulgaria was by far the largest provider of microcredit in the microfinance sector and one of the largest providers within the banking sector. The rapid expansion of the bank's lending operations to small firms and its large market share can be regarded as a further indicator for the limited access to finance of small firms in Bulgaria during the period of evaluation. Moreover, small firms indicating access to bank loans were to a considerable extent clients of ProCredit Bank and can be consequently excluded from the control group. Finally, if some control firms received loans from other providers than ProCredit Bank and if loan effects are positive, we should expect to slightly under-estimate the impact of the program.

### 2.3 The Program - Pro Credit Bank Bulgaria

The program assessed in this paper consists of micro-loans by ProCredit Bank Bulgaria[8] and incorporates many of the characteristics that are typical for microfinance in Eastern Europe. The bank is a for-profit microfinance bank[9] and part of ProCredit Group, which consists of banks in Eastern Europe, Latin America and Germany. It was established in 2001. As expressed by the chairman of its supervisory board in the bank's first annual report, its establishment was motivated by the under-supply of credit to small businesses. The bank's mission statement further specifies the focus on micro, small- and medium-sized enterprises based on the conviction that these businesses create the largest number of jobs and make a vital contribution to the economy. Job creation is therefore a specific objective of the program. The bank's owners do not expect short-term profit maximization, but nevertheless a sustainable return on investment. The bank follows an approach of individual lending to entrepreneurs with a minimum experience of 6 months. Its lending operations expanded rapidly. The bank started in 2001 with 955 loans worth about EUR 5 million. In 2004, the beginning of the period under review, it already had a total of 19,390 business loans outstanding worth EUR 127 million. By 2010, the end of the review period, the bank had 31,675 business (and agricultural) loans outstanding worth EUR 533 million. Accordingly, the average loan size for business loans is relatively large even for Eastern Europe (5,236 in 2001, 6,550 in 2004 and 16,827 in 2010). As for geographical outreach, ProCredit Bank Bulgaria started operations with 7 branches. The network grew to 35 branches by the end of 2004, covering 19 out of 28 Bulgarian districts. In 2010, the country-wide network numbered 75 branches and outlets.

---

[7] http://www.mixmarket.org/mfi/country/Bulgaria (Accessed: 16 November 2015).
[8] For more information see http://www.procreditbank.bg/en (Accessed: 16 November 2015).
[9] In 2012, ProCredit Group introduced a strategic shift away from its original mission as microfinance bank to becoming a financial service provider for small and medium-sized businesses.

# 3 DATA

## 3.1 FIRM-LEVEL DATASET

We constructed a unique dataset by matching two different sources: (i) program data from the microfinance bank, and (ii) a firm-level database for Bulgaria. The program data provides us with information on loans (amount, date of disbursement, loan purpose, and maturity) and clients. Client data on financials and the number of employees is only available for about 75 percent of clients. It is furthermore usually recorded only once at loan approval so that a development over time cannot be tracked. We therefore use the program data to identify our sample of program participants and to obtain loan information. As for all other information on clients' characteristics, in particular their number of employees, we rely on the second source described below. Using one single source of information for characteristics of both participants and non-participants also has the advantage of reducing an important bias in impact studies as Heckman et al. (1997) point out.

We define treatment as disbursement[10] of a loan from ProCredit Bank between 1 January and 31 December 2004. The decision to focus on 2004 results from a trade-off between our research design and data availability. In a context of continuously expanding credit markets from early 2000s onwards, we on the one hand preferred an early period of bank operations where the availability of loans to non-participants was still very limited. Our intention was to be better able to attribute the impact to microcredit instead of under- or overestimating it as discussed above. On the other hand, we were not able to choose yet earlier years of operation (the bank started in October 2001), because data availability in the firm-level dataset sharply reduces when going more backwards in time. To further define treatment, we chose to focus on new clients only, i. e. firms that have not benefitted from earlier loans by ProCredit Bank as we consider the first loan to best kick-start a process of growth. Control firms in contrast did not receive a microcredit by ProCredit Bank in any year prior to or during our period under review.

With regard to the evaluation period, the firm level database provides us with year-end data (31 December) from 2000 until 2010. Our baseline hence dates from 31 December 2003 just before start of the treatment on 1 January 2004. We then allow the treatment effect to take several years to materialize. As the average duration of program loans was 24 months, we accordingly define the end of 2006 as our main evaluation period, i.e. 24 months after disbursement of the last program loan on 31 December 2004. In addition, we will also show the dynamic development of treatment effects for other periods.

To further construct our sample, we cleaned the dataset by loan purposes and kept only business loans. We deleted loans for consumption and housing, because we assume that non-productive loans for private purposes will not have an impact on labor demand in client firms. The core business loan products under the program in 2004 were the micro-loans 'Sprint' and 'Dynamo'. The Sprint loan was especially popular, because no collateral was required and applications took only 24 hours to process. Overall, the average size of a business loan in our sample was EUR 7,092 (equivalent to USD 9,363 at market exchange rates or USD 23,605 PPP-adjusted). This is considerably above the loan sizes of up to about USD 1,616 PPP-adjusted which have been studied so far by other assessment of wage-employment effects.

The bank disbursed loans to 1,153 new formal business clients in 2004. These firms constitute the focus of our research because information in the Bulgarian firm-level dataset is confined to formally registered businesses. The bank furthermore disbursed a considerable amount of 8,117 business loans in the same year to private individuals. These clients are either professionals such as farmers or doctors who are not required to register their business or they are owners of informal and formal firms

[10] One could argue that firms already start hiring employees at loan approval. As on average only 9 days elapsed from loan application to disbursement, it does not seem to make a large difference which stage we use.

who received loans for business purposes as private individuals.[11] Since the bank provided us with information on the formal businesses which are related to private individual clients and vice versa, we were able to exclude any participants and non-participants from our dataset who had already received a microloan as a private individual in previous years (or during the evaluation period in the case of non-participants). Firms which received loans as private individuals differ from those which are legal entities in terms of firm size, but both groups belong to the micro firm-size category (<10 employees).[12] Based on employee figures from the bank, formal businesses among clients in 2004 had on average 8.1 employees (7.5 including the owner according to Amadeus) while private individuals had 2.5 employees. Our sample consequently provides evidence on employment effects of microcredit for businesses with larger initial firm sizes compared to existing research.

The second data source is the firm-level database Amadeus (provided by the Bureau van Dijk). It contains financial and other information on companies in Europe. Information on the number of employees per firm in Bulgaria originates from the national social security institute. This means that only formal jobs for which firms pay social security are accounted for. The Amadeus database has the decisive advantage for the purpose of our research on microcredit to contain a very large number of micro-enterprises with less than 10 employees. The dataset at hand contains year-end firm-level information on registered Bulgarian enterprises from 2000-2010. The coverage of firms in terms of number and characteristics increases for more recent years. In 2004, the year of treatment, the database contains 106,894 firms. Out of these observations, information on our outcome of interest (difference in employees 2003 to 2006) and the necessary covariates for our empirical strategy of matching (location, sector and ownership type) is available for 65,496 firms.[13] Data on sales or profits is available to a much lesser extent, in particular for micro-sized firms and we decided to not consider its availability as requirement for forming our sample. In a next step, we deliberately eliminated any firms receiving a business loan from ProCredit Bank from 2001-2003 for participants and 2001-2003 as well as 2005-2006 for non-participants either as formally registered business or as private entity (3,833 firms). This leaves us with 61,663 firms which were merged with program data using the national firm identification number. We could identify a large majority (87 percent) of the new clients in 2004 in the Amadeus database. In a final step, we excluded large firms (591 firms) because they are too few to carry out precise estimations and were not the target group of microloans in the first place and controlled for outliers in terms of employees (66 firms).[14] Eventually, we obtain a considerably long, balanced panel dataset of 974 participating firms and 60,032 non-participating firms with annual information before, during and after treatment on number of employees, sector, location, ownership type and gender of owner.

Our sample represents 26% of formally registered Bulgarian firms in 2003. As shown below in Table 4 with regard to firm size before treatment, the set of participating firms is almost identical to the average Bulgarian firm with 7.2 employees. It is also representative of average firms in terms of sectoral and regional distribution.[15] The set of non-participating firms on the other hand is of larger size than

---

[11] The informal sector in Bulgaria amounted to about one third of GDP in 2007 according to Williams (2014). The disbursement of loans for formal businesses to private individuals was a common practice in the early years of bank operations when it took authorities very long to issue the necessary documents for loan applications.

[12] In terms of sectoral activity, 51% of private individuals are classified as "activities of private households as employers". The remaining firms are characterized by a higher share in agriculture (8% compared to 3% for legal entities) and a lower share in manufacturing (7% versus 17%), while other sectoral distributions are similar.

[13] The most substantial reduction in sample size (31,667) is caused by missing information on the number of employees in 2003. About half of these firms (13,797) are included in the database although they are non-active. Further 8,492 firms exited in 2003 and did not have any more employees at the end of 2003. The remainder of 9,378 firms did not report information for 2003, but for previous and subsequent years.

[14] We excluded 1 large firm from treated and 590 from controls. We deleted firms with a relative change of 5,000% or more in employees from 2003-2006 (35 firms from controls and none from treated) as well as firms with an absolute change of 300 or more employees (31 firms from controls and none from treated).

[15] https://infostat.nsi.bg/infostat/pages/reports/query.jsf?x_2=1376 (Accessed: 24 February 2017).

average firms because only medium-sized and large enterprises were obliged to file official accounts. As far as differences between participants and non-participants in our sample are concerned, they would however only matter to the extent as they are not controlled for by matching. In that regard, one could argue that we for example do not control for management practices although badly managed firms are probably less likely to report information to authorities (and thus enter the database) than well managed ones. Again, this would only bias our results if among matched non-participants poorly managed ones were less likely to provide the full information to official sources than among participants or vice versa. We do not find reason to believe that this would be the case.

The dataset has three important advantages: First, sample size for both participants and non-participants is very large. The number of non-participating firms relative to the number of participating firms amounts to 62:1. As we will apply propensity score matching, this increases the likelihood of finding one or more non-participating firms that are very similar, if not identical, to each of the participating firms. Moreover, compared to other evaluations of microfinance, we also have a large number of participating firms. Among other things, this enables us to analyze sub-groups of participants in terms of firm size, loan features and geographic location for heterogeneous program effects. Second, the panel structure of the dataset allows us to use a difference-in-differences approach which controls for time-constant unobservables that may have determined the participation in the program and the respective performance of firms. In section 4.1 we will further elaborate on this advantage. A third significant feature of the dataset is that it allows us to account for firm exits of both participants and non-participants. The database contains historical records of inactive companies and we interpret as exit if data on employees or any financials is missing from a specific year onwards until the final year of the database. We set employee values to zero from the year of exit onwards so as to make firm exits result in a negative change of employees. This way, we can avoid the common survivor bias in panel data related to attrition of sample firms. Certainly, some of the exits could rather result from the acquisition of a company by another one. We cannot distinguish between job destruction due to firms that dismiss employee or due to a transfer of employees to other legal entities. In a similar manner, job creation captures total firm growth, regardless of whether the increase in employment is the result of organic growth (internal) or acquired (external) growth. Overall however, exit rates in the Amadeus database correspond very much to what is reported by Eurostat (2009).

Table 3 shows firm exit rates for participants and non-participants separately. They are much higher among non-participants during the evaluation period 2004-2006. The cumulative exit rate for participants amounts to 2.9 percent, while the corresponding rate for non-participating firms is 13.7 percent. Interestingly, for later years from 2007 onwards, the exit rates increase in particular for participants and become very similar to non-participants. One explanation could be decreasing effects of microcredit over time on firm survival. Since average exit rates increase for all firms in our sample, this might also be triggered by the accession of Bulgaria to the European Union in 2007 and increased competitive pressure in particular on small domestic firms. Moreover, the effects of the global financial crisis on small firms also start to increasingly show from 2007/2008 onwards.[16]

*(Table 3 about here)*

The main limitation of the dataset is that it does not provide us with more information on outcome variables such as sales and profit or data on part-time employment and wages. The literature on small firm growth points to the fact that for very small firms, it usually takes a long time to increase one more formal employee or as Coad and Hölzl (2012) put it 'indivisibilities are substantial for very small firms.' They argue that growth effects of microcredit in terms of formal employees will thus not quickly show.

---

[16] Another explanation for the difference in exit rates could be that the program is successful in selecting firms that are not going bankrupt in the short-term. It would however presuppose that the bank is able to predict exit rates very accurately with regard to loan maturities and that the it furthermore disregards the likelihood of exit from 2007 onwards – even though the bank is probably interested in a several follow-up loans.

Here it would be good to have information on part-time employment or sales which usually mirrors growth faster. In addition, it would be interesting to see how profitability or wage levels are influenced by microcredit. Another constraint is that we need to work with limited data for most of the firms in our sample as concerns pre-treatment trends, because information on employees and financials for earlier years than 2003 is only available for a reduced number of observations. Longer time trends together with additional firm characteristics to match upon such as firm age would add further strength to robustness of results. Finally, we cannot totally exclude crowding-out effects of participants which prosper at the cost of non-participants. The comparison of growth rates for matched and unmatched non-participants nevertheless shows that both develop in the same manner. As we have reason to believe that potential crowding-out effects would be strongest towards the very similar matched non-participating firms, our data does not support the existence of such an effect.

## 3.2 Descriptive Statistics

Table 4 provides the descriptive statistics measured at entry into the microcredit program at the end of 2003 separately for participants and for non-participants.

*(Table 4 about here)*

Looking at participants first, 78 percent were micro-enterprises with a size of less than 10 employees.[17] 20 percent were small firms with 10-49 employees and 2 percent were medium-sized enterprises with 50-249 employees. On average, participating firms had 7.49. Compared to our sample of non-participants, there is a considerable size difference. In fact, non-participants were almost twice the size of participants with 12.74 employees. This size difference is also reflected in the distribution of firms within the three firm size categories. As discussed above, our sample represents 26% of total Bulgarian firms. While participants in our sample correspond very much to the average Bulgarian firm in terms of firm size (7.2 employees), non-participants tend to be larger.

In terms of sectoral breakdown, about half (49 percent) of the microcredit is lent to trade companies (NACE section 'wholesale and retail trade; repair of motor vehicles'), 17 percent into the manufacturing sector, and 10 percent into transportation. Real Estate activities and hotels/restaurants represented only 7 and 5 percent of the participants respectively. Other sectors such as agriculture or construction had a share of less than 5 percent each within the group of participating firms. Relative to non-participants, participants were more concentrated in trade, transportation and hotel/restaurant sectors. Non-participating firms on the other hand had a higher share within agriculture, construction and real estate businesses. For micro-enterprises, it is not surprising that a large share is in trade and other service sectors. As a consequence, the slightly higher share of participating firms in these sectors could simply be a firm size effect.

