

Does crime deter South Africans from self-employment?

Katharina Grabrucker^a and Michael Grimm^{a,b,c} †

^aUniversity of Passau, ^bErasmus University Rotterdam, ^cIZA

Abstract

An often-heard argument is that South Africa's very high crime rate is the main reason for the country's small share of business ownership. Combining a fixed-effects model with an instrumental variable approach, we estimate the effect of crime on self-employment and business performance using a matched data set of census, survey and police data. In contrast to previous studies, which focus on perceived rather than actual crime and often deal with geographically limited areas, we do not find robust evidence that high crime rates have a negative impact on self-employment. Although the impact of crime is statistically significant and negative, it is economically small. Moreover, our results suggest a positive rather than a negative relationship between robbery and burglary and sales and average business profits. These results suggest that crime may not be in general a serious threat for small businesses in low and middle-income countries.

Keywords: Crime • Self-Employment • Microenterprises • South Africa • Informal Sector

JEL classification: D22, J24, J46, K40, L26, O12.

† Katharina Grabrucker, e-mail: katharina.grabrucker@uni-passau.de; Michael Grimm, e-mail: michael.grimm@uni-passau.de, phone: +49 (0)851 5093310, University of Passau, Chair of Development Economics, Innstrasse 29, 94032 Passau, Germany.

1. Introduction

Small and microenterprises, in both the informal and the formal sector, are typically an important source of income in the developing world. Particularly in countries with few formal employment opportunities – which is the case for most countries in Sub-Saharan Africa – informal self-employment absorbs a large share of the otherwise unemployed. South Africa hereby marks an exception: using the expanded definition of unemployment (including people not actively searching for a job),¹ more than one third of the South African labor force was unemployed in 2015. At the same time, the business ownership accounted for only 14% of total employment – compared to on average 50% in other Sub-Saharan African countries.² Self-employment, or business ownership more broadly, is also much more widespread in most parts of Asia and Latin America (International Labour Organization 2015; Statistics South Africa 2016). Understanding the reasons for the small share of self-employment in South Africa is, therefore, essential to curbing the high level of unemployment in that country.³ An often-heard argument in the literature is that South Africa's very high crime rate is the main reason for the small share of business ownership. Indeed, there are numerous examples in the literature that would support this hypothesis. Cullen and Levitt (1999) show for the US that crime can lead to urban flight. Rosenthal and Ross (2010) show, also for the US, that crime is an important determinant in the location decision of business owners.

In this paper we re-visit the impact of crime on business ownership by matching official crime statistics with individual level data from several rounds of the national census and a community survey, as well as from several waves of surveys on employers and the self-employed. Combining a fixed-effects model with an instrumental variable approach, we first analyze the impact of crime on self-employment. Second, we examine the impact of crime on business performance by focusing on sales, profits and investment behavior. Our identification strategy uses the variation of crime rates between and within municipalities or districts over time.

Crime had been neglected for a long time as a potential constraint on self-employment, but has increasingly attracted attention since the publication of the first official crime statistics ('Crime Stats') at the beginning of the 2000s. This report showed that South Africa ranks among the worst countries globally in terms of crime. This finding led to a number of studies that investigated the role of crime on self-employment in South Africa. Cichello et al. (2011), for instance, identified – based on a survey on barriers to self-employment undertaken in the Khayelitsha township in Cape Town – crime as the single most important perceived constraint on starting a business and therefore as a bigger hindrance than lack of access to startup capital or risk of business-failure. Studying the cost of crime for businesses operating in the informal sector in the Diepsloot

¹ Statistics South Africa (2015a) defines the South African labor force as all persons aged between 15 and 64 years. The strict definition of unemployment includes all people who are part of the labor force and capable of working, but – although actively searching for a job or trying to start an own business – do not have work. In the expanded definition of unemployment people who are willing to work, but who are not actively looking for a job – so-called discouraged job seekers – are also included.

² Considering only Sub-Saharan African countries for which up-to-date figures on self-employment rates (between 2010 and 2015) are available, the average self-employment rate for these countries is 50% (International Labour Organization 2015). For instance, in 2013 in Namibia 31.5% and in Uganda 53.7% were self-employed, in 2012 self-employment in Rwanda accounted for 68.9% and in Gambia for 61.7% of the labor force, while it was only 9.9% in Botswana (2010) and 13.7% in South Africa (2015).

³ We use self-employment as a synonym for any type of business ownership, whether formal or informal or with or without hired workers. Although the large majority of this group are owners of informal business with no or very few hired workers.

township in the north of Johannesburg, Bhorat and Naidoo (2015) also found crime to be the biggest perceived obstacle for these businesses. Likewise, McDonald (2008) found in a survey of 466 small businesses in Cape Town, Durban and Johannesburg, that more than 50% of the owners perceived crime as a major constraint on their business. These results are also similar to an earlier study by Gough et al. (2003) on home-based enterprises ('HBEs') in Pretoria. Although only a small share of HBEs had been victims, Gough et al. (2003) found that many HBEs reported to limit their business activity (for instance closing before sunset) due to their fear of crime. They also documented that many of these HBEs reported to regularly face a shortage of inputs, as many suppliers would not be willing to deliver to areas with prevalent crime. These studies are also in line with previous findings by Chandra et al. (2002) on micro, small and medium-sized enterprises ('MSMEs') in Johannesburg. Only the findings provided by Skinner (2005) deviate from the former as she reports that informal business owners in Durban attribute only a moderate relevance to crime as a hindrance to their business.

Although most of these studies found a relatively large, negative impact of crime on self-employment, they have two main limitations. First, the studies are based on surveys with a restricted regional scope and hence it is far from certain that these findings apply to South Africa as a whole. Second, the studies focus on perceptions and self-reported crime as a constraint to self-employment, rather than on the impact of actual crime rates on actual businesses performance. While we cannot ignore the fact that perceptions can have an effect on actions, the extent of the impact of these perceptions on actions, such as starting or expanding a business, is unclear.

In our analysis, we do not find any robust evidence for an impact of crime on self-employment. Although we find a statistically significant, negative relationship between crime and the decision to become self-employed – which is robust to alternative definitions of crime and self-employment, alternative subsets of the data and other robustness checks – the impact is negligibly small. We also do not find robust negative effects on sales, profits and investment; if anything we find slightly positive effects. These results question the hypothesis that South Africa's high crime rate can explain the low share of self-employment. To our knowledge, this is the first study that analyzes this hypothesis in such detail.

Our paper also contributes to the general literature on the economics of crime. One strand of this literature does, as we do, analyze the labor market consequences of crime (see, e.g., Glaeser et al. 2010; Rosenthal and Ross 2010; Lens and Meltzer 2016). Another strand considers the effects of local labor market conditions such as wages, unemployment and, more generally, economic growth on crime (see, e.g., Gould et al. 2002; Machin and Meghir 2004; Islam 2014). Our study also adds to the literature that analyzes the constraints on start-up and expansion of small and micro enterprises (see, e.g., McPherson 1996; Blanchflower and Oswald 1998; Fajnzylber et al. 2006; McKenzie and Woodruff 2006; De Mel et al. 2008; McKenzie 2010; Banerjee and Duflo 2011, ch. 9). Closely related is the study by Benyishay and Pearlman (2014). Benyishay and Pearlman use data on microenterprises and their experience with crime from official victimization surveys and investigate the impact of property crime on expansion and income growth by microenterprises in Mexico. Using the variation of the number of home robberies between and within states (urban areas only) and over time, they found a robust negative correlation between robbery rates and the probability of a microenterprise expanding its business.

The remainder of this paper is structured as follows: In the next section we describe in more detail the context of the study. In section 3 we discuss the possible mechanisms linking crime to self-employment. These mechanisms will guide our empirical analysis. In section 4 we present

the empirical strategy, our database and the exact specifications. In section 5 we show and discuss the results and provide various robustness checks. We conclude in section 6.

2. Unemployment, self-employment and crime in South Africa

2.1. South Africa in an international comparison

In this section, we show that South Africa is indeed special with respect to the share of the workforce that is unemployed and self-employed. Table 1 lists a large set of countries situated in different regions of the developing world and with a GDP per capita very similar to the South African one. The first striking feature about South Africa is the high level of poverty, which in turn is obviously linked to the high level of inequality the country faces. Although the poverty figures are not fully comparable, as they are based on national and not international poverty lines, it is interesting to see that only Mexico has a similarly high rate. South Africa's labor force participation rate is, by international standards, also relatively low at just 57%. This feature, however, is also shared by countries in Northern Africa and probably includes a non-negligible share of discouraged job seekers who do not consider themselves as active anymore. The fourth column shows the unemployment rate which is, of course, not easy to compute in poor and emerging economies, but even considering a reasonable margin of error, the South African rate is higher – in most cases significantly higher – than in all the other countries listed in this table. Yet, despite this high rate, the share of people with informal jobs, i.e. without any working contract and/or social security (in the formal or informal sector, independent or as wagedworkers) in relation to total non-agricultural employment is also very low. This share is reported to be 33%, which is well below neighboring Namibia and most countries in the Asia Pacific region and in Latin America, where in most cases the rate clearly exceeds 50% and often even 60%. This discrepancy does not change if only informal sector employment is considered. At the same time, self-employment (i.e., the share of own account workers and employers in relation to total employment) is comparably low in South Africa. While the self-employment rate in the majority of the considered countries is around 30%, it is only 14% in South Africa. This is only 'topped' by 10% of self-employment in Botswana and 12% in China. Finally, looking at crime rates as measured through homicides per 100,000 of the population, South Africa again stands out with a rate of 31 homicides per 100,000 of the population, with only Brazil and Colombia having similar levels of crime. Namibia and Botswana, which are both reputedly not particularly safe, have a rate which is 50% lower. Even Mexico ranks well below South Africa.

In conclusion, South Africa is not only special within Sub-Saharan Africa; it is also special in a worldwide comparison. Many people tend to argue that the high crime rate is the most important reason for the low level of entrepreneurship, and in particular informal entrepreneurship, despite the high level of unemployment. Perceptions data for different areas in South Africa seems to support this hypothesis (Chandra et al. 2002; Gough et al. 2003; Skinner 2005; McDonald 2008; Cichello et al. 2011; Bhorat and Naidoo 2015).

Table 1

Comparison of selected countries with a similar GDP per capita to South Africa (+/- 33%) – Key financial and labor market indicators, latest available year.

	GDP per capita (intl. \$) ^a	Poverty head-count ratio (%) ^b	Labor force particip. (%) ^c	Unemployment (%) ^d	Self-employment (%) ^e	Informal employment (%) ^f	Informal sector employment (%) ^g	Homicides per 100,000 population (#) ^h
Sub-Saharan Africa								
Botswana	16,099	19.3	79.5	18.2	9.9	-	14.4	15.4
Namibia	9,956	28.7	60.8	18.6	31.5	44.0	-	17.5
South Africa	13,049	53.8	56.7	25.1	13.9	32.7	17.8	30.7
Northern Africa								
Algeria	14,193	-	47.3	9.5	28.9	36.5 ⁱ	-	1.4
Egypt	10,533	25.2	52.8	13.2	27.5	35.6 ⁱ	-	3.4
Libya	15,597	-	55.7	19.2	-	-	-	2.5
Tunisia	11,436	15.5	51.8	13.3	24.1	-	-	3.1
Asia Pacific								
China	13,206	-	77.4	4.7	12.1 ^j	32.6 ^k	21.9 ^k	0.8
Indonesia	10,517	11.3	69.5	6.2	37.6	72.5 ^l	64.8 ^l	0.6
Sri Lanka	11,110	6.7	56.2	4.6	36.3	62.1	50.5	3.2
Thailand	15,735	10.5	77.7	0.9	35.9	42.8	32.2	4.9
Latin America								
Brazil	15,893	7.4	72.5	6.8	24.3	36.8	21.7	26.5
Colombia	13,357	27.8	72.6	10.1	47.3	63.7	55.8	30.7
Costa Rica	14,918	21.7	67.6	8.3	23.4	30.7	28.3	8.4
Dom. Republic	13,262	41.1	69.9	15.0	43.7	51.4	33.7	22.0
Ecuador	11,372	22.5	67.2	4.6	33.0	52.0	36.0	12.4
Mexico	17,315	53.2	65.6	4.9	27.3	53.9	34.6	21.5
Paraguay	8,911	22.6	74.8	4.5	37.1	64.4	34.1	9.7
Peru	11,989	22.7	77.2	4.2	40.2	68.8	47.4	9.6

Source: World Development Indicators (World Bank 2015), Crime and Criminal Justice statistics (UNODC 2015) and Key Indicators of the Labour Market (International Labour Organization 2015) except where indicated otherwise.

^a PPP, current international \$ in 2014. ^b Percentage of population living below the national poverty line in 2014 (exceptions: Colombia and Costa Rica 2015, Dominican Republic 2013, Sri Lanka 2012, Egypt, South Africa and Tunisia 2010, Botswana and Namibia 2009). ^c Proportion of a country's working-age population (15-64 years old) actively engaged in the labor market, either by working or looking for work, in 2015. ^d Share of the labor force that is without work but available for and seeking employment in 2015; modeled ILO estimate. ^e Share of self-employed (i.e., own-account workers and employers) in relation to total employment in 2013 (exceptions: Botswana 2010, Peru 2012). ^f Share of persons with informal jobs (in the formal or informal sector) in relation to total non-agricultural employment in 2013 (exceptions: Ecuador 2011, South Africa 2010, Indonesia and Sri Lanka 2009, Botswana 2006, Egypt 2005 and Algeria 2002). ^g Share of persons employed in the informal sector in relation to total non-agricultural employment in 2013 (exceptions as for informal employment). ^h Number of homicides per 100,000 of the population in 2012 (exceptions: Botswana 2010, Egypt, Thailand 2011). ⁱ Blanc et al. (2007). ^j Including family workers. ^k Six cities only. ^l Main cities/metropolitan areas only.

2.2. Trends in unemployment, self-employment and crime in South Africa

In this subsection we show that the ‘paradox’ relating unemployment and entrepreneurship and the very high level of crime is not the result of very recent transitory developments, but rather is the result of longer term trends which are unlikely to disappear in the near future.

South Africa’s unemployment rate has been at a very high level for a long time. It increased in the 1990s, decreased in the early 2000s, then increased again and is now stagnating at around 25% or even, according to the expanded definition (i.e., including individuals who are not actively looking for work), at around 35%. This is shown in Fig. 1. However, the gap between male and female unemployment has at least narrowed.

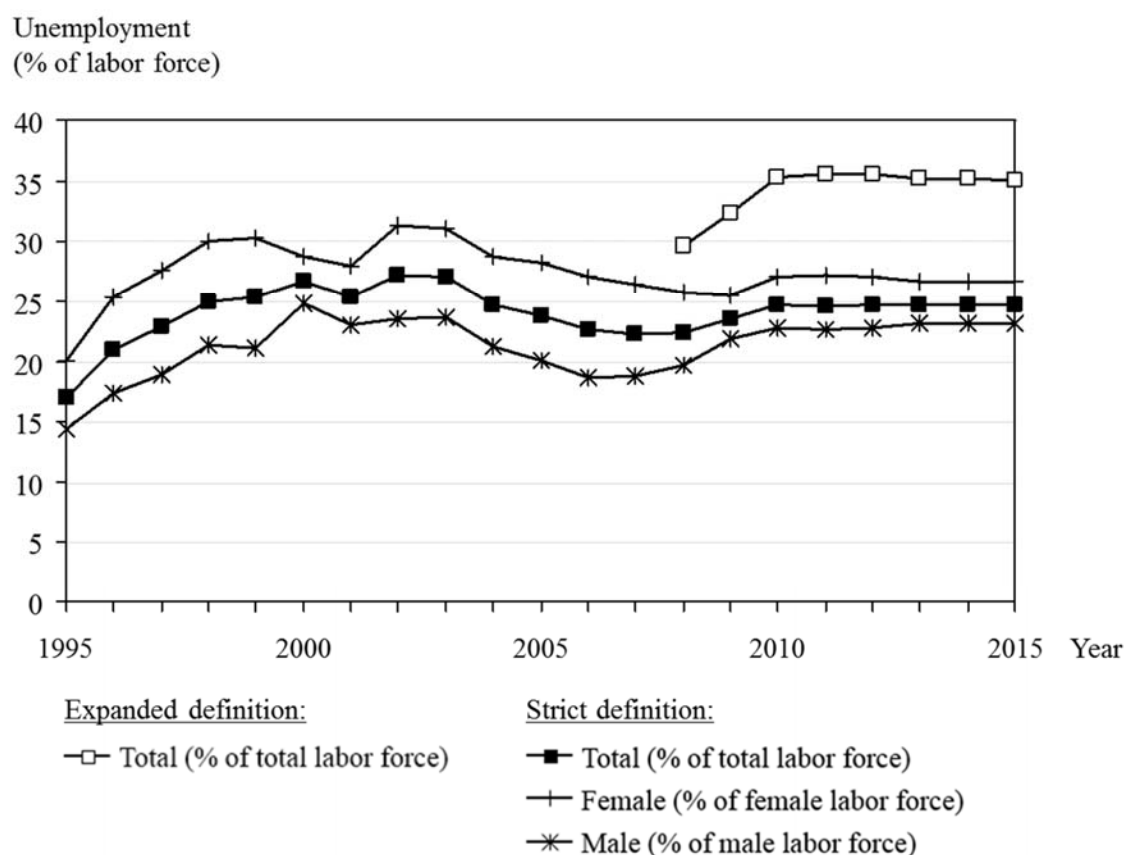


Fig. 1. Development of unemployment from 1995-2015 in South Africa.

Sources: World Development Indicators (World Bank 2015) for 1995-2012 and Quarterly Labor Force Survey (Statistics South Africa 2016) for the strict definition 2013-2015 and expanded definition 2008-2015. Note: Labor force consists of people between 15 and 64 years of age. The unemployment rates for 1995-2012 are end of the year numbers, the rates for 2013-2015 are year averages. The strict definition of unemployment includes all people who are part of the labor force and capable of working, but – although actively searching for a job or trying to start an own business – do not have work. In the expanded definition of unemployment people who are willing to work, but who are not actively looking for a job – so-called discouraged job seekers – are also included.

Fig. 2 presents the development of total, female and male self-employment – i.e., the share of formal and informal employers and own-account workers in relation to total employment – in

South Africa over the last seven years. Despite increasing unemployment rates since 2008, the share of the self-employed in relation to total employment is not only at a very low level, it even decreased – although only slightly – from 15% in 2008 to 14% in 2015. It is also notable that the gap between male and female self-employment is widening. While in 2008 self-employment for both sexes was around 15%, female self-employment dropped to 11% and male increased to 16% in 2015. This is in contrast to other developing and emerging countries, for instance India, where women in particular are the driving force behind increasing self-employment levels.

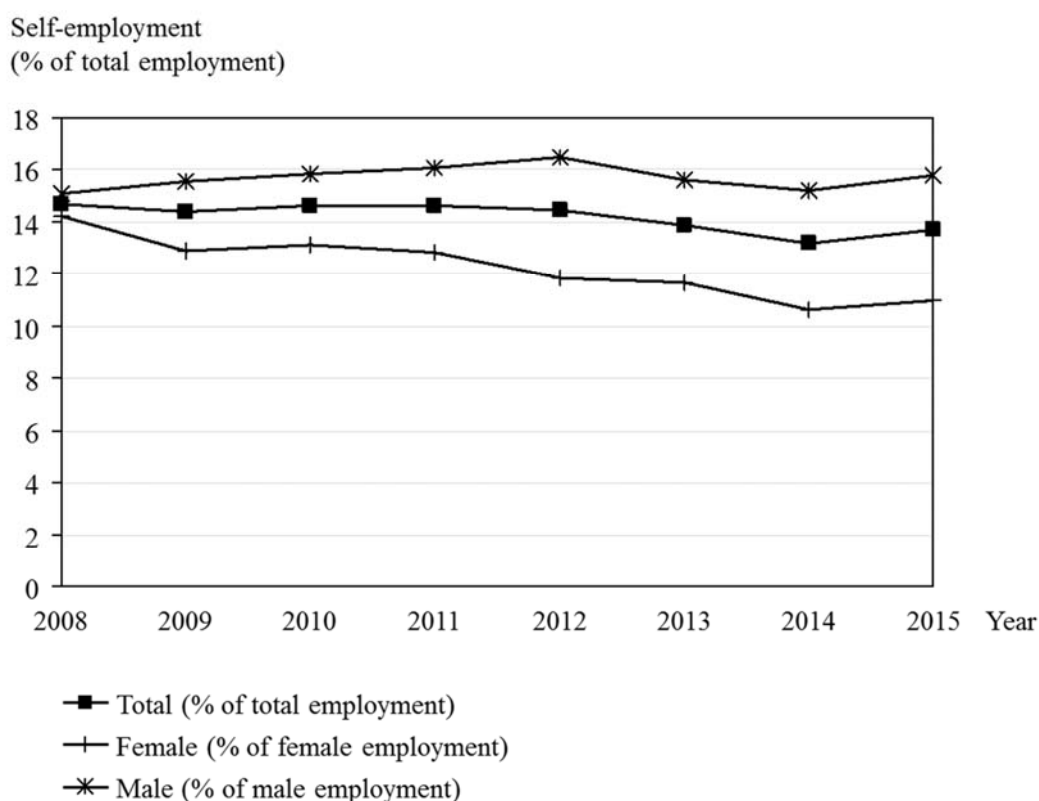


Fig. 2. Development of self-employment from 2008-2015 in South Africa.

Source: Quarterly Labour Force Survey (Statistics South Africa 2016). Note: Self-employment comprises employers as well as own-account workers. The self-employment rates for 2008-2015 are year averages.

Table 2 shows the changes in crime rates for the period 2005 to 2014. It can be noted that crime reduced slightly in almost all crime categories, in particular contact crimes (including murder and sexual assault) as well as in other crime categories (including culpable homicide and kidnapping). Crime incidents related to drugs and the possession of firearms (summarized as ‘Crime detected as a result of police action’), however, experienced an average annual growth of 10% from 2005 to 2014. Overall, it remains clear that the crime level is still particularly high, even by Sub-Saharan African standards.

Table 2

Development of crime rates (# of crimes per 100,000 of the population) in South Africa from 2005-2014 per crime type.

Crime type	Crime rate (# of crimes per 100,000 of the population)										CAGR 2005-14
	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	
Contact crimes	1,800	1,584	1,509	1,411	1,374	1,339	1,245	1,199	1,169	1,158	-5%
Contact-related crimes	338	318	312	294	283	275	257	247	241	234	-4%
Property-related crimes	1,269	1,196	1,131	1,062	1,069	1,095	1,043	1,031	1,068	1,051	-2%
Crime detected as a result of police action	275	298	325	348	376	420	452	500	555	646	+10%
Other serious crimes	1,395	1,155	1,120	1,075	1,109	1,071	1,043	1,033	996	968	-4%
Subcategories of aggravated robbery	55	59	70	81	98	96	84	83	86	94	+6%
Other crime categories	164	136	118	109	102	105	105	99	93	85	-7%
Total crime	5,297	4,747	4,585	4,380	4,411	4,400	4,230	4,192	4,208	4,235	-2%

Source: Crime Stats (South African Police Service 2015).

Note: CAGR = Compound annual growth rate.

In the following section we discuss a theoretical framework linking the exposure to crime and occupational choices. This framework will guide our empirical analysis.

3. Mechanisms linking crime and entrepreneurship

To understand through which channels crime – in particular property-related crime such as robbery, burglary and theft – may impact business ownership, it makes sense to distinguish between direct and indirect effects of crime on potential entrepreneurs as well as on already existing businesses.

The most distinct and direct impact of property-related crime on the victim could be the loss of or damage to property and capital. For existing businesses, the subsequent costs for repair or replacement of property may lead to lower profits and, thus, also to less capital being available for reinvestment. Obviously, a severe loss of property and capital could even push an existing firm into immediate bankruptcy (Edwards and Sundaram 2013). There is some descriptive evidence that such direct effects might be relevant. McDonald (2008) notes that small enterprises in South Africa report facing crime-induced costs that correspond to up to 35% of their total sales. Indirectly, high crime levels may induce increased costs for crime prevention (e.g. CCTV surveillance, security staff, insurance premiums) for existing businesses. Potential entrepreneurs may forego business opportunities altogether as they may fear that the high risk of losing business assets will make their investment unsustainable, or equivalently, may estimate that the costs necessary for prevention will be too high relative to the expected revenues. Obviously, potential or actual entrepreneurs may also fear facing personal injury as a result of crimes such as

robbery, burglary and theft - crimes which may be accompanied by physical violence or even murder (see e.g., Ahwireng-Obeng and Piaray 1999; McDonald 2008; Olawale and Garwe 2010).

In addition to these key channels, the literature evokes further effects of crime on businesses. Kahn (2010), for instance, emphasized that crime may have a negative impact on (urban) agglomeration. Skilled workers in particular may avoid or leave crime-prone areas, which would then lead to a shortage of skilled labor for employers and consequently to lower productivity as well as a lower probability of starting or expanding a business (see also Mahadea and Pillay 2008). Moreover, Gough et al. (2003) and McDonald (2008) point out that the fear of customers to buy from businesses located in areas with high crime rates may negatively impact the sales and profits of a business. Shorter opening hours, in particular the closure of the business before dusk, may additionally decrease their turnover. Suppliers may refuse, for the same reasons, to deliver to businesses in crime-prone areas.

Hence, if all these channels are relevant, we would expect to see the following: Firstly, in areas with relatively high crime rates we should observe that fewer people decide to start a business and that existing entrepreneurs are more likely to keep their business small or close down their business entirely. Econometrically speaking, we would expect a negative impact of crime on the share of self-employment, both across municipalities and across time. Secondly, we would expect in areas with relatively higher crime rates, the costs of doing business to be higher and entrepreneurs to be less willing to invest and, furthermore, to have less resources available for investment and activity expansion, implying lower average sales, profits and investment levels.

Yet, even if these channels sound plausible, it is of course possible that there are also offsetting effects. First, crime prevention may also create business. There is, for example, scope for the retail of certain goods such as video cameras, security doors and locks, as well as for security services. Second, if through the channels mentioned above crime also negatively affects the number of waged jobs that are available, those workers who do not find waged work would possibly seek to become self-employed. Hence, in this case a crime effect would, at first, be visible in relation to waged jobs, but not necessarily in relation to self-employment. The net effect of self-employment could even be positive if more workers entered than quitted because of crime.

Hence theoretically, the effects are unclear. As we have shown above, existing empirical work on South Africa seems to support the hypothesis that crime reduces entrepreneurship, but, again, these studies work with geographically limited samples and base their assessment largely on perception data. All this is reason enough to re-visit the question.

4. Data and empirical strategy

4.1. Data

To investigate the impact of crime on self-employment and business performance, we use four main data sources: official crime statistics from the South African Police Service, data from the Census, from the Community Survey, and from the Survey of Employers and the Self-Employed from Statistics South Africa.

The crime statistics from the South African Police Service ('SAPS'), the so-called 'Crime Stats', are annually published statistics on total reported incidents for 29 different crime categories such as murder, sexual crimes, burglary and robbery, in each of the (around) 1,140 police stations in South Africa. This data is available on an annual basis for the years 2004/05 to

2014/15.⁴ To construct crime rates, i.e., the number of incidents per 100,000 of the population, we use population figures at the municipality and district levels for the years 2001 and 2011 that are published by Statistics South Africa ('Stats SA') and inter- and extrapolate these numbers for the missing years.

As our identification strategy will make use of the variation in crime rates between and within municipalities and districts over time, we now shed some light on these variations. When analyzing, for instance, robbery and burglary at residential and non-residential premises,⁵ we find large differences across municipalities with rates ranging from 100 to almost 5,000 incidents per 100,000 of the population. Moreover, the changes in crime rates over the period 2007/08 to 2011/12 also differ significantly across municipalities. We do not find any correlation between the initial crime level and the change in crime rates over time. Even within urban areas, we find a high variation of crime rates, ranging from 180 to 2.100 incidents per 100,000 of the population and we find no common trend in robbery and burglary crime rates for municipalities with a high population density (the crime growth rate varies from -29% to 49% from 2007/08 to 2011/12 in these areas). Looking at the yearly changes across municipalities, we find strong fluctuations. For instance, while the robbery and burglary crime rate in Overstrand decreased from 2004/05 to 2005/06 by 31%, it increased in the next year by 11%, before it again decreased by 9%. These variations in year-to-year changes, however, can be observed across all municipalities. To smooth out these year-to-year fluctuations we will use aggregated crime rates for the previous two to three years in our estimations.