When considering location, we present clusters of economic development. The 28 districts of Bulgaria are ranked according to their level of economic performance based on a composite index of different indicators and then clustered into five groups. 13 percent of the participants were concentrated in areas with very good economic development, 31 percent in areas with good development, 25 percent with average, 19 percent with unsatisfactory and 12 percent with weak economic development. In comparison to non-participants, participating firms were geographically more equally distributed across the different clusters. Non-participating firms were more concentrated in areas with 'good' economic development. The more even spread of participating firms in terms of location can be interpreted as a sign for the microcredit program's objective to broadly increase access to finance in a development-oriented manner instead of exclusively selecting the economically most promising locations. In total, ProCredit Bank Bulgaria was present with a branch by the end of 2004 in 19 out of 28 Bulgar-

[17] Note that a self-employed firm owner is also counted as employee, so that the minimum number of employees is 1. Zero employees means that a firm went out of business as discussed above.

ian districts, while we have observations for non-participants from all of the districts. This fact is graphically illustrated in Figure 1 and will be later exploited in Section 5.4 for estimating heterogeneous program effects.

*(Figure 1 about here)*

As for the type of ownership, 44 percent of the participating firms were limited partnerships. 31 percent were public joint-stock companies and 24 percent were limited liability companies. A very small number of 1 percent were state enterprises and even less were general partnerships. Non-participants have a slightly higher share of general partnerships and state enterprises. Moreover, non-participants are less often limited partnerships, but instead more often limited liability companies than participating firms. While in a limited partnership at least one managing partner must bear liability for the partnership's actions, in a limited liability company all owners are protected from financial liability, regardless of whether they play an active role in the direction of the company. The higher share of limited partnerships among participants compared to non-participants could reflect preference of the microcredit program for borrowers which have to bear liability for their firm.

Gender of firm owners is finally very similar for both groups of participants and non-participants. If we disregard observations without information on gender of firm owners which is more common among non-participants, then for both participating and non-participating firms about 75 percent are owned by male and 25 percent by female entrepreneurs.

The fact that participating and non-participating firms apparently differ in their characteristics at the start of the microcredit program implies that we should control for these difference when comparing the two groups. At the same time, participating firms do not seem to be a totally different group of firms than are non-participating ones. To have a sufficient amount of non-participants that are inherently similar to participants in relevant ways, is favorable for finding good matches for participants, next to the fact of a merely large number of non-participants relative to participants.

In Table 4 we also report our main outcome of interest which is the change in number of employees for participants and non-participants from before treatment (2003) to after treatment (2006). As we investigate the growth of small firms, we focus on absolute rather than relative change at this stage. Indicators with very low values obviously have a greater tendency to produce large relative change, even when the absolute change is small. Even more important, different base values of indicators will lead to different relative changes even if the absolute change is identical. Applied to our sample with an average firm size before treatment of 7.492 for participants and 12.737 for (unmatched) non-participants, an absolute increase of 2 employees would translate into a relative increase of 27 percent for participants, but only a 16 percent for non-participants. As shown at the bottom of Table 4, employment grows over time for participating firms whereas it decreases for non-participants. In detail, participants have on average 2.73 more employees two years after treatment, while non-participants shrank on average by -0.18 employees over the same time period. We hence observe a considerable raw difference in the development of employees between participants and non-participants. However, these are descriptives only and we do not yet know if the gap is caused by the treatment with microcredit or by differences in key characteristics of the firms.

# 4 EMPIRICAL STRATEGY

## 4.1 IDENTIFICATION OF CAUSAL EFFECTS

For estimating the causal effect of microcredit on wage-employment we use a treatment effects methodology for observational data with non-random assignment of treatment. Our purpose is to account for possible selection biases in order to better identify the counterfactual of how employment in participating firms would have grown in the absence of their participation in the program. A typical bias for microcredit programs arises from self-selection, when firm-owners with better entrepreneurial skills select themselves into the program. But self-selection bias can also go the other way round, when a firm-owner borrows because he has experienced, or expects to experience, a negative shock. Another source of bias for microfinance stems from program selection. Again the bias can go in either direction. As Banerjee, Karlan, et al. (2015) suggest, lenders choose which markets to enter, and depending on their motivation may thus select relatively vibrant and growing markets (because of profitability) or stagnant and poor ones (because of social concerns). ProCredit Bank rather belonged to the first group of lenders with a preference for clients with growth perspectives. The very rapid expansion of its operations on the other hand (from about 1,000 to 20,000 loans from 2001 to 2004 and from 7 to 35 branches) can be taken as indication that this strategy could only be implemented to a limited extent. The even spread of branches across differently developed economic regions additionally points into the same direction. Moreover, as M. Brown et al. (2011) show that only 5 percent of firms which applied for a loan in Eastern Europe in 2004 were rejected, program selection does not seem too acute in our context. Nonetheless, as descriptive statistics showed, characteristics of participants are different from those of non-participants already before participating in the program and some of the change in employees might rather be attributable to these differences rather than the microcredit program.

To minimize the effect of selection biases on our results, we use propensity score matching (see Rubin, 1974; Rosenbaum and Rubin, 1983 and 1985 for conceptual foundations). The basic idea is to find non-participants which are similar to participants in relevant pre-treatment characteristics. That being done, we attribute differences in outcomes between participants and matched non-participants to the program. Matching has become a well-established approach for treatment evaluations. In the area of microcredit it was for example used by Chemin (2008) or Setboonsarng and Parpiev (2008) to reassess datasets which had already been analyzed with methods that accounted less for selection biases. The re-estimated impact with matching was typically found to be lower.

Methodological issues for propensity score matching are discussed in detail by Becker and Ichino (2002), Caliendo and Kopeinig (2008); Dehejia and Wahba (2002) or Smith and Todd (2005). We first of all need to rely on the conditional independence assumption. It implies that systematic differences in outcomes between participants and non-participants with the same values for covariates are attributable to the program, that is $Y^0, Y^1 \perp D|X$. $Y^0$ indicates that a firm does not participate in the microcredit program, $Y^1$ means that a firm participates in the program, $\perp$ denotes independence, D is the indicator for program participation and X a set of observable covariates which are not affected by program participation. In other words, there is no more bias from omitted variables once X is included. This clearly is a strong assumption. It will be justified by additionally applying a difference-in-differences approach to further control for differences in time-invariant characteristics. Moreover, we test the sensitivity of our results to unobserved covariates. A second assumption is that participation probabilities lie in the same domain. This is called the common support assumption and can be expressed as $0 < P(D = 1|X) < 1$. It implies that for each participant there is a non-participant with a similar X. We impose the common support condition in all of our analyses.

Conditioning on all relevant covariates is limited in the case of a high dimensional vector X. When the number of covariates increases, the chances of finding a match reduces. Rosenbaum and Rubin (1983) showed that instead participants may be matched with non-participants using a summary measure of similarity in the form of a propensity score rather than the entire set of covariates. Let $P(X_i)$ be the probability that unit i is assigned to the participating group, conditional on characteristics $X_i$, and X is the multidimensional vector of pre-treatment characteristics or time-invariant stable firm characteristics, and define $P(X_i) \equiv \Pr(D_i = 1 \mid X_i) = E(D_i \mid X_i)$

Given that the conditional independence and common support assumptions hold, the PSM estimator for our parameter of interest ATT (average treatment effect on the treated) can then be written as

$$\Delta_{PSM}^{ATT} = E(Y^1 | P(X), D = 1) - E(Y^0 | P(X), D = 0) \quad (1)$$

where the first term can be estimated from the group of participants and the second term from the matched non-participants.

As we have panel data and both pre- and post-program information on the outcome variable, we can extend this standard approach by a difference-in-differences (DiD) matching estimator as suggested by Heckman et al. (1997). The ATT can then be written as

$$\Delta_{PSM-DiD}^{ATT} = E\left(Y_{t_{after}}^{1} - Y_{t_{before}}^{1} | P(X), D = 1\right) - E\left(Y_{t_{after}}^{0} - Y_{t_{before}}^{0} | P(X), D = 0\right) \qquad (2)$$

where $t_{after}$ is the post-treatment and $t_{before}$ the pre-treatment period.

An important advantage of using a DiD estimator is that it allows us not only to avoid the bias caused by observables as controlled for by standard propensity score matching, but also by time-invariant unobservables (as empirically shown by Heckman et al., 1997; Smith and Todd, 2005). In fact, we are even able to control for unobservables that may vary in time, but affect participating and non-participating firms in the same way as for instance inflation or business cycles. This assumes of course that in absence of the program both groups would have the same trend in outcomes. In our main analysis we use the year 2003 for matching on pre-program trends. In additional estimations for sub-samples we can show that participants and matched non-participants already developed in a similar manner in terms of firm size several years before treatment (see Figure 3).

## 4.2 Estimation Procedure

We estimate the propensity score by means of a probit model[18] for program participation. Choosing the observable variables according to which matching is performed constitutes a critical step. We are guided by previous research and the specific context of our program. In the literature some disagreement exists over the number of covariates to include. Rubin and Thomas (1996) recommend to better include a variable if there are doubts about whether it is related to the outcome or not. On the other hand including too many variables might exacerbate the problem of finding an area of common support. Against this background, we start with a simple specification, that is firm size, sector and location as most evaluations of firm support do. Next, we iteratively add variables, break them down into more detailed categories, test logarithmic expressions instead of levels, squares or interactions. While we try to retain a high number of sample observations, ultimately, the model needs to satisfy the balancing properties as discussed in more detail below. We use the following variables in our base specification: firm size in terms of both absolute number of employees and twelve different categories.[19] Industry is divided into fifty-eight categories based on NACE codes. For geographic location we use the twenty-eight districts of Bulgaria. We add dummy variables for ownership type and gender of owner. For a reduced sample, we additionally estimate a model including further covariates (sales, profit and productivity) and pre-program information backwards until the year 2000. It should be noted that all covariates were collected before participation in the program (at year-end 2003) to avoid the program affecting the firm's characteristics used to simulate the selection process.

In a next step, we carry out the matching for all pairwise combinations. This is not fully straightforward, because the likelihood of observing two units with exactly the same propensity score is in principle zero, since P(X) is a continuous variable. Various matching methods have consequently been proposed to overcome this problem. We test the four most widely used ones (nearest neighbor, caliper, radius and kernel matching). The nearest-neighbor matching method assigns a weight equal to one, takes each participant in turn and identifies the non-participant with the closest propensity score. It can be implemented with replacement, so that a non-participant can be used more than once as a match. This increases the average quality of the matches, but it can lead to higher variance for the

[18] Caliendo and Kopeinig (2008) show that logit and probit models yield similar results for binary treatments.

[19] We use the following firm size categories: 1-4, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-99, and 100-249 employees.

estimated ATT. Although intuitive, the nearest-neighbor technique must be considered very carefully, since it will find a nearest neighbor even with a very different propensity score, if there is no closer unit. The caliper matching method instead chooses the nearest-neighbor within a caliper of a certain bandwidth and thus imposes a quality control on the match by setting a tolerance level on the maximum propensity score distance. A variant of caliper matching is referred to as radius matching with the idea to use not only the nearest-neighbor but all non-participants within the caliper. Kernel matching also uses all non-participants. The kernel is a function that weighs the contribution of each non-participant, so that more importance is attached to those non-participants providing a better match.

With either method we ensure that the matching procedure balances the distribution of observable variables between participants and non-participants. A participant and a matched non-participant might both have the same propensity scores, but this does not guarantee that they are similar in a relevant way: one firm might for instance be active in a sector with a high propensity for program participation, but from a location with a low participation propensity, while for another firm with a similar propensity score it might just be the other way round. We perform several balancing tests: First, we test that each variable has the same mean in the group of participants and non-participants. To this end, two-sided t-tests of equal means are conducted before (unmatched) and after matching (matched). A second test is on pseudo-$R^2$. Its basic idea is to re-estimate the propensity score on the matched sample, that is only on participants and matched non-participants, and compare the pseudo-$R^2$s before and after matching. A successful matching should lead to a low pseudo-$R^2$. Third, we test the Likelihood-ratio (LR) on the joint significance of all regressors in the probit model. The hypothesis we test is that the variables in the probit estimation have no explanatory power after matching. Finally, we also check the mean standardized bias as suggested by Rosenbaum and Rubin (1985) defined as the percentage of the square root of the average sample variances in both groups. If balance is not achieved, we need to revise the model until we attain balance.

# 5 RESULTS

## 5.1 PROPENSITY SCORE ESTIMATION

The probit estimation results for participation probability are presented in Table 5.[20] We briefly discuss the main components that influence the selection into treatment. The binary outcome takes a value of one if a firm receives a microloan from the program. As expected, firm size affects the probability of program participation. The smaller a firm, the more likely it participates in the program. Sectors jointly have a significant influence on participation, but single sectors also influence selection into treatment. In particular, being active in wholesale trade, water transport, financial intermediation as well as electricity, gas and hot water supply positively influences the probability of participation. Location in terms of districts jointly influences participation. At the same time, regional provenance from the majority of districts also has a significant influence. Finally, ownership type and gender of owner contributes to explaining participation in the program. The pseudo $R^2$ of the estimation is 0.095. Note also that the number of observations has decreased from 61,006 to 58,665 compared to the original sample in Table 4. This is caused by the exclusion of some non-participants from the probit estimation that were active in sectors and districts where there are no participants to be compared to.

*(Table 5 about here)*

We next impose the common support assumption by dropping treated observations whose propensity score is higher than the maximum or less than the minimum score of the untreated. This method is

[20] We use the Stata module PSMATCH2 developed by Leuven and Sianesi (2003) for matching.

also described by Caliendo and Kopeinig (2008) as 'Minima and Maxima Comparison'.[21] As indicated in Table 6, this results in 4 out of 974 participants that are off support because their propensity scores are larger than the maximum scores for non-participants. As for minima, no participant has a lower propensity score than any non-participant, and we do not need to drop any further observations. In sum, there is a very substantial region of overlap, and the estimated effect on the remaining firms in terms of support can be viewed as representative of the full sample.