As Crime Stats reflect only reported crime incidents, actual crime numbers and rates might be different. As Kock et al. (2015) emphasized in their guide to South African crime statistics, particularly the unwillingness and inability of victims to report and of police officers to record a crime incident, lead to an underreporting of real crime rates. For instance, the so-called 'Victims of Crime Survey', conducted by Stats SA, asked households across South Africa about their experience of crime. They found that from April 2013 until February 2014 100% of all interviewed individuals who have been victims of car hijacking also reported this crime to the police, while only 31% of robbery and theft victims did so (Statistics South Africa 2014a). This underreporting of some crime types may give us a distorted picture of the actual crime level in a municipality. However, as long as underreporting does not vary within municipalities over time, accounting for municipality fixed-effects can deal with this problem. Time effects can capture general countrywide changes in underreporting. Classical measurement error may lead to a general downward bias in our estimates.

To test whether there is a significant difference in reporting behavior across locations, we use data from the Victims of Crime Survey for the years 2003, 2007, 2011, 2012 and 2013/14 (Statistics South Africa 2003, 2007b, 2011c, 2012, 2014b). We measure underreporting by using the information on whether an individual reported a robbery to the police or not. We find no evidence for significant differences in reporting behavior across provinces (the data does not allow a disaggregation at the municipality level).⁶ Yet this does, of course, not exclude the

⁴ Crime Stats are released annually in September and cover reported crime incidents from April 1 in the previous year to March 31 in the same year. For instance, the Crime Stats published in September 2015, i.e., Crime Stats 2014/15, covered crime incidents from April 1, 2014 until March 31, 2015.

⁵ The crime rate for robbery and burglary at residential and non-residential premises is an aggregated crime rate consisting of the Crime Stats crime types robbery at non-residential premises, burglary at non-residential premises, robbery at residential premises, burglary at residential premises, common robbery and robbery with aggravating circumstances.

⁶ In the Victims of Crime Survey, the lowest geographical area that can be identified is the province. No information on the district or the municipality of the respondent is given.

possibility that there may be a systematic variation of underreporting at the municipality or district level. In general, the reporting rate may, for instance, be lower in municipalities with a high actual crime level, due to the victims' fear of revenge by the offender or because people who are used to crime do not feel the necessity to report an incident. They may also think that the police cannot, in any case, do much about it or does not want to do anything about it. Again, across provinces we do not find evidence for such a bias, as underreporting does not systematically vary across provinces. We may, of course, still face such a bias within provinces over time if, for example, with increasing crime levels, underreporting goes up. This would lead to a downward bias of our estimate of the effect of crime on self-employment. Again, we think such a bias is unlikely, as we do not even see significant differences in reporting behavior across provinces, despite the fact that crime rates differ strongly across provinces (see also Online Appendix, Fig. B1). More problematic for our analysis would be a situation in which changing policies at the municipality level, such as increased police presence or simplified modes of reporting a crime, would affect reporting behavior. Under such circumstances, a municipality with a constant actual crime rate could show an increase in the crime rate based on reported crimes. This would show up as an increase in crime in the Crime Stats and bias our estimates. We are not aware of such cases and such policies would anyway more likely affect entire provinces rather than single municipalities, but we can of course not fully rule out such cases.

The data on employment, specifically on entrepreneurial activities or self-employment, as well as other socio-economic characteristics, are drawn from the Census and the Community Survey. The Community Survey is a mini-census, i.e. the questionnaire is similar to the one used in the Census but it is conducted only on a sub-sample of all communities in South Africa. Next to these data, we use the Survey of Employers and the Self-Employed ('SESE'). The Census and Community Survey data allows measuring self-employment at the individual level and self-employment shares at the municipality level. The SESE data additionally provide the characteristics and performance indicators of businesses.

The Census and the Community Survey are conducted by Stats SA. Data from the Census is publicly available for the years 1996, 2001 and 2011, from the Community Survey for the year 2007. Because of the different timing of the Crime Stats and the Census and Community Survey, we focus in particular on the Census from 2011 and merge it with the Community Survey from 2007. The publically available Census data is a 10% sub-sample (approx. 4.3 million observations). The Community Survey covers about 0.9 million observations. Both data are gathered at the individual level and allow for the identification of the location of respondents down to the municipality level. However, individuals cannot be tracked over time; the data is just a series of repeated cross-sections. Panel data analysis is, however, possible at a higher aggregation level.

The Census and the Community Survey data both contain up to 90 variables relating to the individual's socio-economic characteristics, such as sex, age, family situation, education, health and employment (e.g., employment status and annual income range). We code workers as self-employed if they report either to be an employer or an own-account worker. In our analysis we restrict the dataset to individuals in the 16 to 59 age range and, moreover, exclude individuals who are currently attending school or – according to their own assessment – unable to work due to illness or disability. We thus reduce our dataset to 2.5 million observations for the years 2007 and 2011. See also Table 3 for an excerpt of the summary statistics of some of the key variables in the merged Census data (i.e., the merged data of the Community Survey 2007 and the Census 2011). In terms of the share of respondents based on self-employed status, age, gender and the share of the white population, the merged Census data is relatively balanced for 2007 and 2011 with the exception of the share of individuals with no education or only primary education

(dummy ‘No education’). This share is in 2011 five percentage points smaller than in 2007. In both years, 8% of the individuals are white, less than 50% are male and the individuals are on average 35 years old. Between 6% and 7% are self-employed. The average annual income per capita per municipality increased from 2007 to 2011 from 42,427 South African Rand (ZAR) to 48,389 ZAR (in 2011 real terms). Converted into international Dollars (\$) by using Purchasing Power Parities (PPP) (World Bank 2015), this means an increase in the average annual income per capita by approximately 14% from \$8,372 to \$9,549. Moreover, we can observe a diminishing gap from 2007 until 2011 between the municipality with the lowest average income and the municipality with the highest average income.

Table 3
Excerpt of Summary Statistics of merged Census data.

Variable	2007			2011		
	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.
Self-employment	432,802	0.06		2,083,463	0.07	
Age	432,802	35.32	11.23	2,083,463	35.73	11.31
Male	432,802	0.46		2,083,463	0.47	
White	432,802	0.08		2,083,463	0.08	
No education	432,802	0.23		2,083,463	0.17	
Avg. income (ZAR) ^a	432,802	42,427	21,594	2,083,463	48,389	24,213
Avg. income (\$ PPP) ^b	432,802	8,372	4,261	2,083,463	9,549	4,778

Source: Community Survey 2007 and Census 2011 (Statistics South Africa 2007a, 2011b).

^a Average annual income per capita per municipality in South African Rand and in real terms, with 2011 being the reference year; figures for annual consumer price inflation from World Development Indicators (World Bank 2015).

^b Average annual income per capita per municipality in international \$ and in real terms, with 2011 being the reference year and conversion from ZAR into international \$ by using the Purchasing Power Parity (PPP) conversion factor for 2011; figures for annual consumer price inflation and PPP conversion factor from World Development Indicators (World Bank 2015).

For the sake of simplicity, we henceforth refer to the combined dataset of both sources – Census 2011 and Community Survey 2007 – as ‘merged Census data’.

The Survey of Employers and the Self-Employed (‘SESE’) was conducted by Stats SA in the years 2001, 2005, 2009 and 2013. In our analysis we consider the latest three waves only. The survey focuses on small and micro non-VAT paying businesses throughout South Africa. The surveys contain information on the business (e.g., financial information, type of business and industry, location and number of employees) as well as on its owner (e.g., age, gender and relationship status). The surveyed businesses are drawn from a representative sample provided by the so-called Quarterly Labour Force Survey, which is conducted every three months by Stats SA covering all of South Africa. In each of the three considered SESE waves, between 1,900 and 3,300 businesses were interviewed. We use the information on sales, profits and investment behavior of the businesses as dependent variables for our analysis. The summary statistics for the SESE data (Table 4) show that business performance indicators, such as sales and profits, have – after correcting for inflation – increased over time. This growth can be observed across all municipalities and is not specific to particular municipalities. In all three years, around 50% of businesses are less than 3 years old, less than 50% of the business owners are male and less than

5% white. The share of businesses which had made any investments during the previous year increases from 15% in 2005 to 17% in 2009 and up to 23% in 2013. Moreover, the number of paid employees (excluding the owner) also increases over the years – from on average 0.2 in 2005, to over 0.3 in 2009 and 0.4 per firm in 2013. While the majority of the entrepreneurs stated that they needed capital to start up their business (74% in 2005, 65% in 2009 and 58% in 2013), only about 5% of the business owners have debt in any of the three years. In 2005 96% of the respondents are the single owners of their business – compared to 76% in 2009 and 91% in 2013.⁷ Between 16% and 24% of the owners started their business due to family tradition, existing skills and motivation for the business, or a business idea (dummy ‘business motivation’). While no information is available on the business owners’ highest achieved education level for the year 2005, in 2009 31% and in 2013 25% of the respondents have no formal education.

Table 4
Excerpt of Summary Statistics of SESE data.

Variable	2005			2009			2013		
	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.
Age	2,987	42.826	13.078	1,822	42.432	12.024	1,929	43.217	12.320
Male	2,990	0.390		1,822	0.453		1,929	0.488	
White	2,990	0.024		1,822	0.032		1,929	0.034	
Paid employees (#)	2,990	0.218	0.914	1,822	0.341	1.652	1,929	0.443	2.321
Firm age < 3 years	2,990	0.528		1,822	0.450		1,929	0.422	
Investment in machinery (last year)	2,990	0.028		1,822	0.026		1,929	0.041	
Any investment (last year)	2,990	0.145		1,822	0.168		1,929	0.234	
No education	0	-		1,822	0.308		1,929	0.246	
Single owner	2,990	0.955		1,822	0.755		1,929	0.908	
Business motivation	2,990	0.162		1,822	0.182		1,929	0.238	
Any debt	2,990	0.042		1,822	0.050		1,929	0.040	
Startup capital needed	2,990	0.744		1,822	0.645		1,929	0.579	
Electricity expenses (ZAR) ^a	2,966	103	271	1,822	110	388	1,929	138	361
Sales last month (ZAR) ^a	2,949	2,232	4,764	1,822	3,139	7,876	1,929	3,621	11,567
Profit per month (ZAR) ^a	2,969	1,384	3,106	1,822	1,703	3,758	1,929	2,318	5,737

Source: Survey of Employers and the Self-Employed from 2005, 2009 and 2013 (Statistics South Africa 2015b).

Note: Excluding districts with less than 10 observations in one of the three waves.

^a Electricity expenses and sales last month as well as profit per month in South African Rand and in real terms, with 2013 being the reference year; figures for annual consumer price inflation from World Development Indicators (World Bank 2015).

⁷ Compared to 96% single owners in 2005 and 91% in 2013, the share of only 76% single owners in 2009 is relatively low. We can exclude differences in the questionnaire as the reason for this (in all three years the same question – whether the business is owned by a single owner – was asked), but cannot otherwise explain the difference in a reasonable way.

Table 5 gives an overview of the number of observations per municipality and district (both including metro municipalities) for merged Census and SESE data. The average number of observations per municipality in the merged Census data is – with a mean of around 4,000 in 2007 and 18,500 in 2011 – relatively high. In addition, the lowest number of observations per municipality lies between 500 and 570 in both years. Consequently, we are certain that the number of observations per municipality is high enough to use the variance within municipalities and across time for our analysis. Turning to the SESE data, the number of observations per municipality is – with a mean of sixteen for 2005, eleven for 2009 and twelve for 2013 – relatively small. Moreover, while few municipalities, mostly large urban metro municipalities, have up to 135 observations each, in more than 50% of the municipalities less than ten observations each are available per year. We therefore prefer to focus on within-district differences across time instead of within-municipality differences when using the SESE data. The average number of observations per district is 65 for 2005, 38 for 2009 and 39 for 2013. However, in each of the three years there are some districts (one, ten and eight districts, respectively) with less than ten observations – these districts are excluded from our analysis.

Table 5
Number of observations per municipality and district in merged Census and SESE data.

	Merged Census		SESE		
	2007	2011	2005	2009	2013
Observations per municipality					
Mean	4,030	18,537	16	11	12
Min.	503	568	1	1	1
Max.	64,549	372,764	124	131	135
Total # of municipalities	234	234	212	176	176
thereof municipalities with less than 10 observations	0	0	107	118	120
Observations per district					
Mean	18,136	83,417	65	38	39
Min.	3,062	5,473	7	2	1
Max.	64,549	372,764	160	160	170
Total # of districts	52	52	52	51	52
thereof districts with less than 10 observations	0	0	1	10	8

Source: Survey of Employers and the Self-Employed from 2005, 2009 and 2013, Census from 2011 and Community Survey from 2007 (Statistics South Africa 2007a, 2011b, 2015b).

To use the data for our purpose, four datasets have been merged into two datasets: First, Crime Stats with the Census and Community Survey data to analyze the impact of crime on self-employment, and second, Crime Stats with SESE data to analyze the impact of crime on the performance of businesses. For both merged datasets, substantial adaptation of the data was required.

Due to significant changes to the demarcations and names of both districts and municipalities over the last twenty years, the Census, Community Survey and SESE data had to be adapted to share a common set of administrative areas. To do this we took the Census 2011's demarcations and names as reference and adapted previous demarcations and names accordingly. For instance, while the municipalities Kagisano and Molopo are separated in the Community Survey 2007

data, they are merged into one municipality called Kagisano/Molopo in the Census 2011 data. Today's Tlokwe Local Municipality was in 2007 still called Potchefstroom Local Municipality and Victor Khanye was called Delmas Local Municipality. In order to match old and new demarcations and names, we used geographical metadata from Stats SA and partly fixed municipality codes in the datasets (although changing names) alongside information from the official municipalities' websites.