*(Table 6 about here)*

The balancing tests discussed earlier furthermore show that the matching is successful. Table 7 summarizes the different quality measures after applying caliper matching. First, we test whether each variable has the same mean for participants and non-participants. The t-tests are conducted before and after matching. We clearly see when looking at the first and second column of mean values in Table 7 that the unmatched participants and non-participants had very different characteristics. However, after matching, we cannot reject the hypothesis that the means are equal between the two groups. In fact, we can see that more than half (48 out of 87) of the variables have a mean that is significantly different between participants and non-participants at the 10 percent level before matching takes place, given the p-value in the last column for the two-sided t-test. In the matched sample in turn, we find no significant differences at all. Next to the mean values, we also show the percentage reduction of bias, which is the relative reduction of the difference in mean values for a covariate. It shows that matching reduces the relative bias for all covariates. This result of the t-test of equal means indicates that matching has been successful. The second test is for pseudo-$R^2$. A successful matching should lead to a low pseudo-$R^2$. At the bottom of Table 7, we indicate that the pseudo-$R^2$ decreased substantially after matching from 0.095 to 0.012. Third, the p-value of 1.000 for the test of the LR statistics shows that it cannot be rejected that the variables in the probit estimation have no explanatory power after matching. Finally, it can be seen that the mean standardized bias declines from initially 7.8 percent to 2.1 percent after matching, where a value below 3 to 5 percent generally indicates the success of the matching approach. To summarize, it can be noted that our matching passes all balancing tests. We can proceed to estimating treatment effects.

*(Table 7 about here)*

## 5.2 Main Treatment Effect

We estimate the average treatment effect on the treated (ATT) as defined in Eq. (2) for the main evaluation period until the end of the second year after treatment, 2006. Our outcome variable is the change in number of employees since 2003 before treatment. In Table 8, we report results based on a caliper matching technique with a bandwidth of 0.05. Most important, the results show that the program had a very positive and statistically significant effect on the change in employees for participating firms. To be precise, the participating firms had on average 2.480 more employees after receiving a microcredit than matched non-participants. Observing each group separately, participants increased by 2.722 employees after treatment, while matched non-participants only grew by 0.241 employees. Relative to the number of employees before treatment – participants had 7.478 employees on average in 2003 – this translates into a growth difference in terms of firm size of 33 percent for participants compared to matched non-participants.

A look at the absolute number of employees before treatment in 2003 and after the main evaluation period in 2006, as shown in the lower part of Table 8, complements the picture. First, it reveals that while the difference in firm size in 2003 was large between the two groups before matching (12.497

[21] Another option to impose common support is to use a trimming procedure whereby a certain percentage of the treated observations are dropped for which the propensity score density of the untreated is the lowest. For our analysis, both minima and maxima comparison as well as trimming produces similar results.

for unmatched non-participants), matching was successful in identifying a group of comparable non-participants with a similar firm size (7.478 for matched participants and 7.087 for matched non-participants). Secondly, it shows that unmatched non-participants slightly shrank from 12.497 employees to 12.451 employees over the period 2003 to 2006. At the same time, microcredit participants grew from 7.478 to 10.200 employees, while the matched non-participants grew from 7.087 to 7.328. In sum, microcredit considerably and significantly increased employment two years after treatment, while the average sample firm did not grow.

*(Table 8 about here*)

For comparison, we also present results for other matching algorithms in Table 9.[22] As can be seen, the results from all matching estimators are significant and essentially have the same implications. The main ATT for change in employees until 2006 ranges from 2.280 employees based on nearest neighbor matching without replacement to 2.667 employees based on kernel matching with a bandwidth of 0.05 and a Gaussian weighting function. The closeness of results from different matching algorithms is not surprising according to Caliendo and Kopeinig (2008). With a growing sample size, all matching estimators become closer to comparing only exact matches and should yield the same results. Dehejia and Wahba (2002) even argue that the choice of matching algorithms is not as crucial as the proper estimation of the propensity scores. Nevertheless, finding similar results across different parameters serves as a first robustness check and strengthens our confidence in the results.

*(Table 9 about here*)

### 5.3 DYNAMIC EFFECTS

While the previous results estimate the ATT for the main evaluation period until 2006, we now focus on analyzing the dynamic pattern of treatment effects. In other words, our interest is to disentangle the effect of the program to understand if it is constant or varies over time. We are able to do so because information on the number of employees is available for several years after treatment. In addition, we can track firm exits in our dataset and hence control for sample attrition over a considerable time span. Table 10 provides results of the estimated ATT for each year after treatment until 2010. It also shows the yearly development of absolute employee numbers for participants, matched non-participants and the full sample of non-participants, which is, in addition, graphically illustrated in Figure 2. Since the number of observations varies slightly per year depending on the availability of information, we indicate sample size in the last two rows of Table 10. Note, also, that estimates are again very similar across different matching algorithms.

*(Table 10 and Figure 2 about here*)

As Table 10 and Figure 2 show, the loan effects are substantially positive and long lasting, as significant differences between the participating and matched non-participating firms are maintained several years after reception of the microcredit. A largely positive and statistically significant effect already takes place in the treatment year 2004 with a difference of 2.598 more employees for microcredit participants. The strong program effect lessens towards 2005 to a difference of 1.289, before recovering in 2006 with 2.480. Afterwards, the impact even further increases to 2.964 in 2007 and 3.803 in 2008 before it decreases to 2.040 in 2009. In 2010, the last available point in time in our dataset and six years after treatment, the ATT still amounts to 2.110.

What attracts attention is the remarkable dip in treatment effect for 2005. As can be seen in the second column of Table 10, the microcredit participants first grow strongly from a baseline size of 7.478

---

[22] According to Imbens and Wooldridge (2009) bootstrapping of standard errors is not valid for matching estimators with a fixed number of matches such as nearest-neighbor and caliper matching. It is likely that the problems invalidating the bootstrap disappear if the number of matches increases with sample size. We therefore bootstrapped standard errors with 50 replications for the kernel matching results only.

in 2003 to 10.286 employees until the end of the treatment year in 2004. Afterward they contract to 8.790 employees in 2005 before recovering to 10.200 in 2006. During the same time period, the matched and unmatched non-participants more or less stagnate at 7.087-7.328 employees and 12.07-12.497 employees, respectively, as shown in the third and fourth column of Table 10. The development of non-participants is in line with an economic context during that period that did not give reason for notable yearly changes. Moreover, the increase and subsequent drop in employees also does not mirror the underlying business fundamentals of participants. Financial indicators on sales and profit for participants show constant growth over the period 2003 to 2006.[23] There is no dip for financials in 2005 as is the case for employee numbers. We thus suppose that the dramatic upward and downward development of firm size for microcredit participants is rather to be explained by the program itself. In anticipation or immediately after disbursement of the microcredit, participating firms largely expanded their businesses in terms of employees within the year of treatment, 2004. Presumably, this increase was oversized and unsustainable with regard to underlying business fundamentals, as financials did not increase to the same degree. This consequently triggered a clear contraction in terms of employees within the first year after treatment, 2005, although the average firm size remained higher than before treatment. As with the previous increase, the decrease was again disproportional compared to the financials, and the average firm size for participants settles down at an apparent equilibrium of 10.200 employees in 2006. Another conjecture for the dip in 2005 would be that the expansion into new business areas as enabled by the microloan required a larger number of employees in 2004 than in 2005, when new business processes had become more efficient and more experienced employees needed less time to perform the same tasks. Either way, it seems desirable to consider an evaluation period of several years after treatment for microfinance programs in order to allow treatment effects on firm size to stabilize.

Another noticeable development that is mirrored in the dynamic treatment effects is the impact of the global economic crisis. If we look at the numbers of employees in Table 10, we see that after several years of growth, 2009 marks a clear drop in firm size. Most plausibly, this reduction of employees is to be explained by the economic downturn that followed the global financial crisis of 2007/2008. The crisis was manifested in the downward movement of GDP as well as employment in the Bulgarian economy in the last quarter of 2008. GDP shrank by around 5 percent, and unemployment jumped in 2009. As studies by Wagner and Winkler (2013) show, microfinance institutions have been no less vulnerable to the crisis than other financial institutions. While the crisis negatively affected both participants and non-participants, there is a difference in timing. Unmatched non-participants in our dataset start contracting in terms of firm size one year earlier in 2008 than do participants and matched non-participants. One conjecture for this earlier sensitivity of the average Bulgarian firm to the economic shock is that microcredit clients and their matched comparators are more growth oriented and withstand economic pressures longer than other firms. The time lag in decreasing firm sizes could also be related to the unmatched firms originating from sectors and locations that are less resistant to crises.

### 5.4 Heterogeneous Effects

In the following, we take a closer look at the heterogeneity of the estimated treatment effect. We re-estimate the ATT for various sub-groups of our sample. In particular, we are interested if different loan and firm sizes influence effects of microcredit on wage-employment. In addition, we also investigate heterogeneous effects based on loan purposes and firm location. The results are summarized in Table 11. Note that we show the ATT both in terms of absolute change in the number of employees as well as in terms of relative change compared to the baseline number of employees. For sub-groups with very different baselines, it is helpful to have information on both.

---

[23] Sales increase from 172 Thousand Euros (2003) to 193 TEUR (2004) to 234 TEUR (2005) to 295 TEUR (2006). Profit increases from 86 TEUR (2003) to 108 TEUR (2004), 129 TEUR (2005) and 119 TEUR (2006).

### 5.4.1 Loan Amount

We first investigate whether the effect of microcredit varies with loan size. As argued in the beginning, existing studies point into the direction that larger loans have a more positive impact on wage-employment than smaller loans. There might even exist a certain minimum loan amount necessary for positive wage-employment effects to unfold. In order to test this hypothesis we conducted our propensity score matching analysis for three equally large sub-groups of participants based on loan amount: small (< EUR 2,000 EUR), medium (EUR 2,000 – 6,000) and large (> EUR 6,000). Since some participants received more than one loan during treatment, loan amounts reflect the total amount of loans disbursed in 2004 per client. As can be seen in Table 11, the impact of all three size categories of microcredit is significant, positive and very similar. The ATT is 27.6 percent for small, 30.0 percent for medium and 25.0 percent for large program loans. We consequently do not find any evidence that larger loans unfold larger effects. At the same time, our smallest loan size category is already considerably above the average loan amount for other regions. Loans up to EUR 2,000 at market exchange rates would be equivalent to USD 6,650 at purchasing power parity. We therefore tested the hypothesis again for loan amounts between EUR 500-1,000 (USD 1,664 – 3,328 PPP adjusted) as well as below EUR 500 (USD 1,664 PPP adjusted). With regard to loans between EUR 500-1,000 the estimated effect is significant at the 10% level and again similar to that for larger loan amounts (23.9 percent). As far as loans below EUR 500 are concerned, we also find a similar but insignificant effect (26.8 percent). The number of sample firms with very small loans is too small for carrying out reliable estimates. This is also the reason why we cannot assess the impact of yet smaller loan amounts which would have been very interesting as loan sizes of existing impact assessments range between USD 726-1,616 PPP adjusted. Therefore, while we do not find evidence, that larger loans have a more positive impact, our dataset with a large average loan size and a very small number of participants with loans below EUR 500 might also be ill-suited for testing such a hypothesis.

*(Table 11 about here)*

### 5.4.2 Firm Size

We next investigate whether the effect of microcredit varies with firm size. Empirical evidence as discussed in section 2 shows that the smaller a firm, the less access it has to formal bank loans. The more a firm is constrained by credit, the higher the relative impact of microcredit on its growth should consequently be. We use the typical firm size categories micro (1-9 employees), small (10-49) and medium (50-249). The results in Table 11 confirm findings from other studies: The smaller a firm, the higher the impact of microcredit on employee growth. Micro-sized firms have on average an ATT of 50.4 percent more employees for participants than non-participants in 2006 compared to 2003. Small-sized participating firms increase by 24.3 percent more over the same time period, while the treatment effect for medium-sized firms is even negative but insignificant due to the small number of observations for medium-sized participants.

### 5.4.3 Loan Purposes

An additional analysis is based on different loan purposes. The bank distinguishes between loans for working capital and investment. Working capital refers to current or short-term assets to be used within one year (e.g., wood for carpentry or food and goods for retail vending). Investment loans are those made for the purchase of fixed assets that are used over time of the business. These assets typically have a life span of more than one year, such as machinery, equipment or property. Loans for investments in fixed assets are generally larger and have longer terms than working capital loans. Since the productive activity does not directly use up the fixed asset (that is, it is not sold as part of the product), economic theory suggests that the impact of investment loans upon profitability (and business expansion in terms of employees) is felt over a longer period of time (Ledgerwood, 2013). The impact of working capital loans, on the other hand, is supposed to show instantaneously and diminish over time. Our dataset largely supports this hypothesis. The estimated impact in 2006 of 2.301 for working capital loans and 1.799 for investment loans in Table 11 points to the more immediate impact of working capital over investment loans. A longer-term perspective however shows that from 2008 onwards the relation reverses, and investment loans have a higher impact than working capital loans (2.932 in 2008, 2.972 in 2009 and 2.040 in 2010 for investment loans compared to 1.627 in 2008, 1.054 in 2009 and 1.608 in 2010 for working capital loans). The indicated impact for working capital loans from 2008 onwards furthermore becomes insignificant. Although it remains debatable for how long after treatment an impact can be attributed to the original microloan, the findings suggest that the period over which loans display their impact is affected by the loan purpose. It therefore seems advisable to take the type of loan purpose into consideration when deciding over the appropriate evaluation period for microfinance impact assessments. In addition, our findings do not add support to the frequently made caveat that because money is fungible (e.g., Banerjee, Duflo, et al., 2015; Ledgerwood, 2013), loans intended for a specific purpose are often used for a different purpose within a household or business.

#### 5.4.4 Location

Finally, we make use of the fact that the microcredit program was not present in all Bulgarian districts at the time of treatment. As Figure 1 shows, ProCredit Bank Bulgaria had branches in 19 out of 28 Bulgarian districts during treatment in 2004. Observations for non-participants are available for all 28 districts. We now exploit this institutional feature of the program by including only firms from the 9 districts without program presence as potential comparators in the matching regressions. It should be noted that we had to drop 28 of the 974 participants because they were located in districts with no actual presence of the program but still participated. Overall, propensity score matching again results in a significantly positive ATT of an increase in employees of 2.396 more for participants than for non-participants from districts without the program. Receiving a similar ATT to our main estimation when matching on a sample of non-participants from areas without the program is an important finding with regard to our research design. For areas without program presence, we have good reason to believe that access to alternative credit (a concern addressed in section 2) was in general much more limited than in program areas. This is the case because ProCredit Bank followed a typical strategy for financial institutions by focusing on operationally more accessible areas first. Hence, it is very likely that other providers of microcredit also had limited presence in areas without the program. Obtaining an ATT of 2.396, which is similar to the ATT of 2.480 from our main estimation, supports our hypothesis that in general, little alternative credit was available for non-participants in both program and non-program areas.

In sum, the purpose of this section was to investigate which sub-groups of our sample have experienced enhanced growth as a result of microcredit. This way we were able to test several hypotheses: We first tested whether larger microloans have stronger impacts on wage-employment. We do not find evidence for this hypothesis which might however be due to the relatively large loan amounts by the assessed program. The number of very small microloans is too limited for fully testing the hypothesis. We additionally assessed if smaller firms are more constrained by credit than larger firms. We can confirm this hypothesis as the treatment effect increases with decreasing firm size. Moreover, our findings point to the fact that the periods over which loans display their impact is affected by the purpose of the loan. Working capital loans show a more immediate effect, while loans for investment purposes increase their impact over time. Finally, we are able to show that restricting the sample of non-participants to areas without program presence results in similar treatment effects as those from our main estimation. This finding supports our hypothesis that alternative credit to non-participants was still limited during the period of evaluation, increasing confidence in our research design.

# 6 SENSITIVITY TESTS

After having presented strong positive effects of the program, we test the robustness of our results. The advantage of well-performed matching is that it can account for selection on observables. If participation in microcredit can be understood with observable variables, then matching produces consistent results. To add further strength to our estimates in this regard, we will test the appropriateness of the specified model and its sensitivity to small variations. In addition, we will compare our main results with propensity score matching to treatment effects estimated through a linear regression model with ordinary least squares.

Even if selection on observables is well accounted for, one major drawback of matching in contrast to randomized experiments remains nevertheless: selection on unobservables is not controlled for. By complementing propensity score matching with difference-in-differences, we try to control for any pre-existing constant (time-invariant) differences in the outcome variable before treatment, even if those differences are caused by unobservable attributes. Still, participants and non-participants could differ in terms of time-variant unobserved characteristics. This would violate the conditional independence assumption and bias our results. It is not possible to directly test the conditional independence assumption and the magnitude of selection bias with our non-experimental data. We can however assess the sensitivity of our results to remaining unobservables and will do so by applying three types of analyses: a recently developed test of exogeneity by G. Caetano et al., 2016, the so-called bounding approach as suggested by Rosenbaum (2005) and simulation analyses as applied by Ichino et al. (2008).

*(Table 12 and Figure 3 about here)*

## 6.1 VARIATIONS WITHIN MATCHING PROCESS

To start with, an initial robustness check in terms of variation within the matching process was already built in our presentation of main results. We jointly considered different matching methods in Table 9. The fact that matching with different methods such as caliper, nearest neighbor, radius and kernel estimators produces very similar results, is a positive sign for robustness. Here, we additionally extend the set of variables in the propensity score estimation in order to see whether this has an impact on the causal estimates. As discussed above, data for sales and profit as well as historical trends before start of the treatment in 2003 are only available for a limited amount of sample firms. We consequently restricted our main specification to variables which allow for a large sample size. Now, we make use of this additional information.

First, we go further backwards in time, so as to match on longer trends before start of the treatment. The point of reference is the base ATT for change in employees until 2006 of 2.480 as shown in the first row of Table 12. Adding information on the number of employees for 2002, 2001 and 2000, that is until 3 years before start of the treatment, increases the estimated program effect to 2.768 compared to baseline results. Using the change in employees from 2003 compared to 2000 instead as further variable for matching, this results in an ATT of 3.057. Here is an important point to make: the inclusion of longer time trends not only confirms the positive effect of microcredit from our main estimates. It additionally allows us to show for a sub-sample that participants and matched non-participants developed almost identically in terms of firm size for several years before treatment. After treatment, only participants start to grow considerably while matched non-participants stagnate. This strongly supports our hypothesis that the increase in firm size for participants compared to otherwise identical matched non-participants (based on observable characteristics) is causally related to treatment with microcredit. Figure 3 graphically illustrates this fact for the period 2000 until 2006 based on an extension of the specification for propensity score estimation by the change in employee numbers from 2003 compared to 2000 (Δ N employees 2003 – N employees 2000). The sub-sample consists of 331 participants and 34,052 non-participants (compared to 970 and 57,691 respectively for base effects).

The finding that participants and their matched non-participants had a very similar development of firm size before treatment finally also relates to the applied difference-in-differences method. It adds support to our assumption that in absence of the program both participants and non-participants would have the same trend in the outcome variable.

We furthermore assess how the inclusion of additional variables for the baseline year 2003 influences results. If we use information on sales, profit and productivity (defined as sales per employee) in 2003 for propensity score matching, the impact increases to 2.780. Finally, we modify the specification with regard to geographic location. As Heckman et al. (1997) show, geographical mismatch of participants and non-participants is typically a large source of bias in impact studies. As a consequence, we here exactly match on district by only including matches of participants and non-participants that are identical in terms of district. It should be noted that we already use districts as covariate for the estimation of propensity scores in our main sample. Many matches used for estimating the base effect therefore already fulfill this condition. If we now explicitly force an exact match on location by district, the effect amounts to a significant increase of 1.691 more employees in 2006 for participants compared to before treatment which is about in the same range as the base effect of 2.480.

## 6.2 OLS ESTIMATES

As another robustness test, we compare our main results with propensity score matching (PSM) to treatment effects estimated through a linear regression model with ordinary least squares (OLS). Both methods are supposed to produce similar results, although PSM has two particular advantages compared to OLS according to Chemin (2008). First, matching compares a participant to a synthetic non-participant only if it is comparable enough, that is in the area of common support. A regression with OLS on the other hand will give a result by interpolating a linear relationship between the two groups, no matter if participants and non-participants are very different. This is an undesirable result if the two groups are not comparable. Second, the matching technique is non-parametric as opposed to a linear regression which imposes a linear structure on the data. Matching does not impose a particular structure and heterogeneous treatment effects are allowed for.

Table 13 summarizes the OLS estimates for different models and compares them to results from PSM. We first estimate a model without control variables, i. e. with only a constant and the treatment variable, while the second model includes all covariates in OLS which were also used for estimating the propensity score. In addition, we performed the OLS estimation for two types of samples: the full sample with all available non-participants (61,006 observations) and the reduced sample with those non-participants that were used for matching (58,665 observations). The full sample uses more information, whereas the reduced sample should provide a better approximation of the ATT with PSM.

*(Table 13 about here)*

As expected, the results before matching (without controls) as shown in the first column of Table 13 are very similar for PSM and OLS. If we look at the matched sample only, the estimates for PSM and OLS are identical with a change of 2.770 in employees when no observable differences between participants and non-participants are accounted for. When we compare our main results from PSM after matching to the OLS estimates with controls in the second column, we receive a lower, but still considerable and significant ATT for OLS on the matched sample of 1.887 which further points to the robustness of our results.

## 6.3 EXOGENEITY TEST, BOUNDING AND SIMULATION ANALYSIS

We first apply a recently developed test of exogeneity by C. Caetano (2015)[24] to assess the sensitivity of our results to unobservables. This test has power when unobservable confounders are discontinuous with respect to the explanatory variable of interest, and it is particularly suitable for applications in which that variable has bunching points.

Applied to our context, we assume that change in employees is a continuous function of microloans. For this purpose, we use loan amount as explanatory variable instead of the binary treatment assignment. While the 974 participants receive a continuous amount of euros, the 60,032 non-participants receive zero euros of loan. Loan amount therefore bunches at zero euros. If now the expected change in employees conditional on loan amount is discontinuous at zero, this cannot be due to loan amount, since loan amount has a continuous effect on the number of employees. This discontinuity may rather be due to selection on observable and unobservable firm characteristics which may be discontinuous at a loan amount of zero. If we control for observable characteristics and the number of employees is no longer discontinuous, we fail to reject exogeneity. If on the other hand the discontinuity remains after observables are included, the only explanation is that there is at least one unobserved confounder and hence the model suffers from endogeneity.

Figure 4 compares results with and without controlling for observable characteristics. The left graph indeed shows a discontinuity of the outcome at a loan amount of zero compared to a loan amount of EUR 1-500 (smallest loan amount category for participants). In detail, non-participants have increased by 1.399 less employees in 2006 than participants with the smallest loans amounts (see Table 14). Controlling for observables in the right graph on the other hand shows that the change in employees is no longer discontinuous and we fail to reject exogeneity.

*(Figure 4 and Table 14 about here*)

Another approach to test sensitivity of our results with regard to unobservables is to assess how strong an unobserved component would need to be so as to undermine our results. For this purpose, we apply bounding as well as simulation analysis. The so-called bounding approach was initially suggested by Rosenbaum (2005) who described the main ideas as follows: bounding introduces a sensitivity parameter Γ that measures the degree of departure from random assignment of treatment. Two subjects with the same observed covariates may differ in their odds of receiving the treatment by at most a factor of Γ (≥ 1). In an experiment, random assignment of treatment ensures that Γ = 1. In an observational study with for example Γ =2 for two subjects which were matched on observed covariates, then one might be twice as likely as the other to receive the treatment because they differ in terms of a covariate not observed. As Γ is of course unknown, bounding tries out several values of Γ to see how results change. For each value of Γ bounds are placed on a statistical inference. We implement bounding by using the *rbounds* procedure in Stata developed by DiPrete and Gangl (2004) for continuous outcome variables. *rbounds* uses the results from the matching estimates to calculate maximum and minimum p-values from a Wilcoxon's signed rank test between matched pairs of participating and non-participating firms for an artificial unobserved variable with different values of Γ. The p-values then represent the bound on the significance level of the treatment effect in the case of endogenous selection into treatment.

It is important to recognize that the results from the Rosenbaum bounds are a ''worst-case'' scenario. DiPrete and Gangl (2004) provide an example where p-values become significant at Γ = 1.15. They point to the fact that such a result does not mean that there is no positive effect of treatment on outcome. It rather means that the confidence interval for the effect would include zero if an unobserved variable caused the odds ratio of treatment assignment to differ between participants and non-participants by 1.15 and if this variable's effect on outcome was so strong as to almost perfectly determine whether the outcome would be bigger for the participants or the non-participants in each pair

[24] See also G. Caetano et al. (2016) for an application of the test to a multivariate context

of matched cases in the data. In the case where an unobserved variable had an equally strong effect on selection into treatment but only a weak effect on the outcome variable, the confidence interval for employment would not contain zero. Nonetheless, the Rosenbaum bounds convey important information about the level of uncertainty contained in matching estimators by showing just how large the influence of an unobserved variable must be to undermine the conclusions of a matching analysis.

*(Table 15 about here)*

Table 15 summarizes the results of the bounding analysis. If unobserved factors lead to negative or positive selection, i. e. those who participate always have a lower or higher growth rate of employees than matched non-participants even in the absence of treatment, the p-values will become significant for a certain value of Γ. In the absence of unobserved heterogeneity, that is Γ =1.00, the p-values are insignificant indicated by a $p < 0.05$. Starting from that point, we stepwise increase the value of Γ and simulate an ascending influence of unobserved factors. At Γ =2.00 the range of possible p-values is from 0.0000 to 0.0004, so a bias of this magnitude could not explain the larger firm size increase for participants. If we had failed to control by matching a variable strongly related to firm growth and two times more common among participants, this would not have been likely to produce a difference in employee growth as large as the one we observe (see Rosenbaum, 2005 for a similar example). The upper bound of the p-value becomes finally significant at the 5% level when Γ = 2.35. Two firms that have the same observable characteristics would then differ in their odds of participating in the microcredit program by a factor of 2.35 or 135 percent. In other words, unobserved factors would need to have about one and a third times the influence as observed factors in order to undermine the results. This can be considered to be a reasonably large number given that we have adjusted for many important observed factors (see also Aakvik, 2001 or Caliendo and Künn, 2011). Consequently, we conclude that our results on the change in employees are robust against strong unobserved selection bias. Put in another way, the influence of an unobserved factor would have to be considerable to undermine the conclusions of our matching analysis.

Since the critical values from bounding are rather abstract, we additionally implement the so-called simulation analysis to better illustrate the magnitude of potential hidden bias which would cause us to revise our findings of causal effects of microfinance on employment levels. Simulation analysis was originally applied as sensitivity test to matching by Ichino et al. (2008). Its basic idea is to start from the conditional independence assumption and then examine how the results change as this assumption is weakened in specific ways, whereas for the bounding analysis the conditional independence assumption was entirely dropped. We perform the analysis by simulating an unobserved variable or so-called confounder that is adapted to the distribution of an observable variable. Since we exactly know the influence of the observable characteristics on outcomes and selection, we therefore have a direct linkage to the potential unobserved effect for interpretation. Results are shown in Table 16.

*(Table 16 about here)*

The first two columns in Table 16 show the influence of each confounder on the untreated outcome and on the selection into treatment. Thereby, a value below (above) 1 indicates a negative (positive) impact. The third column shows the resulting ATT given the existence of a confounder with a certain distribution. The last column contains the standard error of the ATT. To facilitate a comparison between actual and simulated results, the first row of Table 16 shows the baseline ATT estimate obtained with no confounder in the matching set. In the absence of a confounder, its influence on outcome and selection is obviously zero and the ATT is 2.361 from nearest-neighbor matching. The following rows of Table 16 show how the baseline estimate changes when the confounding factor is calibrated to mimic different observable covariates and is then included in the set of matching variables. If a confounder is introduced which has the identical distribution as the firm size category 5-9 employees, the influence on outcome (0.714) and on selection (1.017) would be negative. This means that such an unobserved term would have a negative effect on the change in employees in case of no treatment and also on being treated at all. Including this simulated unobserved confounder leads to an

ATT of 1.948 which is within the range of the ATT in the absence of unobserved heterogeneity. We tested other confounders as well. A confounder with the distribution of 'Ownership type (Joint Stock Company)' would lead to an ATT of 1.873. The simulation of a confounder with a distribution of 'Gender of owner (female)' finally results in an ATT of 1.887. Taken in conjunction, these simulations result in ATTs which are of a similar dimension as the baseline ATT. This means that the existence of any unobservables which influence selection or outcome by a similar magnitude as certain observable variables, would still not largely alter the positive effect of the microfinance program on creating jobs for participating firms.

To repeat, no test can directly justify the conditional independence assumption. However, the performed exogeneity test as well as the bounding and simulation analyses convey an impression of robustness of the baseline matching estimates. Overall, the presented sensitivity analyses in this section find that the direction and dimension of our estimated treatment effect holds which increases confidence in the robustness of our results even with respect to unobserved heterogeneity.

# 7 CONCLUSIONS

In this paper, we analyze the impact of a microfinance program in Bulgaria. The program is intended to ease credit constraints of existing firms and thus contribute to the creation of wage-employment. The unique firm-level panel data provides an important opportunity to address two research gaps with regard to microfinance impact assessments: First, we are able to add to the scarce literature on wage-employment effects of microfinance rather than self-employment as most other studies are limited to. We do so by providing evidence for microloans which are larger in size than what has been studied so far. Second, drawing on empirical evidence from a program in Bulgaria with typical features for the region, the study also contributes to the largely neglected area of impact assessments for microfinance in more developed contexts such as Eastern Europe.