Another need for adaptation was driven by the different geographical levels that can be identified in the datasets. As we intend to exploit the variation of crime over time and between administrative areas in our estimation, we had to identify the lowest possible geographical level across all datasets. Crime Stats are published at the police station jurisdiction and province level, however, the lowest geographical level that can be identified in the merged Census data is the municipality – which is a larger geographical area – and in the SESE data is the so-called sub place – which is a smaller geographical area than the police station jurisdiction (see Appendix, Fig. A1). Consequently, we needed to aggregate Crime Stats to the municipality level, the lowest possible geographical level that can be identified in both the merged Census and SESE data. To match the police stations to their respective municipality, we used the crime maps from the 'ISS CrimeHub'⁸ as well as information on police station locations from the South African Police Service. Yet in some cases, police station jurisdictions are located not only in one but in two or sometimes even more municipalities. However, as this concerns only very few of the municipalities and as usually only small parts and not very densely populated areas fall into another municipality, this inaccuracy in matching does not significantly affect our analysis.

4.2. Empirical strategy

We test the hypothesis that crime influences the decision to set up a firm by estimating the impact of crime on 1) business ownership or being self-employed, and on 2) business performance.

4.2.1. Impact of crime on self-employment

In a first step, we use a linear probability model with municipality and time fixed-effects to estimate the impact of crime on self-employment using the dataset containing Crime Stats and merged Census data. Our model takes the following form:

$$SelfEmployment_{ikt} = \alpha + \beta Crime_{kt-1} + \gamma X_{ikt} + \mu Z_{kt} + \delta_k + \theta_t + \varepsilon_{ikt} \quad (1)$$

where the subscript i denotes the individual, k the municipality, and t the year.

We define our dependent variable for self-employment, $SelfEmployment_{ikt}$, as a binary variable, which equals 1 if the individual is an employer or own-account worker and 0 otherwise (i.e., employed (formally or informally), unpaid family worker or unemployed).⁹ As a proxy for

⁸ See <https://www.issafrica.org/crimehub/>.

⁹ In the literature as well as in official data on self-employment, the self-employment share is usually defined as the share of employers and own-account workers in relation to total employment (i.e., excluding the unemployed). In our analysis, we also include unemployed individuals in the dataset. Therefore, the dummy $SelfEmployment_{ikt}$ shows whether an individual is either an employer or own-account worker ('1'), or an employee, unpaid family worker or unemployed ('0').

the crime level, we include the independent variable $Crime_{kt-1}$. We generate a crime rate for burglary and robbery in residential and non-residential premises by aggregating the distinct Crime Stats for robbery at non-residential premises, burglary at non-residential premises, robbery at residential premises, burglary at residential premises, common robbery and robbery with aggravating circumstances. The crime variable is expressed as the log of crime incidents per 100,000 of the population for the two years before the point of observation. We prefer the two-year aggregation as it smooths out short-term fluctuations. X_{ikt} is a vector of control variables including age, gender, education and the population group of the individual. The age (age squared) of the individual is included as a continuous variable. Gender is included as a dummy, which equals 1 if the individual is male. We also construct a dummy to control for education, which equals 1 if the person has no education or only primary education and 0 otherwise (i.e., secondary education or higher). In some specifications, we additionally include an interaction term of education and the crime rate as the effect of crime on self-employment may vary with education, for instance because individuals that are more educated may have a different level of access to information about crime or process this information differently. We also include a dummy for the population group, which equals 1 if the individual is white and 0 otherwise. To control for the effect of income we calculate the log average annual income per capita in the municipality the individual lives in (Z_{kt}). Municipality fixed-effects and year effects are denoted δ_k and θ_t respectively. In some of the specifications we also account for district-specific time trends (i.e., interact year with district dummies); this is more conservative than just accounting for general time trends. Standard errors of the fixed-effects regression are robust and clustered at the municipality level.

Although our approach allows accounting for many potential confounders, equation (1) might still be subject to reverse causality, i.e., business activities may attract crime, and may still be subject to an omitted variable bias. To deal with these problems, we additionally use an instrumental variable approach. We considered various potential instrumental variables, such as police expenditure, criminal law, incarceration rates and liquor stores regulations. However, either appropriate data was not available, or we could not explicitly attribute the observations to the different municipalities, or there was not enough variation between and within municipalities across time. Consequently, we use – as is frequently done in the literature on the economic effects of crime (e.g., Tita et al. 2006)¹⁰ – data on other types of crime as an instrument. The identifying assumption is that municipalities with a high robbery and burglary crime rate are also likely to have a high overall crime rate, and therefore also a high level of contact and, more specifically, sexual crime. While there should be a strong correlation between the instrument (contact or sexual crime) and our robbery and burglary crime rate, the instrument should have no direct influence on the decision to start or close a business and on business performance, and, therefore, our dependent variable $SelfEmployment_{ikt}$. Hence, in the first stage, we regress our potentially endogenous variable $Crime_{kt-1}$ on the instrumental variable, I_{kt-1} , as well as on the already in equation (1) included controls:

$$1^{st} \text{ stage: } Crime_{kt-1} = \alpha + \beta I_{kt-1} + \gamma \tilde{X}_{ikt} + \mu \tilde{Z}_{kt} + \delta_k + \theta_t + \mu_{ikt} \quad (2)$$

¹⁰ Analyzing the impact of crime on housing prices, Tita et al. (2006) use homicide as an instrument for total, violent and property crime, thereby particularly addressing the problem of existing underreporting of the instrumented crime types.

Our instrument I_{kt-1} is the aggregated contact or sexual crime rate for the previous two years in the municipality in which the individual resides. Since the first stage uses only variation on both the left- and right-hand side of the equation at the level of municipalities, we estimate the first stage using the municipality as the observation unit. All control variables are municipality averages; this is denoted by a tilde above the variables. In the second stage we use the predicted variable of the burglary and robbery crime rate, \widehat{Crime}_{kt-1} , as the explanatory variable:

$$2^{\text{nd}} \text{ stage: } SelfEmployment_{ikt} = \alpha + \beta \widehat{Crime}_{kt-1} + \gamma X_{ikt} + \mu Z_{kt} + \delta_k + \theta_t + \varepsilon_{ikt} \quad (3)$$

All variables, dependent and independent, are defined as in equation (1).

To further test the robustness of our results, we estimate a large set of alternative specifications and use alternative variables. Specifically, we test our hypothesis using other crime types as independent variables, such as the total crime rate or total property crime rate. We also vary crime rate calculations, e.g., we use the growth of the crime rate over the previous two years or, alternatively, the aggregated crime rate for the previous three years. In further checks, we use subsets of the data, such as data with men or urban areas only.

4.2.2. Impact of crime on business performance

In a second step, we estimate the impact of crime on business performance indicators, again using a linear model with fixed-effects and the dataset containing Crime Stats and SESE data. Our identification strategy uses the variation of crime rates within districts across municipalities and over time. Our model takes the following form:

$$Y_{ikt} = \alpha + \beta Crime_{kt-1} + \gamma X_{ikt} + \delta_l + \theta_t + \varepsilon_{ikt} \quad (4)$$

where the subscript i denotes the individual, k the municipality, l the district, and t the year.

We measure businesses' performance (Y_{ikt}) by looking, alternatively, at sales, profits and total investment. Sales is defined as total sales in the last month and profit as average monthly profits in the last year (both in log). Total investment is a dummy variable, which equals 1 if the business has made any investment (i.e., in machinery, tools, vehicles, buildings or furniture) in the last year and 0 otherwise. The independent variable is defined – as before – as the log of the aggregated burglary and robbery crime rate (i.e., crime incidents per 100,000 of the population) in the municipality in which the business is located, for the two years before the point of time when the business was surveyed. X_{ikt} is a vector including the control variables age, gender and population group of the business owner. These variables are defined as in equation (1). X_{ikt} further includes variables which capture the potential impact of firm age, ownership, motivation for self-employment, debt, startup capital, electricity expenses and firm size on the performance of the business. Firm age, ownership, motivation for self-employment, debt and startup capital enter as dummy variables (for a description of the variables see previous section). Electricity expenses enter as a continuous variable. We account for firm size through the categorical variable 'Employees', which equals 0 if the individual has no employee, 1 for one employee, 2 for two to four, 3 for five to ten and 4 for more than ten employees. We also include district, δ_l , and year effects, θ_t . Furthermore, in the last of our five specifications, we account for province-specific time trends. Standard errors in the fixed-effects regression are robust and clustered at the district level. Since some of the control variables are potentially endogenous but leaving them out could

lead to an omitted variable bias, we test the robustness of the estimated effect associated with the crime rate by alternatively including or excluding these controls. If this is without implications for the effect associated with crime, it can be assumed that they will not bias our assessment.

As in our previous model, we re-estimate the model using the same instrumental variable strategy as before. We also conduct several robustness checks to test the sensitivity of our results with respect to alternative specifications and data subsets.

5. Results

5.1. Impact of crime on self-employment

Col. 1 of Table 6 shows the results for the OLS regression, while in cols. 2 to 5 we present, step-by-step including fixed-effects and additional control variables, the results for the fixed-effects model as described in equation (1) (for details see Online Appendix, Table B1).¹¹

Table 6
Impact of crime on self-employment (merged Census) – OLS and fixed-effects regression results.

Dependent variable	Self-employment (dummy)				
	OLS	Fixed-effects			
	(1)	(2)	(3)	(4)	(5)
Crime: Robbery & burglary (ln)	-0.0066*** (0.0004)	-0.0135** (0.0066)	-0.0128* (0.0065)	-0.0146** (0.0066)	-0.0062 (0.0065)
Municipality FE	No	Yes	Yes	Yes	Yes
Year FE	No	No	Yes	Yes	Yes
No education x Crime (ln)	No	No	No	Yes	Yes
District-specific time trends	No	No	No	No	Yes
Observations	2,516,265	2,516,265	2,516,265	2,516,265	2,516,265
Number of municipalities	234	234	234	234	234
R-squared	0.025	0.022	0.022	0.022	0.022

Note: The variables male (dummy), age, age squared, average income (ln), white (dummy) and no education (dummy) are included as control variables in each of the regressions. Robust standard errors clustered at municipality level in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

Starting with the OLS regression (col. 1), we note that the coefficient for crime is negative and significant at the 1% level, indicating that a 1% increase in the crime rate leads to a decrease in the probability of being self-employed by 0.0066 percentage points. Put differently, an increase by one standard deviation in the crime rate leads to a decrease in self-employment by 0.05 percentage points. With respect to our controls, the probability of self-employment is higher for men and white individuals (0.03 and 0.09 percentage points, respectively) and slightly increases

¹¹ We tested whether a random-effects estimation is more suitable for our model compared to a fixed-effects estimation, by running the Hausman test as well as another test of over-identifying restrictions developed by Schaffer and Stillman (2006). The latter test allows us – in contrast to the Hausman test – to also include clustered robust standard errors. The results of both tests (see Appendix, Table A1) show that the use of a fixed-effects estimation is more appropriate for our model than a random-effects model.

with age (but at a decreasing rate) and with education (self-employment is lower for individuals with no education by 0.01 percentage points). The average income level in a municipality has a positive impact on self-employment, i.e. a 1% increase in income leads to an increase of the probability of being self-employed by 0.02 percentage points.

When we include municipality fixed-effects (col. 2), the coefficient associated with crime more than doubles (from -0.0066 to -0.0135) and the precision of the coefficient is somewhat lower. It further decreases if we include year effects (col. 3). In the fourth specification (col. 4), we include the interaction term ‘No education’ interacted with ‘Crime (ln)’. The coefficient of the no education-crime interaction is positive, indicating that crime has a less negative impact on self-employment for individuals with no education compared to educated individuals. We can of course only speculate as to why this is the case. Maybe educated individuals are more aware of the prevalence of crime or are more risk-averse than non-educated individuals are, and as a consequence are less likely to become self-employed at a given level of crime. In the last specification (col. 5), district-specific time trends are included. In this estimation, the coefficient of crime becomes insignificant. Hence, based on the results shown in Table 6, we must conclude that although the effect of crime on self-employment is negative, in terms of its economic significance it is negligibly small. If confirmed, this would suggest that crime is unlikely to be – as it is often claimed – one of the major reasons for the low self-employment in South Africa.

To further deal with potential problems of omitted variables and reverse causality, we also apply an instrumental variable approach as specified in equations (2) and (3). We test two instruments – contact crime and sexual crime – and take the third fixed-effects specification (Table 6 col. 3) as the benchmark specification. The results are shown in Table 7 (for details see Online Appendix, Table B2).

Table 7

Impact of crime on self-employment (merged Census) – Instrumental variable regression results.

	Recap: Fixed-effects	Instrumental Variable	
		Contact crime	Sexual crime
	(1)	(2)	(3)
2 nd stage – Self-employment (dummy):			
Crime: Robbery & burglary (ln)	-0.0128* (0.0065)	-0.0088*** (0.0005)	-0.0114*** (0.0005)
Municipality FE	Yes	No	No
Year FE	Yes	Yes	Yes
Observations	2,516,265	2,516,265	2,516,265
Number of municipalities	234	234	234
R-squared	0.022	0.025	0.025
1 st stage – Crime: Robbery & burglary (ln)			
Contact/Sexual crime (ln)		0.8076*** (0.0462)	0.8207*** (0.0542)
Year FE		Yes	Yes
F statistic on instrument		305	229

Note: We manually estimate the instrumental variable model. In the first stage, we use an aggregated dataset (i.e., one observation per municipality and per year) and regress the instruments on our crime rate of interest, Crime: Robbery & burglary (ln), to predict the crime variable. In the second stage we use the main dataset (i.e., disaggregated with many observations per municipality) and regress the business performance indicators on the predicted crime variable and other controls. The variables male (dummy), age, age squared, average income (ln),

white (dummy) and no education (dummy) are included as control variables in each of the regressions. Robust standard errors clustered at municipality level in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

The first stage results (second part of Table 7), i.e., the regression of our robbery and burglary crime rate on the instruments, suggest a sufficiently strong correlation between robbery and burglary for both instruments. Moreover, the F statistics for both regressions are clearly above the required thresholds of roughly ten (Stock and Yogo 2005), indicating that contact as well as sexual crime are relevant instruments in our context.

The results of the second stage of the instrumental variable approach are presented in the first part of Table 7. To ease the comparison of the results from the instrumental variable approach with the results from the fixed-effects model, the results from our fixed-effects benchmark specification are also shown in this table.

It is noticeable that there is no considerable difference in the results from our instrumental variable approach (Table 7 cols. 2 and 3) compared to the results without instrumentation (Table 7 col. 1). The coefficients associated with the crime rate are just slightly lower when using contact or sexual crime as an instrument (-0.0088 and -0.0114 compared to -0.0128), but the significance level increases from 10% to 1%. The coefficients of the control variables all have the same signs as before. The associated coefficients are identical or almost identical.