We base our analysis on propensity score matching and extend the standard approach to a difference-in-differences matching estimator to assess the effectiveness of participating in the microcredit program as opposed to non-participation. Our identifying assumption is that conditional on the information in our dataset, selection into the program can be assumed to be random such that differences in outcome between participants and non-participants can be interpreted as causal effects. To justify this very strong assumption, we carefully assess the sensitivity of our results. The application of different matching estimators does not alter results. The extension of variables for the baseline year or the inclusion of longer time trends neither has an impact on the causal estimates. In fact, the consideration of longer time trends before treatment suggests that in the absence of the program, both participants and non-participants would have the same trend in the outcome variable. Our results with propensity score matching are furthermore similar to treatment effects estimated through a linear regression model with ordinary least squares. Regarding sensitivity to unobservables, the application of a recently developed test of exogeneity shows that we cannot reject exogeneity in our model. The outcome is no longer discontinuous once we control for observable characteristics. The so-called bounding approach furthermore detects that unobserved factors would need to be large (1.35 times the influence of observables) in order to undermine the results. In the same manner, the simulation of an unobserved variable that is adapted to the distribution of an observable variable would not largely alter our estimates.

As main result, we find a relatively large positive and significant effect of microcredit on wage-employment. Participating firms have on average 2.48 more employees two years after receiving a microcredit than do matched non-participants. Related to the number of employees before treatment, this translates into a growth difference in terms of firm size of 33 percent for participants. The estimated effect seems very plausible when comparing labor costs of an additional employee to loan size. Average annual labor costs in Bulgaria amounted to about EUR 2,600 at the end of 2006 (including wage, social contributions and taxes paid by the employer).[25] An increase in 2.48 employees would hence cost EUR 6,448 which is very close to the average loan size of EUR 7,092.

The microcredit program also has a long lasting impact on wage-employment. Significant differences between the participating and matched non-participating firms exist for at least six years after reception of the microcredit. At the same time, the increase in the number of employees for participants is characterized by initial upward and downward adjustments during the first two years after treatment before stabilizing at a higher level of firm size. Consequently, only an evaluation period of two or more years after treatment seems able to capture the sustainable impact of microfinance on firm growth.

Heterogeneous effects for sub-groups first of all support the finding from other studies that smaller firms are more constrained by credit than are larger firms. The treatment effect increases with decreasing firm size and is largest for micro-enterprises. Our results for different loan purposes moreover point to the fact that the time at which the impact occurs is affected by what the loan is used for. While working capital loans show a more immediate effect, loans for investment purposes increase their impact over time. It therefore seems advisable to take into consideration the type of loan purpose when determining the appropriate evaluation period for microfinance impact assessments. As far as loan size is finally concerned, our data only provides limited scope for assessing whether larger loans lead to larger impacts due to very few small loans in our sample. With respect to loans above what has been studied so far, i.e. loans above EUR 500 (or USD 1,664 PPP-adjusted) we find a very similar impact of 25-30 percent on wage-employment irrespective of the loan amount. This finding might rather point to a certain loan size threshold for positive effects to unfold and that the loans assessed in this study are above that threshold. That would also explain why the impact on wage-employment found in this study is considerably larger than what has been estimated so far. Another type of explanation could be related to the specific macroeconomic and institutional environment for firms in Eastern European which helped microcredit to unfold such positive effects on wage-employment. Ahlin et al. (2011), for example, find that the success of microfinance is very context-specific. While they focus on the performance of microfinance providers, this may be true in a more general manner for microfinance's impact. In this regard, it would be very interesting to see further studies in the future for programs in similarly developed countries as Bulgaria and compare our results to. It furthermore could be very worthwhile for future research to study effects in a similar context but for a program with a larger number of very small loans comparable to existing studies. This would allow to test if there indeed exists a minimum loan amount for positive effects of microfinance on wage-employment.

With regard to the broader economic implications for the labor market, an increase of 2.5 employees (33%) per loan recipient (i.e., 2,416 jobs for 974 recipients) might at first not appear transformative. Given that our sample of participants was limited to legal entities among borrowers and that the program was rapidly expanding, we expect the full economic impact to be more pronounced. By assuming that the private individuals among borrowers in 2004 in relative terms grew just as much as the evaluated legal entities the total effect would rather amount to almost 10,000 jobs for a single year. Finally, in terms of not only benefits but also costs of microfinance, it must be stressed that the bank was operating in a profitable manner and the program can be regarded as financially self-sustaining. In light of these findings, it seems that the high expectations of policymakers for the job creating effects of microcredit can indeed be justified for programs which are characterized by relatively large loan sizes in more developed contexts.

[25] See: https://infostat.nsi.bg/infostat/pages/module.jsf?x_2=79 (Accessed: 24 February 2017).

# 8 ACKNOWLEDGEMENTS

I am grateful to Thorsten Schank for his advice and encouragement. I would also like to thank the editor Edwin Leuven and an anonymous referee, Adalbert Winkler as well as participants of the 2nd Microfinance and Rural Finance Conference 2016 (Aberystwyth), the EALE Conference 2016 (Ghent) and the seminar at the Gutenberg School of Management and Economics (Mainz) for a number of very useful comments and discussions. I also thank the team of ProCredit Bank Bulgaria for their valuable feedback regarding all questions related to the program. Financial support from ProCredit Bank Group and from the German Academic Exchange Service (DAAD) is gratefully acknowledged.

# 9 REFERENCES

Aakvik, A. (2001). Bounding a matching estimator: the case of a Norwegian training program. *Oxford Bulletin of Economics and Statistics, 63* (1), 115-143. doi:10.1111/1468-0084.00211

Abou-Ali, H., El-Azony, H., El-Laithy, H., Haughton, J., & Khandker, S. (2010). Evaluating the impact of Egyptian Social Fund for Development Programmes. *Journal of Development Effectiveness, 2* (4), 521-555. doi:10.1080/19439342.2010.529926

Ahlin, C., Lin, J., & Maio, M. (2011). Where does microfinance flourish? Microfinance institution performance in macroeconomic context. *Journal of Development Economics, 95* (2), 105-120. doi:10.1016/j.jdeveco.2010.04.004

Angelucci, M., Karlan, D., & Zinman, J. (2015). Microcredit impacts: Evidence from a randomized microcredit program placement experiment by Compartamos Banco. *American Economic Journal: Applied Economics, 7* (1), 151-182. doi:10.1257/app.20130537

Armendáriz de Aghion, B., & Morduch, J. (2000). Microfinance beyond group lending. *Economics of Transition, 8* (2), 401-420. doi:10.1111/1468-0351.00049

Armendáriz de Aghion, B., & Morduch, J. (2010). *The Economics of Microfinance* (2nd ed.). Cambridge, Massachusetts: MIT Press.

Attanasio, O., Augsburg, B., De Haas, R., Fitzsimons, E., & Harmgart, H. (2015). The impacts of microfinance: Evidence from joint-liability lending in Mongolia. *American Economic Journal: Applied Economics, 7* (1), 90-122. doi:10.1257/app.20130489

Augsburg, B., De Haas, R., Harmgart, H., & Meghir, C. (2015). The Impacts of Microcredit: Evidence from Bosnia and Herzegovina. *American Economic Journal: Applied Economics, 7* (1), 183-203. doi:10.1257/app.20130272

Ayyagari, M., Demirguc-Kunt, A., & Maksimovic, V. (2010). Formal versus informal finance: Evidence from China. *Review of Financial Studies, 23* (8), 3048-3097. doi:10.1093/rfs/hhq030

Ayyagari, M., Demirgüç-Kunt, A., & Maksimovic, V. (2008). How important are financing constraints? The role of finance in the business environment. *The World Bank Economic Review, 22* (3), 483-516. doi:10.1093/wber/lhn018

Banerjee, A. (2013). Microcredit under the microscope: what have we learned in the past two decades, and what do we need to know? *Annual Review of Economics, 5* (1), 487-519. doi:10.1146/annurev-economics-082912-110220

Banerjee, A., & Duflo, E. (2014). Do firms want to borrow more? Testing credit constraints using a directed lending program. *Review of Economic Studies, 81* (2), 572-607. doi:10.1093/restud/rdt046

Banerjee, A., Duflo, E., Glennerster, R., & Kinnan, C. (2015). The miracle of microfinance? Evidence from a randomized evaluation. *American Economic Journal: Applied Economics, 7* (1). doi:10.1257/app.20130533

Banerjee, A., Karlan, D., & Zinman, J. (2015). Six randomized evaluations of microcredit: introduction and further steps. *American Economic Journal: Applied Economics, 7* (1), 1-21. doi:10.1257/app.20140287

Beck, T., Demirgüç-Kunt, A., Laeven, L., & Maksimovic, V. (2006). The determinants of financing obstacles. *Journal of International Money and Finance, 25* (6), 932-952. doi:10.1016/j.jimonfin.2006.07.005

Beck, T., Demirgüç-Kunt, A., & Maksimovic, V. (2005). Financial and Legal Constraints to Growth: Does Firm Size Matter? *Journal of Finance, 60* (1), 137-177. doi:10.1111/j.1540-6261.2005.00727.x

Becker, S., & Ichino, A. (2002). Estimation of average treatment effects based on propensity scores. *Stata Journal, 2*(4), 358–377. (http://www.stata-journal.com/sjpdf.html?articlenum=st0026 [accessed 12 October 2016])
Brown, J. D., Earle, J. S., & Lup, D. (2005). What makes small firms grow? Finance, human capital, technical assistance, and the business environment in Romania. *Economic Development and Cultural Change, 54* (1), 33-70. doi:10.1086/431264
Brown, M., Ongena, S., Popov, A., & Yeşin, P. (2011). Who needs credit and who gets credit in Eastern Europe? *Economic Policy, 26* (65), 93-130. doi:10.1111/j.1468-0327.2010.00259.x
Bruhn, M., & Love, I. (2014). The real impact of improved access to finance: Evidence from Mexico. *Journal of Finance, 69* (3), 1347-1376. doi:10.1111/jofi.12091
BSMEPA. (2004). Small and Medium-Sized Enterprises in Bulgaria, 2002-2003  Sofia: Bulgarian Small and Medium Enterprises Promotion Agency. Retrieved from http://images.mofcom.gov.cn/bg/accessory/200704/1176275888467.pdf [accessed 12 October 2016]
Budina, N., Garretsen, H., & De Jong, E. (2000). Liquidity Constraints and Investment in Transition Economies. *Economics of Transition, 8* (2), 453-475. doi:10.1111/1468-0351.00051
Caetano, C. (2015). A Test of Exogeneity Without Instrumental Variables in Models With Bunching. *Econometrica, 83* (4), 1581-1600. doi:10.3982/ecta11231
Caetano, G., Kinsler, J., & Teng, H. (2016). Towards Causal Estimates of Children's Time Allocation on Skill Development. (https://media.terry.uga.edu/socrates/publications/2016/05/kinsler_WP_time_allocation.pdf [accessed 2 February 2017])
Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys, 22* (1), 31-72. doi:10.1111/j.1467-6419.2007.00527.x
Caliendo, M., & Künn, S. (2011). Start-up subsidies for the unemployed: Long-term evidence and effect heterogeneity. *Journal of Public Economics, 95* (3-4), 311-331. doi:10.1016/j.jpubeco.2010.11.003
CGAP. (2016). Taking Stock: Recent Trends in International Funding for Financial Inclusion. *CGAP Brief*. (http://www.cgap.org/sites/default/files/Brief-Taking-Stock-Recent-Trends-in-International-Funding-for-Financial-Inclusion-Dec-2016.pdf [accessed 22 March 2017])
Chemin, M. (2008). The Benefits and Costs of Microfinance: Evidence from Bangladesh. *Journal of Development Studies, 44* (4), 463-484. doi:10.1080/00220380701846735
Coad, A., & Hölzl, W. (2012). Firm growth: Empirical analysis. In M. Dietrich & J. Krafft (Eds.), *Handbook on the Economics and Theory of the Firm* (pp. 324-338). Cheltenham, UK: Edward Elgar.
Coleman, B. E. (1999). The impact of group lending in Northeast Thailand. *Journal of Development Economics, 60* (1), 105-141. doi:10.1016/S0304-3878(99)00038-3
Coleman, B. E. (2006). Microfinance in Northeast Thailand: Who benefits and how much? *World Development, 34* (9), 1612-1638. doi:10.1016/j.worlddev.2006.01.006
Crépon, B., Devoto, F., Duflo, E., & Parienté, W. (2015). Estimating the impact of microcredit on those who take it up: Evidence from a randomized experiment in Morocco. *American Economic Journal: Applied Economics, 7* (1), 123-150. doi:10.1257/app.20130535
Cull, R., Demirguc-Kunt, A., & Morduch, J. (2009). Microfinance Meets the Market. *Journal of Economic Perspectives, 23* (1), 167-192. doi:10.1257/089533009797614036
de Kok, J., Vroonhof, P., Verhoeven, W., Timmermans, N., Kwaak, T., Snijders, J., & Westhof, F. (2011). Do SMEs create more and better jobs? doi:10.13140/2.1.3308.1282
De Mel, S., McKenzie, D., & Woodruff, C. (2008). Returns to capital in microenterprises: evidence from a field experiment. *Quarterly Journal of Economics, 123*(4), 1329-1372. (http://www.jstor.org/stable/40506211 [accessed 12 October 2016])
Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and Statistics, 84* (1), 151-161. doi:doi:10.1162/003465302317331982
DiPrete, T. A., & Gangl, M. (2004). Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments. *Sociological methodology, 34* (1), 271-310. doi:10.1111/j.0081-1750.2004.00154.x
Duenwald, C., Gueorguiev, N., & Schaechter, A. (2007). Too Much of a Good Thing? Credit Booms in Transition Economies: The Cases of Bulgaria, Romania, and Ukraine. In C. Enoch & İ. Ötker-

Robe (Eds.), *Rapid Credit Growth in Central and Eastern Europe: Endless Boom or Early Warning?* (pp. 236-263). London: Palgrave Macmillan UK.