To further test the robustness of our results, we also re-estimate the model using different definitions of the crime rate and alternative subsets of the data (for an overview of the robustness checks see Appendix, Table A2 and Table A3). As before, we use the third fixed-effects specification (Table 6 col. 3) as our benchmark specification.

When using the full dataset (Appendix, Table A2), but including the aggregated robbery and burglary crime rate for the previous three years as an explanatory variable – instead of the previous two years as in our benchmark specification – the sign and size of the crime rate coefficient as well as the significance level remain the same. Other robustness checks involve the inclusion of crime growth rates in the regression, as individuals might not only be sensitive to absolute levels of the crime rate, but also to changes in the crime rate. When including the growth rate of robbery and burglary levels for the previous two years as well as a variable for the crime level two years before that, the coefficients for both variables are negative, but insignificant. Also, the growth rate for the previous three years seems to have no significant impact on self-employment, while the level of crime three years before that – included as control in the same regression – is significant at the 5% level and the coefficient has a similar size as in our benchmark specification. Both coefficients are negative. The results of our main analysis and these robustness checks suggest that the crime level, i.e. the number of burglary and robbery incidents per 100,000 of the population, has a statistically significant, though economically negligible, negative impact on self-employment. We also test to what extent other types of crime may impact self-employment. Using aggregated crime rates for the previous two years, our regression results show that total crime (i.e., all crime incidents, ranging from, e.g., insult to murder) does have a negative, but not significant impact on self-employment. Likewise, the aggregated crime rate for property crime, which includes, besides robbery and burglary, theft of motor vehicles and stock theft, does not significantly impact the probability of being self-employed.

To further test the robustness of our results, we construct various subsets of the data (Appendix, Table A3). On the one hand, we want to understand whether the impact of robbery and burglary is similar when excluding from our dataset individuals who are employed in the formal sector. We find no significant impact of crime on self-employment in this case. On the other hand, we test whether crime impacts self-employment when only considering self-employment in the informal sector – in contrast to self-employment in both the formal and the informal sector as in our benchmark specification. Therefore, we exclude all individuals who are self-employed in the formal sector from our dataset, as we assume that these individuals are unlikely to become self-employed in the informal sector. We find a slightly positive impact of crime on self-employment in the informal sector. The coefficient is, however, not significant. We also tested whether crime reduces dependent wage work. Recall that a possible scenario could be that crime deters some people from self-employment, but that this reduction is offset by people who cannot find a job as a waged worker due to crime-induced labor demand shortages. Hence, these individuals move into self-employment, so that the net effect is small or insignificant. Again, we do not find any evidence for this scenario (see Online Appendix, Table B3).

When we exclude rural municipalities (defined as municipalities with less than 400 people per square kilometer) from our main sample, the coefficient associated with the crime rate is negative but again not significant. An estimation for men alone does also not yield different results. In another robustness check, we exclude all white individuals from the sample. Again, the results show a very small, albeit negative and significant (at 5%), effect. We also re-estimate the model only for those municipalities in which previous studies found a negative correlation between perceived crime levels and self-employment.¹² In fact, for these areas we find a slightly positive, significant effect but not – as one would expect – a negative effect. Moreover, we test our hypothesis by analyzing the impact of crime on self-employment at an aggregated level. To do this we collapse our data at the municipality level for each year. Again, we find no significant impact of crime on the share of self-employment using this specification (Appendix, Table A3).

Finally, we analyze the variation of the impact of crime on self-employment between low and high crime areas, by dividing the municipalities into quintiles and running separate estimations for the five groups. Our results (Appendix, Table A4) suggest a positive, though not significant, impact of crime for the first, third and fifth quintiles. For the second and fourth quintiles the coefficients are negative, but also not significant.

Overall, the significance and – in particular – the magnitude of the impact of crime on self-employment suggests that crime does not deter individuals from self-employment. Although the coefficient associated with robbery and burglary is in some specifications negative and significant, the magnitude of the effect is negligible. Other types of crime play no role at all.

5.2. *Impact of crime on business performance*

In this subsection, we analyze whether crime affects the performance of businesses as measured by sales, profit and total investment (see equation (4)). The results are presented in Table 8 (for details see Online Appendix, Table B4 to Table B6). As before, we show in col. 1,

¹² These municipalities are Johannesburg (Chandra et al. 2002; McDonald 2008) incl. Diepsloot township (Bhorat and Naidoo 2015), Cape Town (McDonald 2008) incl. the township of Khayelitsha (McDonald 2008; Cichello et al. 2011), Durban/eThekweni (Skinner 2005; McDonald 2008) and Pretoria/City of Tshwane (Gough et al. 2003).

the OLS results, while cols. 2 to 5 show the results of the fixed-effects model using alternative specifications.¹³ Robust standard errors are shown in parentheses and are clustered at the district level.

Table 8

Impact of crime on business performance (SESE): Sales, profit and total investment – OLS and fixed-effects regression results.

	OLS		Fixed-effects		
	(1)	(2)	(3)	(4)	(5)
Sales last month (ln)					
Crime: Robbery & burglary (ln)	0.1920*** (0.0243)	0.0914 (0.0616)	0.1918*** (0.0511)	0.3006** (0.1144)	0.2806** (0.1124)
Observations ^a	6,411	6,411	6,411	3,608	3,608
R-squared	0.289	0.263	0.297	0.245	0.249
Average profit per month in last year (ln)					
Crime: Robbery & burglary (ln)	0.1731*** (0.0232)	0.0003 (0.0669)	0.1219** (0.0504)	0.1959* (0.1004)	0.1550 (0.1007)
Observations ^a	6,448	6,448	6,448	3,570	3,570
R-squared	0.310	0.283	0.334	0.295	0.2987
Total investment last year (dummy)					
Crime: Robbery & burglary (ln)	-0.0082 (0.0074)	-0.0151 (0.0176)	-0.0080 (0.0167)	0.0032 (0.0271)	-0.0080 (0.0252)
Observations ^a	6,714	6,714	6,714	3,751	3,751
R-squared	0.106	0.107	0.111	0.120	0.122
District FE	No	Yes	Yes	Yes	Yes
Year FE	No	No	Yes	Yes	Yes
No education x crime (ln)	No	No	No	Yes	Yes
Province-specific time trends	No	No	No	No	Yes
Number of districts	39	39	39	39	39

Note: The variables male (dummy), age, age squared, white (dummy), firm age < 3 years (dummy), single owner (dummy), business motivation (dummy), debt (dummy), startup capital needed (dummy), electricity expenses and employees (categorical) are included as control variables in each of the regressions. ^a Information on education is available for 2009 and 2013 only; observations from 2005 are, therefore, dropped when including ‘No education x Crime (ln)’ in the regression (cols. 4-5). Robust standard errors clustered at district level in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

In the first part of Table 8 we present the results of the regression of sales in the last month on crime and other control variables. It can be seen that the coefficient associated with crime is positive in all five specifications, in particular in our preferred specification where we control for district fixed-effects and time effects (col. 3). The coefficient implies an increase in sales by 0.1918% for an increase in the crime rate by 1%, and is significant at the 1% level. The coefficient is also positive and significant if we introduce province-specific time trends (col. 5). As in the analysis of self-employment above, the coefficients of the control variables have, largely, the expected signs, which we also take as an indication that measurement error in

¹³ As before with the merged Census dataset, we also test whether a random-effects estimation is more suitable for our model compared to a fixed-effects estimation, by running the Hausman test and a test developed by Schaffer and Stillman (2006). The results of the Schaffer-Stillman test (see Appendix, Table A5) suggest the fixed-effects estimation to be the more appropriate specification.

reported sales is moderate. We see, for instance, that men generate higher sales than women, whites higher than blacks, older firms higher than younger firms and so on.

Turning to the regressions with average profits per month over the last year as the dependent variable (second part of Table 8), the OLS results are quite similar to the previous estimates based on sales. The coefficient associated with the crime rate is positive and significant at 1% when applying OLS, but becomes insignificant when including district fixed-effects (col. 2). However, the coefficient turns significant when we also account for time effects. If we include, in addition, province-specific time trends (col. 5), the effect of crime again becomes insignificant. The coefficients of the control variables are again largely as expected.

Finally, we use total investment as the dependent variable (third part of Table 8). Across four of the five specifications, the coefficient associated with crime is negative, but insignificant. There are no major changes with respect to the effects associated with the control variables.

To account for reverse causality and unobserved heterogeneity, we again use an instrumental variable approach. As in the previous analysis with merged Census data, we use two instruments – contact crime and sexual crime. We use the fixed-effects specifications in col. 3 of Table 8 as the benchmark specification. The instrumental variable regressions results are presented in Table 9 (for details see Online Appendix, Table B7 to Table B9).

Table 9

Impact of crime on business performance (SESE): Sales, profit and total investment – Instrumental variable regression results.

	Recap: Fixed-effects (1)	Instrumental Variable	
		Contact crime (2)	Sexual crime (3)
2 nd stage – Sales last month (ln):			
Crime: Robbery & burglary (ln)	0.1918*** (0.0511)	0.2208*** (0.0279)	0.2106*** (0.0364)
Observations	6,411	6,411	6,411
R-squared	0.297	0.320	0.317
2 nd stage – Average profit per month in last year (ln):			
Crime: Robbery & burglary (ln)	0.1219** (0.0504)	0.1988*** (0.0262)	0.1888*** (0.0339)
Observations	6,448	6,448	6,448
R-squared	0.334	0.361	0.358
2 nd stage – Total investment last year (dummy):			
Crime: Robbery & burglary (ln)	-0.0080 (0.1144)	-4.17e-5 (0.0084)	-0.0125 (0.0109)
Observations	6,714	6,714	6,714
R-squared	0.111	0.110	0.110
District FE	Yes	No	No
Year FE	Yes	Yes	Yes
Number of districts	39	39	39
1 st stage – Crime: Robbery & burglary (ln):			
Contact/Sexual crime (ln)		1.0699*** (0.0435)	1.0620*** (0.0642)
Year FE		Yes	Yes
F statistic on instrument		606	273

Note: We manually estimate the instrumental variable model. In the first stage, we use an aggregated dataset (i.e., one observation per municipality and per year) and regress the instruments on our crime rate of interest, Crime: Robbery & burglary (ln), to predict the crime variable. In the second stage, we use the disaggregated dataset and regress the business performance indicators on the predicted crime variable as well as other controls, district fixed-effects and time effects. The variables male (dummy), age, age squared, white (dummy), firm age < 3 years (dummy), single owner (dummy), business motivation (dummy), debt (dummy), startup capital needed (dummy) and electricity expenses are included as control variables in each of the regressions. Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

In the fourth part of Table 9 we show the first stage results; as before with the merged Census data, there is a strong correlation between our variable of interest and the instrumental variables. The second stage results show that, compared to the non-instrumented results, the impact of robbery and burglary on sales (first part of Table 9) slightly increases from 0.1918 (fixed-effects approach) to 0.2208 (contact crime as instrument) and 0.2106 (sexual crime as instrument). The statistical significance remains at 1% for both instruments. For the included control variables, the signs of the coefficients, their size and the significance levels remain more or less the same across the three specifications. We get similar results when using the instrumental variable model to estimate the impact of crime on average profits (second part of Table 9). Here as well the significance level is at 1% and the size of the coefficient increases when we instrument; from 0.1219 (fixed-effects) to 0.1988 (contact crime as instrument) and 0.1888 (sexual crime as instrument). Lastly, when we use the instrumental variable model to estimate the impact of robbery and burglary on total investment (third part of Table 9) we again find no impact. The coefficient of robbery and burglary remains negative and insignificant across all specifications. As before, the impact of all other included control variables is comparable to the results of our non-instrumented fixed-effects specification.

Based on these findings it seems that robbery and burglary crime rates may have no or even a slightly positive effect on sales and, to a lesser extent, on profits, but not a negative effect. To check again whether these results are sensitive to different specifications, we run various robustness checks, using the fixed-effects specifications from Table 8 col. 3 (including district fixed-effects and year effects) as our benchmark specifications. These robustness checks are shown in the Appendix, Table A6 and Table A7.

We start with the robustness checks for sales in the last month as the dependent variable. Replacing the two years aggregated robbery and burglary crime rate – as used in our benchmark specification – with a variable for robbery and burglary aggregated for the previous three years, the impact of crime remains constant. The results indicate that a 1% increase in the crime rate in the previous three years leads to an increase of sales in the last month by 0.19% – compared to 0.19% when using the two years aggregated crime rate. The significance level remains at 1%. Crime growth rates seem to have no significant impact on sales, while the levels of crime two and three years before are significant at the 1% level each, indicating that a 1% increase in these crime levels leads to a 0.19% increase in sales. Testing the sensitivity of our benchmark specification to different types of crime, we again find a positive effect when we use total crime or aggregate all property crimes (i.e., robbery and burglary, as well as theft of motor vehicles and stock). The coefficients are significant at the 1% level. Using average profit as the dependent variable, we find similar results regarding the size of the coefficients, though at a lower significance level, i.e., 5%. The robustness checks for investment again indicate that crime does not have a robust impact on total investment.

When we restrict our dataset to male entrepreneurs only, we find a significant impact on sales (at the 1% level), indicating that a 1% increase in robbery and burglary levels in the previous two

years leads to an increase of sales by 0.24%. When we exclude white entrepreneurs from our dataset, the impact of robbery and burglary on sales is significant at the 1% level and the size of the coefficient is comparable to our benchmark fixed-effects specification (around 0.19). In a final robustness check, we aggregate the data at the municipality level and re-estimate the model. The results again suggest a positive, though not significant, impact of crime on sales per month. For average profit, the coefficients are all positive, though less significant (10% for male only, 5% when excluding white, and insignificant with aggregated data). We again cannot find any significant impact of crime on investment.

Overall, the results of our analysis suggest that there are – in contrast to reported perceptions by business owners – no substantial negative effects of crime on sales, profit and investment; if anything, the impact is slightly positive for sales and average profits. In our theoretical considerations we listed a few explanations why one may see a positive rather than a negative effect, but a definitive answer would need further analysis. The results may, of course, also reflect an omitted variables bias that our specifications is not able to address.

6. Conclusion

The high crime level in South Africa has often been mentioned as one of the main reasons for the country's low share of micro and small enterprises – low compared to other countries in Sub-Saharan Africa, but also compared to countries in Latin America and Asia with similar levels of GDP per capita. Previous papers studying this relationship are typically based on small datasets, with an often limited regional scope, and mostly focus on the perceptions of crime rather than on actual crime. There are also some papers that study the relationship in a panel cross-country framework, where unobserved variables and parameter heterogeneity make identification difficult. This paper contributes to the literature by linking official crime statistics with data from two rounds of the Census and the Community Survey and data from three waves of the nationwide Survey of Employers and the Self-Employed. Using the variation of crime rates within municipalities or districts over time, while controlling for year and municipality or district fixed-effects as well as a wide range of potential confounders, we study the impact of crime on the probability of being self-employed and on business performance indicators such as sales, profits and investment. To deal with potential problems of reverse causality and omitted variable bias, we use an instrumental variable approach in addition to the fixed-effects model.

We do not find any robust evidence that the high crime rates in South Africa can explain the low share of self-employment among the South African work force. We also do not find any substantial negative effect on businesses' performance. Although we find a negative relationship between robbery and burglary crime rates and self-employment, the impact is negligibly small. These results are robust to different specifications and estimations using various data subsets. We also do not find any negative impact on formal and informal wage workers. If we estimate the impact of crime on sales, profits and investment, we likewise find no substantial negative impact; if anything, we find a positive impact of crime on sales and average profit. Although our analysis cannot completely rule out any unobserved variable bias or any bias due to measurement error, we believe the consistency of the findings across outcomes, specifications and data subsets makes it unlikely that the high crime rate is the major cause of the low share of entrepreneurship. Also recall that systematic measurement error could only drive our results if the measurement error varied within municipalities over time. More generally, our results suggest that crime may not necessarily be a serious threat for small businesses in low and middle-income countries.

Hence, future research should investigate alternative hypotheses underlying the low level of self-employment in South Africa. The distance of many townships to the business centers of towns and cities (and corresponding transportation costs) is often cited as an alternative hypothesis. Some have speculated that high reservation wages might be behind the high rate of unemployment, but empirical evidence does not support this idea (Kingdon and Knight 2004). Another explanation could be that in South Africa the formal sector is so developed that micro and small informal businesses find hardly any opportunities, i.e., compared to many other countries of a similar GDP per capita, the backward and forward linkages potentially connecting the informal sector to the formal sector might be very weak. Formal firms may simply not collaborate with informal firms. Moreover, the middle class, unlike in other countries with the same income level, may almost exclusively buy their products in the formal sector and not in the informal sector. Future research may explore whether these aspects are more relevant than crime.

Acknowledgements

The authors thank Monet Durieux, Frederick Fourie, Desirée Manamela and Caiphus Mashaba for their help in matching the different data sources used in this paper. We are also grateful for comments received at conferences at the University of Heidelberg (AEL) and the University of Göttingen, Germany, as well as at seminars at the University of Cape Town and the University of the Free State in Bloemfontein, South Africa. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- Ahwireng-Obeng, Fred, and Desmond Piaray. 1999. Institutional obstacles to South African entrepreneurship. *South African Journal of Business Management* 30, no. 3:78–85.
- Banerjee, Abhijit V., and Esther Duflo. 2011. *Poor economics: A radical rethinking of the way to fight global poverty*. 1st ed. New York: Public Affairs.
- Benyishay, Ariel, and Sarah Pearlman. 2014. Crime and microenterprise growth: Evidence from Mexico. *World Development* 56:139–152.
- Bhorat, Haroon, and Karmen Naidoo. 2015. Exploring the relationship between crime-related business insurance and informal firms' performance: A South African case study. Unpublished manuscript, Development Policy Research Unit, University of Cape Town.
- Blanc, Frédéric, Ahmed Galal, Maryse Louis, and Jean-Louis Reiffers. 2007. The challenge of employment in the Mediterranean countries. FEMISE Research Programme, Marseille.
- Blanchflower, David G. and Andrew J. Oswald. 1998. What makes an entrepreneur? *Journal of Labor Economics* 16, no. 1:26–60.
- Chandra, Vandana, Jean-Pascal Nganou, and Claire Marie-Noel. 2002. South Africa - Constraints to growth in Johannesburg's black informal sector: Evidence from the 1999 informal sector survey. Report no. 24449 – ZA, World Bank, Washington, DC.
- Cichello, Paul L., Colin Almeleh, Liberty Mncube, and Morne Oosthuizen. 2011. Perceived barriers to entry into self-employment in Khayelitsha, South Africa: Crime, risk, and start-up capital dominate profit concerns. CSSR Working Paper no. 300, Centre for Social Science Research, University of Cape Town.

- Cullen, Julie B., and Steven D. Levitt. 1999. Crime, urban flight, and the consequences for cities. *The Review of Economics and Statistics* 81, no. 2:159–169.
- De Mel, Suresh, David J. McKenzie, and Christopher Woodruff. 2008. Returns to capital in microenterprises: Evidence from a field experiment. *The Quarterly Journal of Economics* 123, no. 4:1329–1372.
- Edwards, Lawrence, and Asha Sundaram. 2013. Crime and firm dynamics: Evidence from South Africa. Unpublished manuscript, School of Economics, University of Cape Town.
- Fajnzylber, Pablo, William Maloney, and Gabriel V. Montes Rojas. 2006. Microenterprise dynamics in developing countries: How similar are they to those in the industrialized world? Evidence from Mexico. *World Bank Economic Review* 20, no. 3:389–419.
- Glaeser, Edward L., Stuart S. Rosenthal, and William C. Strange. 2010. Urban economics and entrepreneurship. *Journal of Urban Economics* 67, no. 1:1–14.
- Gould, Eric D., Bruce A. Weinberg, and David B. Mustard. 2002. Crime rates and local labor market opportunities in the United States: 1979–1997. *The Review of Economics and Statistics* 84, no. 1:45–61.
- Gough, Katherine V., A. Graham Tipple, and Mark Napier. 2003. Making a living in African cities: The role of home-based enterprises in Accra and Pretoria. *International Planning Studies* 8, no. 4:253–277.
- International Labour Organization. 2015. Key indicators of the labour market (KILM), 9th ed. http://www.ilo.org/ilostat/faces/help_home/data_by_subject (last accessed September 29, 2016).
- Islam, Asif. 2014. Economic growth and crime against small and medium sized enterprises in developing economies. *Small Business Economics* 43, no. 3:677–695.
- Kahn, Matthew E. 2010. New evidence on the trends in the cost of urban agglomeration. In *Agglomeration Economics*, ed. Edward L. Glaeser, 339–354. Chicago: The University of Chicago Press.
- Kingdon, Geeta G., and John Knight. 2004. Unemployment in South Africa: The nature of the beast. *World Development* 32, no. 3:391–408.
- Kock, Chris de, Anine Kriegler, and Mark Shaw. 2015. A citizen's guide to SAPS crime statistics: 1994 to 2015. http://www.criminology.uct.ac.za/sites/default/files/image_tool/images/225/CRI_pdf_Optimised-Citizen's%20guide%20to%20the%20SAPS%20stats.pdf (last accessed September 29, 2016).
- Lens, Michael C., and Rachel Meltzer. 2016. Is crime bad for business? Crime and commercial property values in New York City. *Journal of Regional Science* 56, no. 3:442–470.
- Machin, Stephen, and Costas Meghir. 2004. Crime and economic incentives. *The Journal of Human Resources* 39, no. 4:958–979.
- Mahadea, Darma, and Morgan Pillay. 2008. Environmental conditions for SMME development in a South African province. *South African Journal of Economic and Management Science* 11, no. 4:431–448.
- McDonald, Kerri. 2008. The impact of crime on small businesses in South Africa: A study commissioned by the South African presidency. SBP, Johannesburg.

- McKenzie, David J. 2010. Impact assessments in finance and private sector development: What have we learned and what should we learn? *World Bank Research Observer* 25, no. 2:209–233.
- McKenzie, David J., and Christopher Woodruff. 2006. Do entry costs provide an empirical basis for poverty traps? Evidence from Mexican microenterprises. *Economic Development and Cultural Change* 55, no. 1:3–42.
- McPherson, Michael A. 1996. Growth of micro and small enterprises in southern Africa. *Journal of Development Economics* 48, no. 2:253–277.
- Olawale, Fatoki, and David Garwe. 2010. Obstacles to the growth of new SMEs in South Africa: A principal component analysis approach. *African Journal of Business Management* 4, no. 5:729–738.
- Rosenthal, Stuart S., and Amanda Ross. 2010. Violent crime, entrepreneurship, and cities. *Journal of Urban Economics* 67, no. 1:135–149.
- Schaffer, Mark E., and Steven Stillman. 2006. xtoverid: Stata module to calculate tests of overidentifying restrictions after xtreg, xtivreg, xtivreg2 and xthtaylor. <http://ideas.repec.org/c/boc/bocode/s456779.html> (last accessed September 29, 2016).
- Skinner, Caroline. 2005. Constraints to growth and employment in Durban: Evidence from the informal economy. Research Report no. 65, School of Development Studies, University of KwaZulu-Natal, Durban.
- South African Police Service. 2015. Crime statistics. http://www.saps.gov.za/resource_centre/publications/statistics/crimestats/2015/crime_stats.php (last accessed September 29, 2016).
- Statistics South Africa. 2003. Victims of crime survey 2003. <https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/178> (last accessed September 29, 2016).
- Statistics South Africa. 2007a. Community survey 2007. <https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/101> (last accessed September 29, 2016).
- Statistics South Africa. 2007b. Victims of crime survey 2007. <https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/179> (last accessed September 29, 2016).
- Statistics South Africa. 2011a. Census 2011 geography metadata. <http://microdata.worldbank.org/index.php/catalog/2067/download/31411> (last accessed September 29, 2016).
- Statistics South Africa. 2011b. South African census 2011, 10% Sample. <https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/485> (last accessed September 29, 2016).
- Statistics South Africa. 2011c. Victims of crime survey 2011. <https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/176> (last accessed September 29, 2016).
- Statistics South Africa. 2012. Victims of crime survey 2012. <https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/414> (last accessed September 29, 2016).

- Statistics South Africa. 2014a. Victims of crime survey 2013/14: Statistical release.
<http://www.statssa.gov.za/publications/P0341/P03412013.pdf> (last accessed September 29, 2016).
- Statistics South Africa. 2014b. Victims of crime survey 2013-14.
<https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/511> (last accessed September 29, 2016).
- Statistics South Africa. 2015a. Quarterly labour force survey: Quarter 3: 2015.
<http://www.statssa.gov.za/publications/P0211/P02113rdQuarter2015.pdf> (last accessed September 29, 2016).
- Statistics South Africa. 2015b. Survey of employers and the self-employed (SESE): Series 2001-2013. <https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/514> (last accessed September 29, 2016).
- Statistics South Africa. 2016. Quarterly labour force survey (QLFS): Trends 2008-2016.
http://www.statssa.gov.za/publications/P0211/QLFS_Trends_2008-2016.xlsx (last accessed September 29, 2016).
- Stock, James H., and Motohiro Yogo. 2005. Testing for weak instruments in linear IV regression. In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, ed. Donald W. K. Andrews and James H. Stock, 80–108. Cambridge: Cambridge University Press.
- Tita, George E., Tricia L. Petras, and Robert T. Greenbaum. 2006. Crime and residential choice: A neighborhood level analysis of the impact of crime on housing prices. *Journal of Quantitative Criminology* 22, no. 4:299–317.
- UNODC. 2015. Crime and criminal justice statistics: Homicide counts and rates (2000-2014). <https://data.unodc.org/#state:0> (last accessed September 29, 2016).
- World Bank. 2015. World development indicators.
<http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators> (last accessed September 29, 2016).

Appendix (A)

Availability of geographical information in data			Geographical/administrative levels in South Africa	
Crime Stats	Merged Census	SESE		
X	X	X		SA
X	X	X		Province
	X	X		District ^a
	X	X		Municipality ^a
X				Police station jurisdiction
		X		Main place
		X		Sub place

Fig. A1. Overview of geographical level that can be identified in main datasets.
Source: Statistics South Africa (2011a). ^a Including metro municipalities.

Table A1

Impact of crime on self-employment (merged Census) – Test of suitability of fixed-effects vs. random-effects model specification.

Specification from Table 6	Results of fixed-effects vs. random-effects tests			
	(2)	(3)	(4)	(5)
Hausman test - no robust standard errors				
Hausman statistic	65.46	84.01	86.96	49.52
P-value	0.0000	0.0000	0.0000	0.4523
Schaffer-Stillman test - robust standard errors				
Sargan-Hansen statistic	65.67	66.79	82.09	11,000.00
P-value	0.0000	0.0000	0.0000	0.0000

Note: In each of the regressions the variable self-employment (dummy) is used as the dependent variable and the variables crime: robbery & burglary (ln), male (dummy), age, age squared, average income (ln), white (dummy) and no education (dummy) are included as explanatory or control variables. In cols. 4 and 5 we additionally included an interaction of no education and crime. In col. 2 we included municipality fixed-effects only, while in cols. 3 and 4 we added time effects, and in col. 5 additionally district-specific time trends.

Table A2

Impact of crime on self-employment (merged Census) – Robustness checks with full dataset.

Dependent variable	Self-employment (dummy)				
	Full dataset				
Crime type	Robbery and burglary			Total crime	Property crime
Crime calculation	3 years aggregated	2 years growth	3 years growth	2 years aggregated	
	(1)	(2)	(3)	(4)	(5)
Crime rate – Aggregated (ln)	-0.0126* (0.0068)			-0.0099 (0.0094)	-0.0127 (0.0084)
Crime rate – Growth		-0.0161 (0.0144)	-0.0239 (0.0206)		
Crime rate – Level before (ln)		-0.0083 (0.0066)	-0.0153** (0.0068)		
Municipality FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Observations	2,516,265	2,516,265	2,516,265	2,516,265	2,516,265
Number of municipalities	234	234	234	234	234
R-squared	0.022	0.022	0.022	0.022	0.022

Note: The variables male (dummy), age, age squared, average income (ln), white (dummy) and no education (dummy) are included as control variables in each of the regressions. Robust standard errors clustered at municipality level in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

Table A3

Impact of crime on self-employment (merged Census) – Robustness checks with subsets of data and estimation with aggregated data.