Dunn, E., & Arbuckle, J. G. (2001). Microcredit and microenterprise performance: impact evidence from Peru. *Small Enterprise Development, 12* (4), 22-33. doi:10.3362/0957-1329.2001.043

Duval, A., & Goodwin-Groen, R. (2005). Experienced Consortium Deepens Bulgarian Financial System by Creating ProCredit Bank, a Commercial Bank for Microenterprises. *CGAP Case Studies in Donor Good Practices, 20*. (http://documents.worldbank.org/curated/en/407381468232154822/pdf/764280BRI0CGAP00Box374367B00PUBLIC0.pdf [accessed 12 October 2016])

Duvendack, M., & Palmer-Jones, R. (2012). High noon for microfinance impact evaluations: re-investigating the evidence from Bangladesh. *Journal of Development Studies, 48* (12), 1864-1880. doi:10.1080/00220388.2011.646989

Duvendack, M., Palmer-Jones, R., Copestake, J. G., Hooper, L., Loke, Y., & Rao, N. (2011). What is the evidence of the impact of microfinance on the well-being of poor people? (https://assets.publishing.service.gov.uk/media/57a08aeeed915d622c0009bb/Microfinance2011Duvendackreport.pdf [accessed 12 October 2016])

EU. (2005). The New SME Definition: User Guide and Model Declaration. *Enterprise and Industry Publications*. (http://ec.europa.eu/DocsRoom/documents/10109/attachments/1/translations/en/renditions/native [accessed 12 October 2016])

EU. (2010). Europe 2020: A Strategy for Smart, Sustainable and Inclusive Growth - Communication from the Commission. Publications Office of the European Union. Retrieved from http://ec.europa.eu/eu2020/pdf/COMPLET%20EN%20BARROSO%20%20%20007%20-%20Europe%202020%20-%20EN%20version.pdf [accessed 12 October 2016]

Eurostat. (2009). Business Demography: employment and survival. *Statistics in focus, 70/2009*. (http://ec.europa.eu/eurostat/documents/3433488/5284008/KS-SF-09-070-EN.PDF/d43f5156-61fb-4f3a-8bd7-e6902e48c8de [accessed 15 December 2016])

Fafchamps, M., & Schündeln, M. (2013). Local financial development and firm performance: Evidence from morocco. *Journal of Development Economics*. doi:10.1016/j.jdeveco.2013.01.010

Grimm, M., & Paffhausen, A. L. (2015). Do interventions targeted at micro-entrepreneurs and small and medium-sized firms create jobs? A systematic review of the evidence for low and middle income countries. *Labour Economics, 32*, 67-85. doi:10.1016/j.labeco.2015.01.003

Gubert, F., & Roubaud, F. (2011). The Impact of Micro Finance Loans on Small Informal Enterprises in Madagascar. A Panel Data Analysis. *Social Protection Series. The World Bank, No. 77931*. (http://siteresources.worldbank.org/INTLM/Resources/390041-1212776476091/5078455-1398787692813/9552655-1398787856039/Gubert-Roubaud-Impact_of_Microfinance_Loans_on_Small_Informal_Enterprises-Madagascar.pdf [accessed 12 October 2016])

Guirkinger, C. (2008). Understanding the coexistence of formal and informal credit markets in Piura, Peru. *World Development, 36* (8), 1436-1452. doi:10.1016/j.worlddev.2007.07.002

Hartarska, V., Caudill, S. B., & Gropper, D. M. (2006). The cost structure of microfinance institutions in eastern Europe and central Asia. *William Davidson Institute Working Paper No. 809*. doi:10.2139/ssrn.905911

Hartarska, V., & Nadolnyak, D. (2008). An Impact Analysis of Microfinance in Bosnia and Herzegovina. *World Development, 36* (12), 2605-2619. doi:10.1016/j.worlddev.2008.01.015

Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *Review of Economic Studies, 64* (4), 605-654. doi:10.2307/2971733

Ichino, A., Mealli, F., & Nannicini, T. (2008). From temporary help jobs to permanent employment: What can we learn from matching estimators and their sensitivity? *Journal of Applied Econometrics, 23* (3), 305-327. doi:10.1002/jae.998

Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature, 47* (1), 5-86. doi:10.1257/jel.47.1.5

Karlan, D., & Morduch, J. (2010). Access to Finance. In D. Rodrik & M. Rosenzweig (Eds.), *Handbook of Development Economics* (Vol. 5, pp. 4703 – 4784.). Kidlington, UK: Elsevier.

Karlan, D., & Zinman, J. (2011). Microcredit in Theory and Practice: Using Randomized Credit Scoring for Impact Evaluation. *Science, 332* (6035), 1278-1284. doi:10.1126/science.1200138

Ledgerwood, J. (2013). *The New Microfinance Handbook: a Financial Market System Perspective*. Washington, D.C.: The Worldbank.

Leuven, E., & Sianesi, B. (2003). PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing, Version 4.0.10. (https://ideas.repec.org/c/boc/bocode/s432001.html [accessed 20 January 2014])

Madestam, A. (2014). Informal finance: A theory of moneylenders. *Journal of Development Economics, 107*, 157-174. doi:10.1016/j.jdeveco.2013.11.001

Maes, J. P., & Reed, L. R. (2012). State of the microcredit summit campaign report 2012. (http://www.microcreditsummit.org/resource/46/state-of-the-microcredit-summit.html [accessed 12 October 2016])

McKenzie, D., & Woodruff, C. (2008). Experimental evidence on returns to capital and access to finance in Mexico. *The World Bank Economic Review, 22* (3), 457-482. doi:10.1093/wber/lhn017

Ministry of Economy. (2004). Annual report on the condition and development of SMEs in Bulgaria 2004. (http://old.mee.government.bg/eng/ind/econ/docs.html?id=162491 [accessed 12 October 2016])

MIX-Market. (2013). Convergences Microfinance Barometer 2013. (http://www.convergences.org/en/bibliotheque/microfinance-barometer-2013/ [accessed April 13 2016])

Montgomery, H. (2005). Meeting the Double Bottom Line: The Impact of Khushali Bank's Microfinance Program in Pakistan. *Asian Development Bank Policy Papers, 8*. doi:10.2139/ssrn.1337277

Nannicini, T. (2007). Simulation-based sensitivity analysis for matching estimators. *Stata Journal, 7*(3), 334. (http://www.stata-journal.com/article.html?article=st0130 [accessed 12 October 2016])

Pande, R., Cole, S., Sivasankaran, A., Bastian, G. G., & Durlacher, K. (2012). Does poor people's access to formal banking services raise their incomes? (http://eppi.ioe.ac.uk/cms/Portals/0/PDF%20reviews%20and%20summaries/Banking-access-2012Pande.pdf?ver=2015-06-05-140218-057 [accessed 12 October 2016])

Pissarides, F., Singer, M., & Svejnar, J. (2003). Objectives and constraints of entrepreneurs: evidence from small and medium size enterprises in Russia and Bulgaria. *Journal of Comparative Economics, 31 (2003)* 503–531. doi:10.1016/S0147-5967(03)00054-4

Pitt, M., & Khandker, S. (1998). The Impact of Group-Based Credit Programs on Poor Households in Bangladesh: Does the Gender of Participants Matter? *Journal of Political Economy, 106*(5), 958-996. (http://www.jstor.org/stable/10.1086/250037 [accessed 12 October 2016])

Roodman, D., & Morduch, J. (2014). The Impact of Microcredit on the Poor in Bangladesh: Revisiting the Evidence. *Journal of Development Studies, 50* (4), 583-604. doi:10.1080/00220388.2013.858122

Rosenbaum, P. R. (2005). Observational study. In B. Everitt & D. Howell (Eds.), *Encyclopedia of Statistics in Behavioral Science* (Vol. 3, pp. 1451-1462). Chichester: John Wiley & Sons.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70* (1), 41-55. doi:10.1093/biomet/70.1.41

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *American Statistician, 39*(1), 33-38. (http://amstat.tandfonline.com/doi/abs/10.1080/00031305.1985.10479383 [accessed 12 October 2016])

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology, 66* (5), 688. doi:10.1037/h0037350

Rubin, D. B., & Thomas, N. (1996). Matching using estimated propensity scores: relating theory to practice. *Biometrics*, 249-264. doi:10.2307/2533160

Setboonsarng, S., & Parpiev, Z. (2008). Microfinance and the Millennium Development Goals in Pakistan: Impact assessment using propensity score matching. *ADB Institute Discussion Papers, 104*. (https://www.econstor.eu/handle/10419/53481 [accessed 12 October 2016])

Simeonova-Ganeva, R., Vladimirov, Z., Boeva, M., Panayotova, N., Ganev, K., & Peneva, R. (2011). Analysis of the Situation and Factors for Development of SMEs in Bulgaria: SMEs in the Crisis Context doi:10.2139/ssrn.2290717 (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2290717 [accessed 11 October, 2016])

Smith, J. A., & Todd, P. E. (2005). Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics, 125* (1), 305-353. doi:10.1016/j.jeconom.2004.04.011

Tedeschi, G. A., & Karlan, D. (2010). Cross-Sectional Impact Analysis: Bias from Dropouts. *Perspectives on Global Development and Technology, 9* (3), 270-291. doi:10.1163/156914910X499714

van Rooyen, C., Stewart, R., & de Wet, T. (2012). The Impact of Microfinance in Sub-Saharan Africa: A Systematic Review of the Evidence. *World Development, 40* (11), 2249–2262. doi:10.1016/j.worlddev.2012.03.012

Wagner, C., & Winkler, A. (2013). The vulnerability of microfinance to financial turmoil–evidence from the global financial crisis. *World Development, 51*, 71-90. doi:10.1016/j.worlddev.2013.05.008

Williams, C. C. (2014). Out of the shadows: a classification of economies by the size and character of their informal sector. *Work, employment and society, 28* (5), 735 –753.

Worldbank. (2006). Bulgaria - BEEPS at a glance. Washington, DC: World Bank. Retrieved from http://documents.worldbank.org/curated/en/592901468249644234/Bulgaria-BEEPS-at-a-glance [accessed]

Wydick, B. (2016). Microfinance on the margin: why recent impact studies may understate average treatment effects. *Journal of Development Effectiveness, 8* (2), 257-265. doi:bu10.1080/19439342.2015.1121512

# 10 TABLES AND FIGURES

**TABLE 1**

Overview of impact assessments on microfinance and employment

| Authors | Country | Program | | | Research Design [a] | Effects | |
|---|---|---|---|---|---|---|---|
| | | Loan | Borrower | MFI | | self-employment[b] | wage-employment |
| Abou-Ali et al. (2010) | Egypt | individual | households | NGO/ bank | PSM | + | [c] |
| Angelucci et al. (2015) | Mexico | group | women | bank | RCT (expansion of program into new areas using randomized loan promotion) | 0 | |
| Attanasio et al. (2015) | Mongolia | individual / group | women | bank | RCT (expansion of program into new areas) | 0/+ | |
| Augsburg et al. (2015) | Bosnia & Herzegovina | individual | private individuals | NGO | RCT (randomized approval for marginally creditworthy loan applicants) | +/0 | |
| Banerjee, Duflo, et al. (2015) | India | group | women | NGO | RCT (expansion of program into new areas) | + | 0/- |
| Bruhn and Love (2014) | Mexico | individual | private individuals | bank | DiD by exploiting cross-time and cross-municipality variation in simultaneous rollout of 800 branches in new areas | +/0 | [c] |
| Chemin (2008) | Bangladesh | group | households | bank | PSM (re-estimation of Pitt and Khandker, 1998) | +/0 | |
| Coleman (1999) and Coleman (2006) | Thailand | group | women | NGO | DiD and 'pipeline design' (treated are clients in program villages; controls are future clients in new program villages surveyed between loan application and disbursement and randomly selected non-clients in both type of villages) | 0 | |
| Crépon et al. (2015) | Morocco | group | private individuals | NGO | RCT (expansion of program into new areas) | 0 | |
| Dunn and Arbuckle (2001) | Peru | individual | firms | bank | DiD and 'pipeline design' (treated are clients in program villages; controls are future clients in new program villages surveyed between loan application and disbursement and randomly selected non-clients in both type of villages) | + | + |
| Duvendack and Palmer-Jones (2012) | Bangladesh | group | households | bank | PSM (replication of Chemin, 2008) | 0/0 | |
| Gubert and Roubaud (2011) | Madagascar | individual | firms | NGO | PSM combined with DiD | - | 0 |
| Karlan and Zinman (2011) | Philippines | individual | private individuals | bank | RCT (randomize approval for marginally creditworthy loan applicants) | - | - |
| Montgomery (2005) | Pakistan | group | private individuals | bank | DiD and 'pipeline design' (treated are clients in program villages; controls are future clients in new program villages surveyed between loan application and disbursement and randomly selected non-clients in both type of villages) | +/0 | |
| Pitt and Khandker (1998) | Bangladesh | group | households | bank | Regression on discontinuity based on eligibility rule (land ownership) by comparing individuals just above and below eligibility | -/+ | |

**TABLE 1** *(continued)*

| Authors | Country | Program | | | Research Design[a] | Effects | |
|---|---|---|---|---|---|---|---|
| | | Loan | Borrower | MFI | | self-employment[b] | wage-employment |
| Roodman and Morduch (2014) | Bangladesh | group | households | bank | Replication of Pitt and Khandker (1998) by dropping outliers, using a robust linear estimator and a new program for estimation of mixed process maximum likelihood models | -/0 | |
| Setboonsarng and Parpiev (2008) | Pakistan | group | private individuals | bank | PSM (re-estimation of data used in Montgomery, 2005) | 0 | |
| Tedeschi and Karlan (2010) | Peru | individual | firms | bank | DiD and 'pipeline design' (treated are clients in program villages; controls are future clients in new program villages surveyed between loan application and disbursement and randomly selected non-clients in both type of villages; re-estimation of data in Dunn and Arbuckle, 2001) | + | + |

[a] PSM = Propensity score matching, RCT= Randomized Control Trial, DiD = Difference in Differences, IV=Instrumental Variables

[b] Self-employment includes business ownership and business creation: negative, + positive, 0 no or insignificant effect)

[c] Wage-employment effects are measured on the level of aggregate employment rates for program areas. No information available on wage-employment effects in borrowers firms.

**TABLE 2**

Comparison of effects on wage-employment in microfinance impact assessments

| Authors | Country | Outcome variable | Impact on wage-employment | Loan size (market exchange rate) | Loan size (PPP adjusted)[a] | Evaluation period | Firm size[b] |
|---|---|---|---|---|---|---|---|
| Banerjee, Duflo, et al. (2015) | India | Number of employees of new and existing firms | 0: insignificant (-0.05 employees or -20%) for existing businesses<br>-0.2 employees (or -20%) for new businesses | USD 200 | USD 777 | 18 months | 0.3 |
| Karlan and Zinman (2011) | Philippines | Number of paid employees in all household businesses | - 0.27 employees (-39%) | USD 220 | USD 726 | 11-22 months | 0.7 |
| Dunn and Arbuckle (2001) | Peru | Total days worked per month for non-household members in top 3 client firms | +2.49 days (0.13 employees or 4%) | USD 586 | USD 1,368 | 24 months | 2.2 |
| Tedeschi and Karlan (2010) | Peru | | +1.18 days (0.06 employees or 3%) | | | 24 months | |
| Gubert and Roubaud (2011) | Madagascar | Number of workers per firm | 0: insignificant (around +9%) | USD 443 | USD 1,616 | 36 months | 2.3 |

[a] Purchasing power parity (PPP) adjusted exchange rates are calculated by using the conversion factor to market exchange rate for a specific country in 2006 as indicated by the Worldbank: http://data.worldbank.org/indicator/PA.NUS.PPPC.RF (Accessed 13 March 2017)

b Number of employees before start of the program

**TABLE 3**

Exit of sample firms 2004-2009

| | Participants | Non-Participants |
|---|---|---|
| | Sample size | |
| 2003 | 974 | 60,032 |
| | | |
| Year of exit | Exit rate (Percentage) | |
| 2004 | 0.2% | 4.6% |
| 2005 | 1.3% | 4.5% |
| 2006 | 1.4% | 4.6% |
| 2007 | 5.5% | 6.1% |
| 2008 | 6.4% | 5.3% |
| 2009 | 4.8% | 6.1% |

Note: Exit indicates that no data on variables employees, turnover or profit exists for respective and any following years.