Dependent variable	Self-employment (dummy) ^a						
Dataset	Excl. formally employed	Informal self-empl.	Urban only ^b	Perceived impact of crime ^c	Male only	Excl. white	Municipality aggreg.
Crime type	Robbery and burglary						
Crime calculation	2 years aggregated						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Crime rate – Aggreg. (ln)	0.0060 (0.0077)	0.0011 (0.0060)	-0.0253 (0.0403)	0.0804*** (0.0115)	-0.0058 (0.0080)	-0.0126** (0.0064)	0.0055 (0.0067)
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,657,672	2,410,994	1,166,299	776,436	1,196,403	2,294,616	468
Number of municipalities	234	234	14	4	234	234	234
R-squared	0.018	0.006	0.023	0.027	0.024	0.008	0.106

Note: The variables male (dummy), age, age squared, average income (ln), white (dummy) and no education (dummy) are included as control variables in each of the regressions. Robust standard errors clustered at municipality level in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

^a For the regression in col. 2, the self-employment dummy is defined as the share of all individuals who are employers or own-account workers in the informal sector in relation to the total labor force (excluding individuals employed in the formal sector). In cols. 1 and 3 to 7 the self-employment share consists of all self-employed in the informal and the formal sectors (total labor force also including individuals employed in the formal sector). ^b Urban is defined as a municipality with more than 400 residents per square kilometer. ^c Only municipalities considered in which previous surveys found a negative impact of the perceptions of crime on self-employment. These are Johannesburg (Chandra et al. 2002; McDonald 2008) incl. Diepsloot township (Bhorat and Naidoo 2015), Cape Town (McDonald 2008) incl. the township of Khayelitsha (McDonald 2008; Cichello et al. 2011), Durban/eThekweni (Skinner 2005; McDonald 2008) and Pretoria/City of Tshwane (Gough et al. 2003).

Table A4

Impact of crime on self-employment (merged Census) – Variation of impact for different crime levels (quintiles).

Dependent variable Crime level (quintiles) ^a	Self-employment (dummy)				
	1 st quintile (1)	2 nd quintile (2)	3 rd quintile (3)	4 th quintile (4)	5 th quintile (5)
Crime: Robbery & burglary (ln)	0.0113 (0.0117)	-0.0017 (0.0189)	0.0177 (0.0114)	-0.0109 (0.0155)	0.0075 (0.0115)
Municipality FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Observations	241,822	314,588	231,740	489,058	1,239,057
Number of municipalities	47	47	47	47	46
R-squared	0.015	0.020	0.028	0.017	0.025

Note: The variables male (dummy), age, age squared, average income (ln), white (dummy) and no education (dummy) are included as control variables in each of the regressions. Robust standard errors clustered at municipality level in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

^a The municipalities are allocated into quintiles according to the aggregated crime rate for the two years previous to 2007. The 1st quintile includes the 20% of municipalities with the lowest crime rate, while the 5th quintile the 20% of municipalities with the highest crime rate. As before, the crime rate for burglary and robbery in residential and non-residential premises is used. The number of observations per quintile varies due to the different number of observations per municipality (particularly rural areas with high crime rates with a large number of observations).

Table A5

Impact of crime on business performance (SESE): Sales, profit and total investment – Test of suitability of fixed-effects vs. random-effects model specification.

Specification from Table 8	Results of fixed-effects vs. random-effects tests			
	(2)	(3)	(4)	(5)
Sales last month (ln):				
Hausman test - no robust standard errors				
Hausman statistic	6.93	11.62	12.03	22.96
P-value	0.7322	0.4764	0.5253	0.3461
Schaffer-Stillman test - robust standard errors				
Sargan-Hansen statistic	18.80	11.88	50.55	249.19
P-value	0.0936	0.5380	0.0000	0.0000
Average profit per month in last year (ln):				
Hausman test - no robust standard errors				
Hausman statistic	17.50	12.67	26.79	22.71
P-value	0.0641	0.3932	0.0133	0.3599
Schaffer-Stillman test - robust standard errors				
Sargan-Hansen statistic	28.38	20.84	113.36	477.46
P-value	0.0049	0.0761	0.0000	0.0000
Total investment last year (dummy):				
Hausman test - no robust standard errors				
Hausman statistic	10.84	11.76	29.51	41.83
P-value	0.3705	0.4655	0.0055	0.0044
Schaffer-Stillman test - robust standard errors				
Sargan-Hansen statistic	33.89	44.70	101.47	838.99
P-value	0.0007	0.0000	0.0000	0.0000

Note: The variables crime: robbery & burglary (ln), male (dummy), age, age squared, white (dummy), firm age < 3 years (dummy), single owner (dummy), business motivation (dummy), debt (dummy), startup capital needed (dummy), electricity expenses and employees (categorical) are included as explanatory or control variables in each of the regressions. In cols. 4 and 5 we additionally included an interaction of no education and crime. In col. 2 we included municipality fixed-effects only, while in cols. 3 and 4 we added time effects, and in col. 5 additionally province-specific time trends.

Table A6

Impact of crime on business performance (SESE): Sales, profit and total investment – Robustness checks with full dataset.

Dataset	Full dataset				
	Robbery and burglary			Total crime	Property crime
	3 years aggregated	2 years growth	3 years growth	2 years aggregated	
	(1)	(2)	(3)	(4)	(5)
Sales last month (ln):					
Crime rate – Aggregated (ln)	0.1949*** (0.0513)			0.1950*** (0.0547)	0.2026*** (0.0510)
Crime rate – Growth		-0.0801 (0.3685)	-0.1794 (0.3920)		
Crime rate – Level before (ln)		0.1860*** (0.0533)	0.1887*** (0.0543)		
Observations	6,411	6,411	6,411	6,411	6,411
R-squared	0.297	0.297	0.298	0.297	0.298
Average profit per month in last year (ln):					
Crime rate – Aggregated (ln)	0.1288** (0.0501)			0.1046** (0.0504)	0.1086** (0.0489)
Crime rate – Growth		-0.3062 (0.3639)	-0.4131 (0.4180)		
Crime rate – Level before (ln)		0.1186** (0.0528)	0.1215** (0.0514)		
Observations	6,448	6,448	6,448	6,448	6,448
R-squared	0.334	0.335	0.335	0.334	0.334
Total investment last year (dummy):					
Crime rate – Aggregated (ln)	-0.0086 (0.0165)			-0.0034 (0.0188)	-0.0158 (0.0157)
Crime rate – Growth		-0.0458 (0.1081)	0.1164 (0.1247)		
Crime rate – Level before (ln)		-0.0080 (0.0177)	-0.0077 (0.0173)		
Observations	6,714	6,714	6,714	6,714	6,714
R-squared	0.111	0.111	0.111	0.111	0.111

Note: The control variables male (dummy), age, age squared, white (dummy), firm age < 3 years (dummy), single owner (dummy), business motivation (dummy), debt (dummy), startup capital needed (dummy), electricity expenses and employees (categorical), as well as district and year fixed-effects are included in each of the regressions. Robust standard errors clustered at district level in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

Table A7

Impact of crime on business performance (SESE): Sales, profit and total investment – Robustness checks with subsets of data and estimation with aggregated data.

Dataset	Male only	Excluding white	Municipality aggregation
Crime type	Robbery and burglary		
Crime calculation	2 years aggregated		
	(1)	(2)	(3)
Sales last month (ln):			
Crime rate – Aggregated (ln)	0.2414*** (0.0842)	0.1954*** (0.0512)	0.0698 (0.0884)
Observations	2,770	6,243	421
R-squared	0.282	0.276	0.477
Average profit per month in last year (ln):			
Crime rate – Aggregated (ln)	0.1021* (0.0592)	0.1171** (0.0513)	0.0340 (0.0644)
Observations	2,789	6,271	422
R-squared	0.316	0.315	0.506
Total investment last year (dummy):			
Crime rate – Aggregated (ln)	-0.0195 (0.0215)	-0.0090 (0.0164)	-0.0384 (0.0260)
Observations	2,926	6,520	424
R-squared	0.095	0.094	0.160

Note: The control variables male (dummy), age, age squared, white (dummy), firm age < 3 years (dummy), single owner (dummy), business motivation (dummy), debt (dummy), startup capital needed (dummy), electricity expenses and employees (categorical), as well as district and year fixed-effects are included in each of the regressions. Robust standard errors clustered at district level in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

Online Appendix (B)
Supplemental Appendix

Underreporting
 (share of robbery incidents not reported),
 between 2010 and 2014

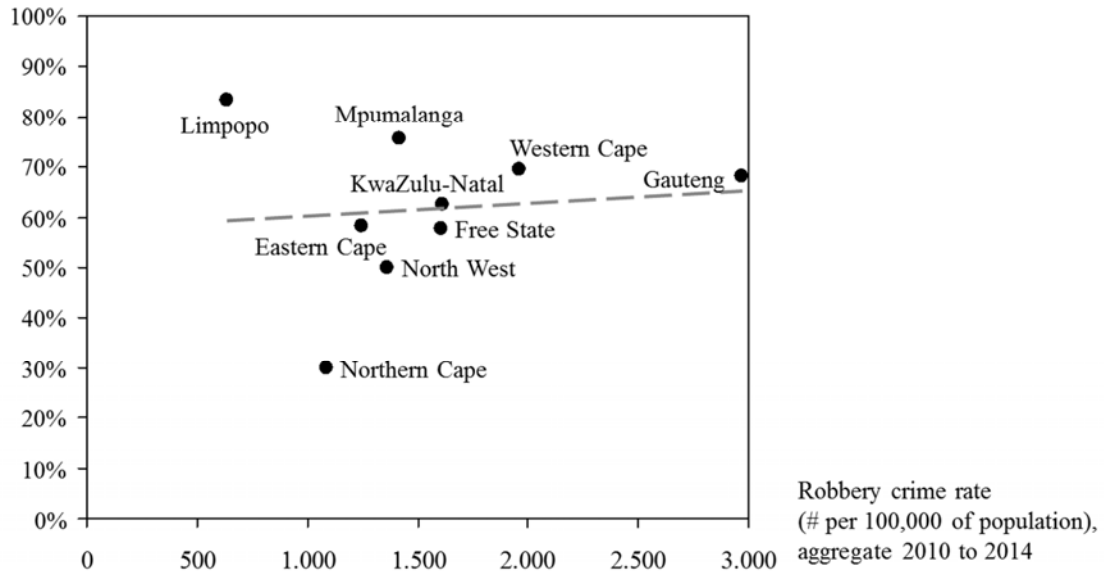


Fig. B1. Relationship between the share of underreporting and the robbery crime rate at province level, 2010 to 2014.

Sources: Crime Stats (South African Police Service 2015) and Victims of Crime Survey (Statistics South Africa 2014b).

Table B1

Impact of crime on self-employment (merged Census) – OLS and fixed-effects regression results.

Dependent variable	Self-employment (dummy)				
	OLS	Fixed-effects			
	(1)	(2)	(3)	(4)	(5)
Crime: Robbery & burglary (ln)	-0.0066*** (0.0004)	-0.0135** (0.0066)	-0.0128* (0.0065)	-0.0146** (0.0066)	-0.0062 (0.0065)
Male (dummy)	0.0275*** (0.0003)	0.0277*** (0.0015)	0.0277*** (0.0015)	0.0277*** (0.0014)	0.0278*** (0.0014)
Age	0.0081*** (0.0001)	0.0080*** (0.0003)	0.0080*** (0.0003)	0.0080*** (0.0003)	0.0080*** (0.0003)
Age squared	-8.97e-5*** (0.0000)	-8.84e-5*** (0.0000)	-8.83e-5*** (0.0000)	-8.83e-5*** (0.0000)	-8.83e-5*** (0.0000)
Average income (ln)	0.0175*** (0.0004)	0.0191*** (0.0032)	0.0270*** (0.0062)	0.0271*** (0.0062)	0.0183*** (0.0059)
White (dummy)	0.0936*** (0.0008)	0.0954*** (0.0065)	0.0954*** (0.0065)	0.0957*** (0.0065)	0.0958*** (0.0065)
No education (dummy)	-0.0149*** (0.0004)	-0.0149*** (0.0013)	-0.0150*** (0.0013)	-0.0658*** (0.0124)	-0.0687*** (0.0124)
No education x Crime (ln)				0.0069*** (0.0017)	0.0072*** (0.0017)
Constant	-0.2413*** (0.0028)	-0.2041*** (0.0674)	-0.2874*** (0.0853)	-0.2742*** (0.0852)	-0.2742*** (0.0852)
Municipality FE	No	Yes	Yes	Yes	Yes
Year FE	No	No	Yes	Yes	Yes
District-specific time trends	No	No	No	No	Yes
Observations	2,516,265	2,516,265	2,516,265	2,516,265	2,516,265
Number of municipalities	234	234	234	234	234
R-squared	0.025	0.022	0.022	0.022	0.022

Note: Robust standard errors clustered at municipality level in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

Table B2

Impact of crime on self-employment (merged Census) – Instrumental variable 2nd stage regression results.

Dependent variable	Self-employment (dummy)		
	Recap: Fixed-effects	Instrumental variable: 2 nd stage	
		Contact crime	Sexual crime
	(1)	(2)	(3)
Crime: Robbery & burglary (ln)	-0.0128* (0.0065)	-0.0088*** (0.0005)	-0.0114*** (0.0005)
Male (dummy)	0.0277*** (0.0015)	0.0275*** (0.0003)	0.0276*** (0.0003)
Age	0.0080*** (0.0003)	0.0081*** (0.0001)	0.0081*** (0.0001)
Age squared	-8.83e-05*** (0.0000)	-0.0001*** (0.0000)	-0.0001*** (0.0000)
Average income (ln)	0.0270*** (0.0062)	0.0187*** (0.0004)	0.0195*** (0.0004)
White (dummy)	0.0954*** (0.0065)	0.0936*** (0.0008)	0.0937*** (0.0008)
No education (dummy)	-0.0150*** (0.0013)	-0.0149*** (0.0004)	-0.0150*** (0.0004)
Constant	-0.2874*** (0.0853)	-0.2381*** (0.0028)	-0.2253*** (0.0029)
Municipality FE	Yes	No	No
Year FE	Yes	Yes	Yes
Observations	2,516,265	2,516,265	2,516,265
Number of municipalities	234	234	234
R-squared	0.022	0.0249	0.0250

Note: Robust standard errors clustered at municipality level in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

Table B3

Impact of crime on self-employment, formal employment and unemployment – Multinomial logit regressions results.