**TABLE 4**

Descriptive Statistics

| | **Participants** | **Non-Participants** |
|---|---|---|
| **Firm size category** | | |
| Micro (1-9 employees) | 78.1 | 71.7 |
| Small (10-49 employees) | 20.0 | 22.6 |
| Medium (50-249 employees) | 1.8 | 5.7 |
| **Employees (number)** | 7.492 (13.156) | 12.737 (26.009) |
| **Sector (NACE section)**[a] | | |
| Agriculture, hunting, forestry | 2.7 | 5.8 |
| Fishing | 0.0 | 0.1 |
| Mining and quarrying | 0.0 | 0.2 |
| Manufacturing | 16.8 | 16.6 |
| Electricity, gas and water supply | 0.1 | 0.2 |
| Construction | 3.1 | 7.3 |
| Wholesale and retail trade; repair of motor vehicles | 49.2 | 42.2 |
| Hotels and restaurants | 5.2 | 3.5 |
| Transportation, storage, communications | 10.4 | 5.8 |
| Financial intermediation | 0.5 | 1.8 |
| Real estate and renting activities | 7.1 | 11.3 |
| Public administration and defense | 0.0 | 0.0 |
| Education | 0.3 | 0.3 |
| Health and social work | 2.2 | 2.5 |
| Other services activities | 2.5 | 2.4 |
| **Location (by Economic Development)**[b] | | |
| Very Good | 12.8 | 9.4 |
| Good | 30.7 | 41.3 |
| Average | 25.2 | 18.8 |
| Unsatisfactory | 19.3 | 17.0 |
| Weak | 12.0 | 13.6 |
| **Ownership type**[c] | | |
| General partnership | 0.4 | 5.2 |
| Limited partnership | 44.0 | 34.3 |
| Joint-stock company | 30.9 | 28.1 |
| Limited liability company | 23.5 | 29.1 |
| State enterprise | 1.1 | 3.2 |
| **Gender of owner**[d] | | |
| Male | 72.9 | 61.7 |
| Female | 26.9 | 20.6 |
| No information | 0.2 | 17.8 |
| Sample size | 974 | 60,032 |
| **Change in Employees 2003-2006 (number)**[e] | 2.726 (12.683) | -.181 (18.592) |

Note: Numbers are percentages and measured at the end of 2003 unless otherwise stated; Standard deviation in parentheses.

a Sector is defined according to NACE Rev. 1.1. NACE is the standard classification system of economic activities since 1970 in the European Union. For each observation, we have information on the level of four-digit numerical codes (classes). We aggregated this information into two-digit numerical codes (divisions), in order to keep enough non-participants for the matching with our participants. Here, we present the first level of NACE Rev. 1.1 consisting of so-called sections.

b Bulgaria is divided into 28 districts, 264 municipalities and about 4,360 places with different 4-digit postal codes. We have information on the level of postal codes for each observation. Furthermore we clustered observations by district into five categories of economic development. We developed a composite index based on GDP per capita, unemployment and employment rate of the population aged 15+, number of non-financial companies per 1,000 people, expenditures for acquisition of fixed tangible assets per 1,000 people and annual income per household member. The clustering is then performed by calculating mean ranks of economic development for each district in the period 2003 to 2006 and dividing districts into five groups. All data originates from the National Statistical Institute (NSI) of Bulgaria.

c According to the Bulgarian commercial law the following legal forms exist for commercial companies: General partnership, Limited Partnership, Partnership Limited by shares (regrouped with Limited Partnership due to small number of observations), Limited Liability Company and Joint Stock Company. Legal forms differ in their degree of liability for the companies' obligations by their members, in requirements for minimum capital and in the according level of complexity for management and reporting. In addition, there exist state-owned and municipal enterprises into which we included the small number of associations and co-operations.

d The information on gender of owners is based on first and last names of firm owners in the Amadeus database. Owners are defined as holding at least 25 percent of a company. In all but 41 cases (where the first indicated owner was used), the information in Amadeus comprised only one person as owner.

e The difference in difference estimator of the unmatched sample therefore amounts to 2.907 and a standard error of 0.598

**TABLE 5**

Probit estimation of participation probability

| | **Coefficient** | **Standard Error** |
|---|---|---|
| Number of Employees | -0.008* | 0.004 |
| Firm size categories [a] | $chi^2(11)$ = 17.10* | |
| Sector by NACE divisions [a] | $chi^2(43)$ = 280.58*** | |
| Location by districts [a] | $chi^2$ (24) = 209.94*** | |
| Ownership type (Ref.: General partnership) | | |
| Limited partnership | -0.719** | 0.338 |
| Joint-stock company | -0.575* | 0.338 |
| Limited liability company | -0.673** | 0.336 |
| State enterprise | -1.070*** | 0.356 |
| Gender of owner (Ref.: no information) | | |
| Male | 1.853*** | 0.312 |
| Female | 1.876*** | 0.313 |
| Constant | -3.159*** | 0.417 |
| Number of observations | | |
| Participants | 974 | |
| Non-participants | 57,691 | |
| LR $chi^2(84)$ | 944.26*** | |
| Prob > $chi^2$ | 0.000 | |
| Pseudo $R^2$ | 0.095 | |
| Log likelihood | -4485.3754 | |

Notes: The table reports probit estimations. For definitions of the variables see notes in Table 3.
[a] Depicted is the joint significance for dummy variables based on Wald tests.

* 10%, ** 5%, *** 1% significance level

**TABLE 6**

Firms off and on common support region

| **Treatment assignment** | **Off support** | **On support** | **Total** |
|---|---|---|---|
| Participants | 4 | 970 | 974 |
| Non-participants | 0 | 57,691 | 57,691 |
| Total | 4 | 58,661 | 58,665 |

Note: The propensity score distributions are based on probit estimations in Table 5. Common support is imposed by dropping treated observations whose propensity score is higher than the maximum or less than the minimum pscore of the untreated.

**TABLE 7**

Balancing Tests (after caliper matching)

| **T-test of equal means** | | Mean | | % reduction of bias | t-test | |
|---|---|---|---|---|---|---|
| Variable | | Participants | Non-participants | | t | p>t |
| Employees (number) | Unmatched | 7.492 | 12.496 | | -6.09 | 0.000 |
| | Matched | 7.478 | 7.087 | 92.2 | 0.64 | 0.525 |
| Firm size categories [a] | Unmatched | | | | 7 (12) | |
| | Matched | | | | 0 (12) | |
| Sector by NACE division [a] | Unmatched | | | | 16 (44) | |
| | Matched | | | | 0 (44) | |
| Location by districts [a] | Unmatched | | | | 18 (24) | |
| | Matched | | | | 0 (24) | |
| Ownership type (Ref.: General partnership) | | | | | | |
| Limited partnership | Unmatched | 0.441 | 0.344 | | 6.31 | 0.000 |
| | Matched | 0.442 | 0.463 | 78.7 | -0.91 | 0.362 |
| Joint-stock company | Unmatched | 0.309 | 0.286 | | 1.61 | 0.107 |
| | Matched | 0.307 | 0.300 | 69.3 | 0.35 | 0.730 |
| Limited liability company | Unmatched | 0.235 | 0.291 | | -3.84 | 0.000 |
| | Matched | 0.236 | 0.224 | 78.0 | 0.65 | 0.518 |
| State enterprise | Unmatched | 0.011 | 0.031 | | -3.51 | 0.000 |
| | Matched | 0.011 | 0.013 | 89.4 | -0.41 | 0.681 |
| Gender of owner (Ref.: no information) | | | | | | |
| Male | Unmatched | 0.729 | 0.621 | | 6.88 | 0.000 |
| | Matched | 0.729 | 0.745 | 84.7 | -0.82 | 0.410 |
| Female | Unmatched | 0.269 | 0.208 | | 4.62 | 0.000 |
| | Matched | 0.269 | 0.253 | 72.8 | 0.83 | 0.408 |
| | | | | | | |
| **Pseudo-$R^2$** | Unmatched | 0.095 | | | | |
| | Matched | 0.012 | | | | |
| | | | | | | |
| **LR $chi^2$** | Unmatched | 944.26 | p-value | 0.000 | | |
| | Matched | 32.84 | p-value | 1.000 | | |
| | | | | | | |
| **Mean standardized bias** | Unmatched | 7.8 | | | | |
| | Matched | 2.1 | | | | |

Notes: For definitions of the variables see notes in Table 3.
[a] Depicted is the number of dummy variables with mean values that differ significantly at the 10 percent level between participants and non-participants. The total number of dummy variables is indicated in parentheses.

**TABLE 8**

Main treatment effects with caliper matching

| Sample | Number of Employees 2003 | Number of Employees 2006 | Difference | **ATT (Δ Number of employees 2006-2003)** | S.E. |
|---|---|---|---|---|---|
| Participants (Matched) | 7.478 | 10.200 | 2.722 | **2.480*** ** | 0.564 |
| Non-Participants (Matched) | 7.087 | 7.328 | 0.241 | | |
| Non-Participants (Unmatched) | 12.497 | 12.451 | -0.045 | | |

Notes: Estimates are based on a caliper matching estimator with a bandwidth of 0.05.

* 10%, ** 5%, *** 1% significance level

**TABLE 9**

Main treatment effects with different matching estimators

| Matching Estimator | **ATT**[a] | S.E. |
|---|---|---|
| Caliper Matching (bandwidth 0.05) | 2.481*** | 0.564 |
| Caliper Matching (bandwidth 0.01) | 2.480*** | 0.564 |
| Caliper Matching (bandwidth 0.001) | 2.504*** | 0.567 |
| Nearest Neighbor (1 neighbor, with replacement) | 2.480*** | 0.564 |
| Nearest Neighbor (1 neighbor, without replacement) | 2.280*** | 0.514 |
| Radius Matching (bandwidth 0.05) | 2.592*** | 0.417 |
| Kernel Matching (bandwidth 0.05, Gaussian weighting function) | 2.667*** | 0.415 |
| Kernel Matching (bandwidth 0.05, Epanechnikov weighting function) | 2.485*** | 0.418 |

Note: Depicted are average treatment effects on the treated as the difference in outcome variables between participants and non-participants. For radius and kernel matching estimates the standard error is based on bootstrapping with 100 replications.

[a] Absolute change defined as Δ N employees 2006 - N employees 2003.

* 10%, ** 5%, *** 1% significance level

**TABLE 10**

Dynamic treatment effects

| Year | ATT[a] | S.E. | Number of employees participants | Number of employees matched non-participants | Number of employees full sample of non-participants | Number of participants | Number of non-participants |
|---|---|---|---|---|---|---|---|
| 2003 (before treatment) | 0 | 0 | 7.478 | 7.087 | 12.497 | 970 | 57,691 |
| 2004 (year of treatment) | 2.598*** | 0.292 | 10.286 | 7.296 | 12.333 | 970 | 57,691 |
| 2005 (1 year after treatment) | 1.289*** | 0.404 | 8.790 | 7.109 | 12.207 | 970 | 57,691 |
| 2006 (2 years after treatment) | 2.480*** | 0.562 | 10.200 | 7.328 | 12.451 | 970 | 57,691 |
| 2007 (3 years after treatment) | 2.694*** | 0.717 | 11.052 | 7.702 | 13.313 | 904 | 54,882 |
| 2008 (4 years after treatment) | 3.803*** | 1.054 | 11.378 | 8.052 | 12.309 | 894 | 54,941 |
| 2009 (5 years after treatment) | 2.040*** | 0.767 | 9.790 | 7.827 | 10.885 | 951 | 56,694 |
| 2010 (6 years after treatment) | 2.110*** | 0.711 | 8.350 | 5.973 | 8.966 | 958 | 57,603 |

Notes: Estimates are based on a caliper matching estimator with a bandwidth of 0.05.

[a] Absolute change defined as Δ N employees year Y - N employees 2003.

* 10%, ** 5%, *** 1% significance level

**TABLE 11**

Heterogeneous treatment effects

| | **ATT[a]** | S.E. | **ATT[b] (in percent)** | Baseline number employees (participants 2003) | Number participants | Number non-participants |
|---|---|---|---|---|---|---|
| Base Effect | 2.480*** | 0.562 | 0.332 | 7.478 | 970 | 57,691 |
| Loan size [c] | | | | | | |
| small | 1.058*** | 0.337 | 0.276 | 3.830 | 312 | 40,893 |
| medium | 1.782** | 0.782 | 0.300 | 5.954 | 325 | 48,753 |
| large | 3.802*** | 1.451 | 0.250 | 12.322 | 329 | 52,719 |
| < EUR 500 | 0.594 | 0.445 | 0.268 | 2.219 | 32 | 17,822 |
| EUR 500-1,000 | 0.670* | 0.365 | 0.239 | 2.800 | 100 | 29,862 |
| Firm size | | | | | | |
| micro (1-9 employees) | 1.590*** | 0.302 | 0.504 | 3.156 | 759 | 41,617 |
| small (10-49 employees) | 4.166** | 1.693 | 0.243 | 17.160 | 187 | 11,395 |
| medium (50-249 employees) | -25.471 | 20.448 | -0.305 | 83.529 | 17 | 560 |
| Loan purpose | | | | | | |
| working capital | 2.301*** | 0.881 | 0.283 | 8.129 | 412 | 53,506 |
| investment | 1.799*** | 0.694 | 0.257 | 6.991 | 557 | 54,602 |
| Location (non-participants from districts without program) | 2.396*** | 0.854 | 0.328 | 7.314 | 934 | 6,776 |

Notes: Estimates are based on a caliper matching estimator with a bandwidth of 0.05.

[a] Absolute change defined as Δ N employees 2006 - N employees 2003.

[b] Relative change defined as $\frac{\Delta\text{ N employees 2006} - \text{N employees 2003}}{\text{N employees participants 2003}}$

[c] Loan amount categories are defined as small < 2,000 EUR, medium= 2,0000 -6,000 EUR, and large > 6,000 EUR.

* 10%, ** 5%, *** 1% significance level

**TABLE 12**

Sensitivity analysis for alternative specifications of propensity score estimation

| | **ATT[a]** | S.E. | Number participants | Number non-participants |
|---|---|---|---|---|
| **Base Effect** | 2.480*** | 0.564 | 970 | 57,691 |
| **Additional years prior treatment** | | | | |
| Employees 2002, 2001, 2000 | 2.768* | 1.440 | 314 | 30,186 |
| Δ N employees 2003 – N employees 2000 | 3.057** | 1.201 | 314 | 30,186 |
| **Additional variables at baseline** | | | | |
| Sales, profit, productivity 2003 | 2.780*** | 0.944 | 545 | 22,905 |
| **Exact matching on district** | 1.691*** | 0.621 | 971 | 59,574 |

Notes: Estimates are based on a caliper matching estimator with a bandwidth of 0.05.

[a] Absolute change defined as Δ N employees 2006 - N employees 2003.

* 10%, ** 5%, *** 1% significance level

**TABLE 13**

OLS estimates

| | Unmatched / without control variables | | Matched / with control variables | |
|---|---|---|---|---|
| PSM (ATT[a] / S.E.) | 2.770*** | 0.592 | 2.480*** | 0.564 |
| OLS on full sample (61,006 obs) | 2.907*** | 0.598 | 1.856*** | 0.586 |
| OLS on matched sample (58,665 obs) | 2.770*** | 0.592 | 1.887*** | 0.584 |

Notes: Estimates are based on a caliper matching estimator with a bandwidth of 0.05. Standard errors are in parentheses.

[a] Absolute change defined as Δ N employees 2006 - N employees 2003.

* 10%, ** 5%, *** 1% significance level

**TABLE 14**

Exogeneity test results

| Outcome variable: Change in employees 2006 | No controls | With controls |
|---|---|---|
| Treatment (1= zero loan) | -1.339 (3.004) | -0.436 (2.926) |
| Loan amount categories | yes | yes |
| Number of Employees | | yes |
| Firm size categories | | yes |
| Sector by NACE divisions | | yes |
| Location by districts | | yes |
| Ownership type | | yes |
| Gender of owner | | yes |
| Constant | 1.158 (3.003) | 6.003 (3.310) |
| Observations | 61,006 | 61,006 |
| R-squared | 0.000 | 0.057 |

**TABLE 15**

Sensitivity analysis with bounding

| Outcome variable: Change in Employees 2006 | | |
|---|---|---|
| | Significance levels p-critical (Wilcoxon signed rank test) | |
| Γ | Lower bound | Upper bound |
| 1.00 | 0.000 | 0.000 |
| 1.25 | 0.000 | 0.000 |
| 1.50 | 0.000 | 4.2e-12 |
| 1.75 | 0.000 | 4.4e-07 |
| 2.00 | 0.000 | 0.0004 |
| 2.25 | 0.000 | 0.027 |
| 2.30 | 0.000 | 0.049 |
| **2.35** | **0.000** | **0.081** |
| 2.40 | 0.000 | 0.126 |
| 2.50 | 0.000 | 0.257 |
| 2.75 | 0.000 | 0.690 |
| 3.00 | 0.000 | 0.934 |

Note: Reported results are achieved by using rbounds.ado from DiPrete and Gangl (2004). Results are based on a caliper matching estimator with a bandwidth of 0.05.

**TABLE 16**

Sensitivity analysis with simulation analysis

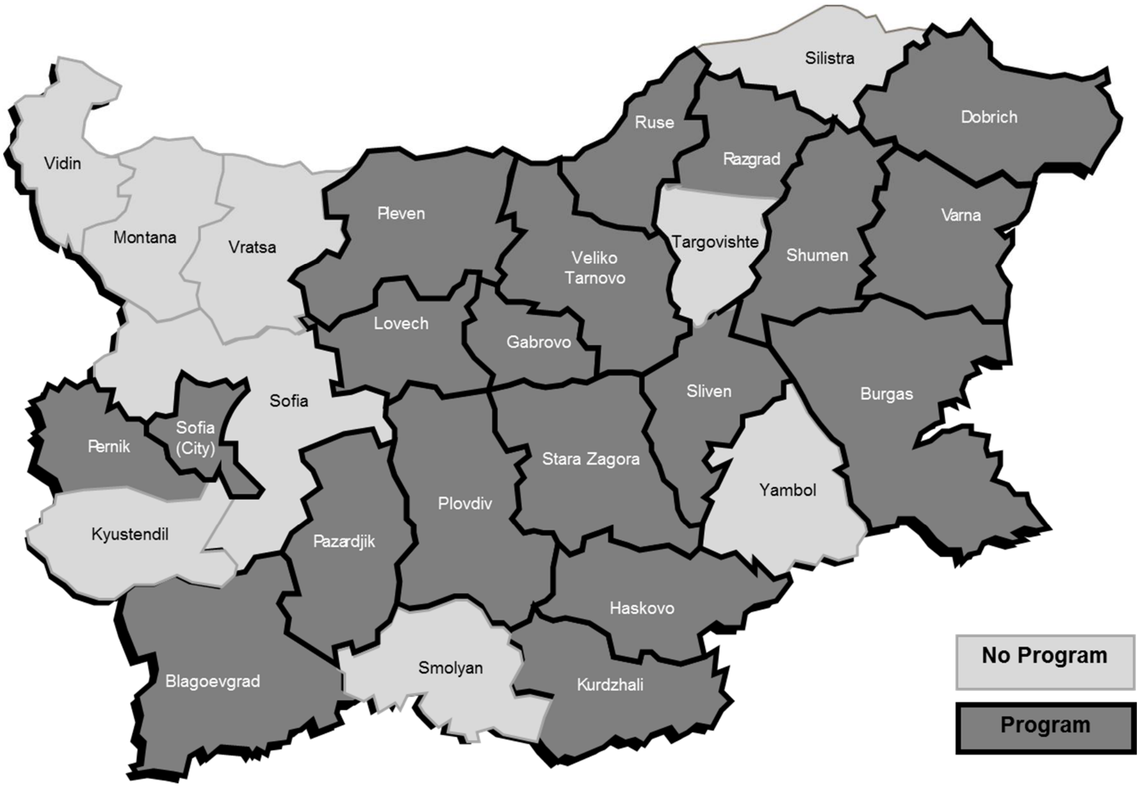
| Confounder | Influence of unobserved confounder on | | **ATT**[a] | S.E. |
|---|---|---|---|---|
| | Outcome | Selection | | |
| No unobserved heterogeneity | 0.00 | 0.00 | 2.361 | 0.427 |
| Confounder with an influence like (see Table 4) | | | | |
| Firm size category (5-9 employees) | 0.714 | 1.017 | 1.948 | 0.553 |
| Ownership type (Joint Stock Company) | 1.307 | 1.111 | 1.873 | 0.570 |
| Gender of owner (female) | 1.187 | 1.677 | 1.887 | 0.556 |

Note: Reported results are achieved by using sensatt.ado from Nannicini (2007) and are related to the continuous outcome variable absolute change in employees 2006. The potential confounder is simulated on the basis of a binary transformation of the outcome: Y=1 if the outcome is above the mean. Results are based on 100 iterations for the simulation of the confounder and the ATT estimation. The underlying matching algorithm is nearest neighbor matching imposing common support based on the Stata command pscore.

[a] Absolute change defined as Δ N employees 2006 - N employees 2003.

**FIGURE 1**

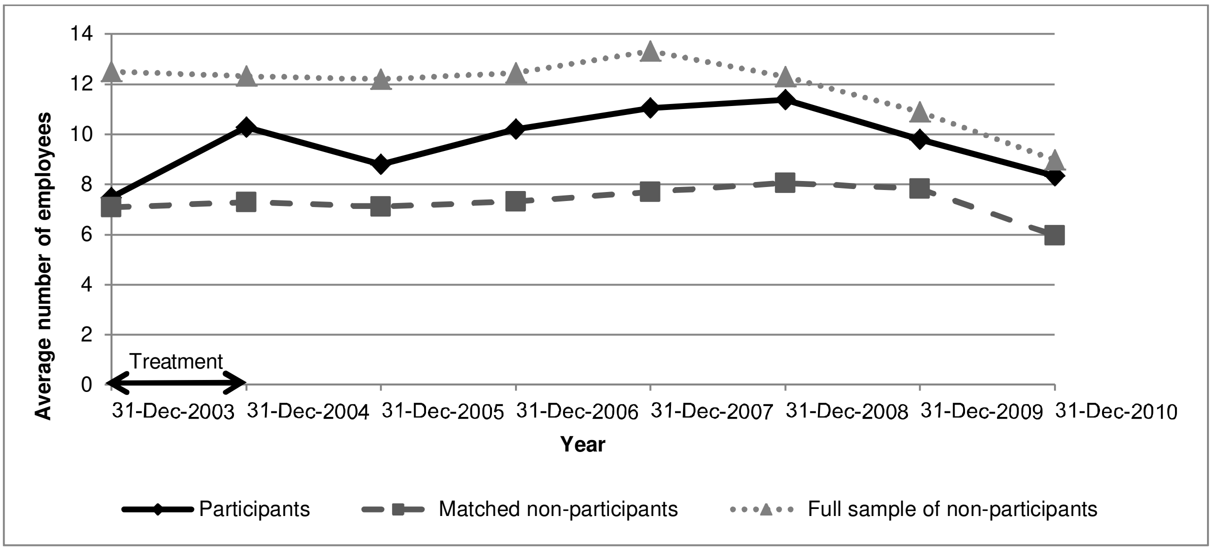
Program Presence on 31 December 2004 (districts with branch)




Source: ProCredit Bank Bulgaria

**FIGURE 2**

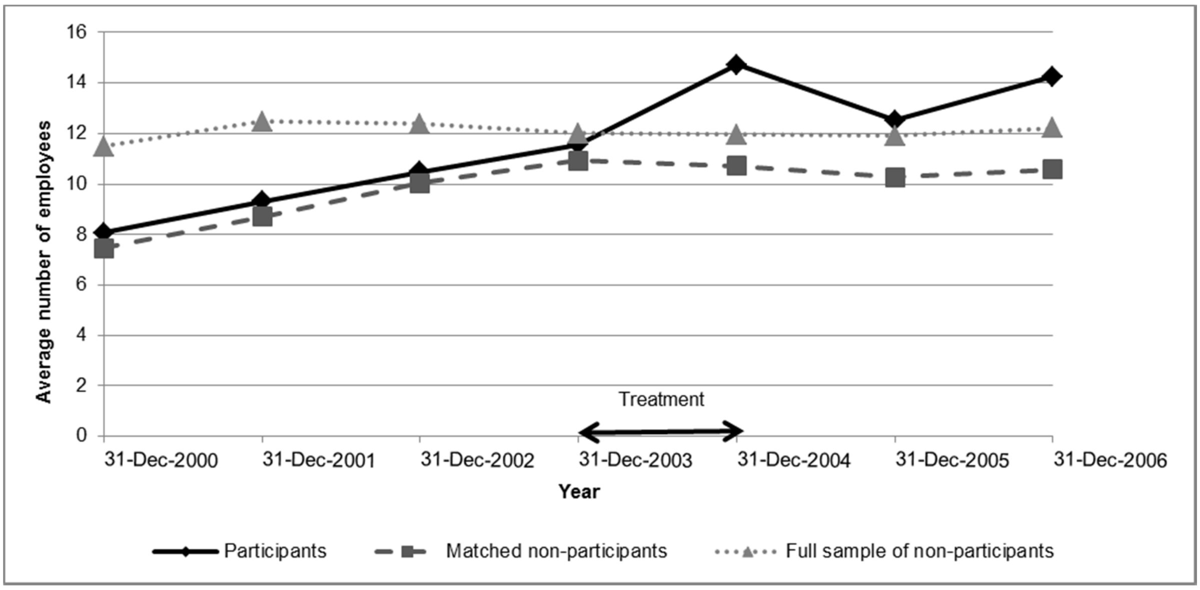
Dynamic development of employees 2003-2010 (participants, matched non-participants, full sample of non-participants)




Notes: Estimates are based on a caliper matching estimator with a bandwidth of 0.05.

**FIGURE 3**

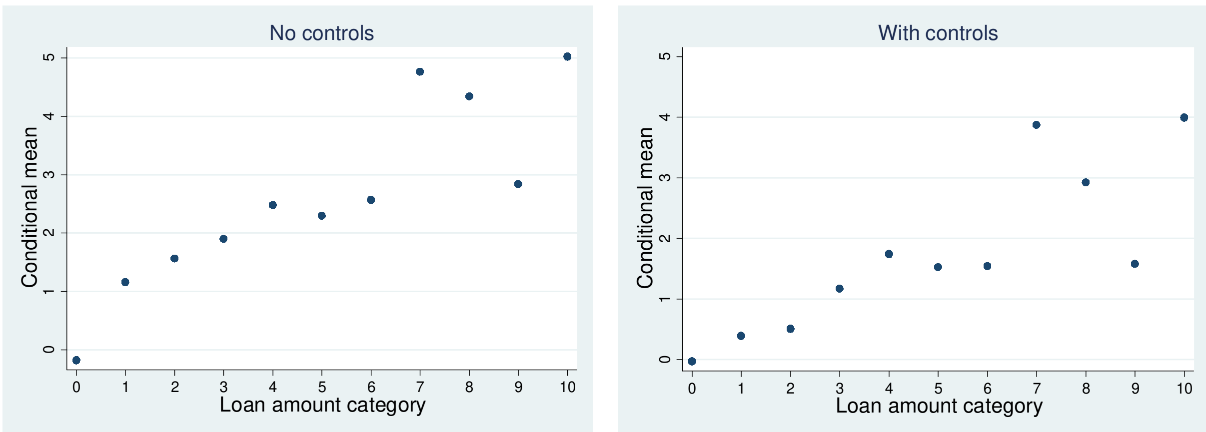
Dynamic development of employees 2000-2006 for sub-samples




Notes: Estimates are based on a caliper matching estimator with a bandwidth of 0.05. The specification for the estimation of the propensity score is extended by Δ N employees 2003 – N employees 2000.

**FIGURE 4**

Exogeneity test results




Note: In the left plot the vertical axis shows the mean change in employees per given loan amount, i.e. horizontal axis variable. The right plot shows on the vertical axis the mean residuals from a linear regression with controls per given loan amount Loan amount is aggregated into ten different categories: EUR 0; EUR 1-500; EUR 501-1,500; EUR 1,501-2,500; EUR 2,501-4,000; EUR 4,001-5,000; EUR 5,001-6,000; EUR 6,001-7,000; EUR 7,001-8,000; EUR 8,001-12,500; EUR 12,501-maximum.