Dependent variable (Employment status) – categories:	Multinomial Logit		
	Self-employment	Formal employment	Unemployment
	(1)	(2)	(3)
Crime: Robbery & burglary (ln)	-0.0477*** (0.0108)	0.0888*** (0.0068)	-0.0343*** (0.0062)
Year FE	Yes	Yes	Yes
Observations	2,516,265	2,516,265	2,516,265
R-squared	0.083	0.083	0.083

Note: The variables male (dummy), age, age squared, average income (ln), white (dummy) and no education (dummy) are included as control variables in the regression. Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

Table B4

Impact of crime on business performance (SESE): Sales – OLS and fixed-effects regression results.

Dependent variable	Sales last month (ln)				
	OLS	Fixed-effects			
	(1)	(2)	(3)	(4)	(5)
Crime: Robbery & burglary (ln)	0.1920*** (0.0243)	0.0914 (0.0616)	0.1918*** (0.0511)	0.3006** (0.1144)	0.2806** (0.1124)
Male (dummy)	0.4913*** (0.0329)	0.4816*** (0.0320)	0.4547*** (0.0293)	0.4106*** (0.0395)	0.4099*** (0.0389)
Age	0.0495*** (0.0066)	0.0488*** (0.0059)	0.0435*** (0.0056)	0.0355*** (0.0101)	0.0347*** (0.0101)
Age squared	-0.0006*** (0.0001)	-0.0006*** (0.0001)	-0.0006*** (0.0001)	-0.0005*** (0.0001)	-0.0005*** (0.0001)
White (dummy)	1.1745*** (0.1129)	1.1304*** (0.1325)	1.1494*** (0.1370)	1.1098*** (0.1792)	1.1106*** (0.1817)
Firm age < 3 years (dummy)	-0.3802*** (0.0312)	-0.3798*** (0.0322)	-0.3476*** (0.0310)	-0.2730*** (0.0497)	-0.2693*** (0.0495)
Single owner (dummy)	-0.2064*** (0.0517)	-0.1723** (0.0822)	-0.0950 (0.0779)	0.0048 (0.0764)	-0.0005 (0.0766)
Business motivation (dummy)	0.2024*** (0.0413)	0.2010*** (0.0421)	0.1693*** (0.0419)	0.1592*** (0.0554)	0.1581*** (0.0532)
Debt (dummy)	0.4783*** (0.0888)	0.4632*** (0.1285)	0.4678*** (0.1259)	0.4433** (0.1951)	0.4420** (0.1907)
Startup capital needed (dummy)	0.1313*** (0.0335)	0.1337*** (0.0417)	0.2138*** (0.0375)	0.1456*** (0.0506)	0.1419*** (0.0509)
Electricity expenses	0.0006*** (0.0001)	0.0006*** (0.0001)	0.0005*** (0.0001)	0.0004*** (0.0001)	0.0004*** (0.0001)
Employees (categorical)	0.6672*** (0.0288)	0.6658*** (0.0313)	0.6524*** (0.0297)	0.6222*** (0.0318)	0.6203*** (0.0320)
No education (dummy)				1.1927** (0.5380)	1.2372** (0.5407)
No education x Crime (ln)				-0.1886** (0.0772)	-0.1947** (0.0774)
Constant	4.2003*** (0.2448)	4.9240*** (0.4883)	3.8887*** (0.4239)	3.6851*** (0.8462)	3.8711*** (0.8349)
District FE	No	Yes	Yes	Yes	Yes
Year FE	No	No	Yes	Yes	Yes
Province-specific time trends	No	No	No	No	Yes
Observations ^a	6,411	6,411	6,411	3,608	3,608
Number of districts	39	39	39	39	39
R-squared	0.289	0.263	0.297	0.245	0.249

Note: Robust standard errors clustered at district level in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

^a Information on education is available for 2009 and 2013 only; observations from 2005 are, therefore, dropped when including 'No education x Crime (ln)' in the regression (cols. 4-5).

Table B5

Impact of crime on business performance (SESE): Profit – OLS and fixed-effects regression results.

Dependent variable	Average profit per month in last year (ln)				
	OLS	Fixed-effects			
	(1)	(2)	(3)	(4)	(5)
Crime: Robbery & burglary (ln)	0.1731*** (0.0232)	0.0003 (0.0669)	0.1219** (0.0504)	0.1959* (0.1004)	0.1550 (0.1007)
Male (dummy)	0.5545*** (0.0313)	0.5439*** (0.0283)	0.5135*** (0.0258)	0.4986*** (0.0343)	0.5035*** (0.0332)
Age	0.0416*** (0.0063)	0.0408*** (0.0060)	0.0345*** (0.0061)	0.0293*** (0.0091)	0.0288*** (0.0090)
Age squared	-0.0005*** (0.0001)	-0.0005*** (0.0001)	-0.0005*** (0.0001)	-0.0004*** (0.0001)	-0.0004*** (0.0001)
White (dummy)	1.0688*** (0.1027)	1.0087*** (0.0601)	1.0227*** (0.0560)	0.9184*** (0.0941)	0.9342*** (0.0976)
Firm age < 3 years (dummy)	-0.4668*** (0.0300)	-0.4579*** (0.0318)	-0.4235*** (0.0306)	-0.3719*** (0.0385)	-0.3705*** (0.0382)
Single owner (dummy)	-0.2721*** (0.0485)	-0.2259** (0.0904)	-0.1691** (0.0803)	-0.0618 (0.0694)	-0.0164 (0.0660)
Business motivation (dummy)	0.2374*** (0.0385)	0.2281*** (0.0471)	0.1911*** (0.0467)	0.1613** (0.0613)	0.1584** (0.0605)
Debt (dummy)	0.3871*** (0.0732)	0.3699*** (0.0788)	0.3654*** (0.0710)	0.4286*** (0.1040)	0.4270*** (0.1035)
Startup capital needed (dummy)	-0.0598* (0.0310)	-0.0614 (0.0453)	0.0351 (0.0373)	-0.0107 (0.0374)	-0.0096 (0.0380)
Electricity expenses	0.0006*** (0.0001)	0.0006*** (0.0001)	0.0005*** (0.0001)	0.0005*** (0.0001)	0.0005*** (0.0001)
Employees (categorical)	0.6177*** (0.0254)	0.6180*** (0.0289)	0.6029*** (0.0264)	0.5658*** (0.0248)	0.5667*** (0.0248)
No education (dummy)				2.0017*** (0.5906)	1.9271*** (0.5967)
No education x Crime (ln)				-0.3006*** (0.0844)	-0.2905*** (0.0850)
Constant	4.3092*** (0.2328)	5.5536*** (0.5262)	4.3649*** (0.4151)	4.3380*** (0.7182)	4.6427*** (0.7199)
District FE	No	Yes	Yes	Yes	Yes
Year FE	No	No	Yes	Yes	Yes
Province-specific time trends	No	No	No	No	Yes
Observations ^a	6,448	6,448	6,448	3,570	3,570
Number of districts	39	39	39	39	39
R-squared	0.310	0.283	0.334	0.295	0.2987

Note: Robust standard errors clustered at district level in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

^a Information on education is available for 2009 and 2013 only; observations from 2005 are, therefore, dropped when including 'No education x Crime (ln)' in the regression (cols. 4-5).

Table B6

Impact of crime on business performance (SESE): Total investment – OLS and fixed-effects regression results.

Dependent variable	Total investment last year (dummy)				
	OLS		Fixed-effects		
	(1)	(2)	(3)	(4)	(5)
Crime: Robbery & burglary (ln)	-0.0082 (0.0074)	-0.0151 (0.0176)	-0.0080 (0.0167)	0.0032 (0.0271)	-0.0080 (0.0252)
Male (dummy)	0.1079*** (0.0097)	0.1086*** (0.0124)	0.1065*** (0.0124)	0.1264*** (0.0176)	0.1270*** (0.0179)
Age	0.0037** (0.0018)	0.0041** (0.0017)	0.0038** (0.0017)	0.0039 (0.0025)	0.0037 (0.0025)
Age squared	-0.0001*** (0.0000)	-0.0001*** (0.0000)	-0.0001*** (0.0000)	-0.0001** (0.0000)	-0.0001** (0.0000)
White (dummy)	0.1975*** (0.0353)	0.2037*** (0.0357)	0.2064*** (0.0357)	0.1940*** (0.0404)	0.1944*** (0.0397)
Firm age < 3 years (dummy)	-0.0323*** (0.0091)	-0.0327*** (0.0113)	-0.0310*** (0.0112)	-0.0369** (0.0151)	-0.0362** (0.0150)
Single owner (dummy)	0.0065 (0.0149)	0.0189 (0.0187)	0.0140 (0.0196)	0.0329 (0.0326)	0.0330 (0.0308)
Business motivation (dummy)	0.0452*** (0.0127)	0.0493** (0.0195)	0.0463** (0.0197)	0.0607*** (0.0216)	0.0594*** (0.0214)
Debt (dummy)	0.0841*** (0.0259)	0.0871*** (0.0269)	0.0879*** (0.0265)	0.1043*** (0.0372)	0.1022*** (0.0369)
Startup capital needed (dummy)	0.0055 (0.0099)	0.0117 (0.0127)	0.0175 (0.0108)	0.0105 (0.0182)	0.0110 (0.0177)
Electricity expenses	0.0000** (0.0000)	0.0000* (0.0000)	0.0000 (0.0000)	0.0000* (0.0000)	0.0000* (0.0000)
Employees (categorical)	0.1209*** (0.0102)	0.1233*** (0.0133)	0.1220*** (0.0133)	0.1196*** (0.0144)	0.1189*** (0.0143)
No education (dummy)				0.1442 (0.2187)	0.1592 (0.2135)
No education x Crime (ln)				-0.0189 (0.0305)	-0.0210 (0.0299)
Constant	0.1107 (0.0714)	0.1351 (0.1352)	0.0746 (0.1295)	-0.0211 (0.1989)	0.0568 (0.1872)
District FE	No	Yes	Yes	Yes	Yes
Year FE	No	No	Yes	Yes	Yes
Province-specific time trends	No	No	No	No	Yes
Observations ^a	6,714	6,714	6,714	3,751	3,751
Number of districts	39	39	39	39	39
R-squared	0.106	0.107	0.111	0.120	0.122

Note: Robust standard errors clustered at district level in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

^a Information on education is available for 2009 and 2013 only; observations from 2005 are, therefore, dropped when including 'No education x Crime (ln)' in the regression (cols. 4-5).

Table B7

Impact of crime on business performance (SESE): Sales – Instrumental variable 2nd stage regression results.

Dependent variable	Sales per month (ln)		
	Recap: Fixed-effects	Instrumental variable: 2 nd stage	
		Contact crime	Sexual crime
	(1)	(2)	(3)
Crime: Robbery & burglary (ln)	0.1918*** (0.0511)	0.2208*** (0.0279)	0.2106*** (0.0364)
Male (dummy)	0.4547*** (0.0293)	0.4661*** (0.0322)	0.4770*** (0.0322)
Age	0.0435*** (0.0056)	0.0449*** (0.0065)	0.0450*** (0.0065)
Age squared	-0.0006*** (0.0001)	-0.0006*** (0.0001)	-0.0006*** (0.0001)
White (dummy)	1.1494*** (0.1370)	1.1930*** (0.1103)	1.2232*** (0.1115)
Firm age < 3 years (dummy)	-0.3476*** (0.0310)	-0.3420*** (0.0305)	-0.3399*** (0.0304)
Single owner (dummy)	-0.0950 (0.0779)	-0.1394** (0.0542)	-0.1529*** (0.0541)
Business motivation (dummy)	0.1693*** (0.0419)	0.1682*** (0.0406)	0.1676*** (0.0407)
Debt (dummy)	0.4678*** (0.1259)	0.4766*** (0.0867)	0.4794*** (0.0862)
Startup capital needed (dummy)	0.2138*** (0.0375)	0.2063*** (0.0329)	0.1994*** (0.0329)
Electricity expenses	0.0005*** (0.0001)	0.0006*** (0.0001)	0.0006*** (0.0001)
Employees (categorical)	0.6524*** (0.0297)	0.6513*** (0.0284)	0.6522*** (0.0285)
Constant	3.8887*** (0.4239)	3.7142*** (0.2626)	3.8121*** (0.3105)
District FE	Yes	No	No
Year FE	Yes	Yes	Yes
Observations	6,411	6,411	6,411
Number of districts	39	39	39
R-squared	0.297	0.320	0.317

Note: Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

Table B8

Impact of crime on business performance (SESE): Profit – Instrumental variable 2nd stage regression results.

Dependent variable	Average profit per month in last year (ln)		
	Recap: Fixed-effects	Instrumental variable: 2 nd stage	
		Contact crime	Sexual crime
	(1)	(2)	(3)
Crime: Robbery & burglary (ln)	0.1219** (0.0504)	0.1988*** (0.0262)	0.1888*** (0.0339)
Male (dummy)	0.5135*** (0.0258)	0.5248*** (0.0300)	0.5343*** (0.0300)
Age	0.0345*** (0.0061)	0.0359*** (0.0061)	0.0360*** (0.0062)
Age squared	-0.0005*** (0.0001)	-0.0005*** (0.0001)	-0.0005*** (0.0001)
White (dummy)	1.0227*** (0.0560)	1.0861*** (0.0969)	1.1126*** (0.0973)
Firm age < 3 years (dummy)	-0.4235*** (0.0306)	-0.4224*** (0.0289)	-0.4199*** (0.0289)
Single owner (dummy)	-0.1691** (0.0803)	-0.2151*** (0.0498)	-0.2272*** (0.0498)
Business motivation (dummy)	0.1911*** (0.0467)	0.1959*** (0.0373)	0.1960*** (0.0374)
Debt (dummy)	0.3654*** (0.0710)	0.3775*** (0.0687)	0.3812*** (0.0684)
Startup capital needed (dummy)	0.0351 (0.0373)	0.0351 (0.0300)	0.0290 (0.0300)
Electricity expenses	0.0005*** (0.0001)	0.0005*** (0.0001)	0.0005*** (0.0001)
Employees (categorical)	0.6029*** (0.0264)	0.6009*** (0.0247)	0.6020*** (0.0247)
Constant	4.3649*** (0.4151)	3.8227*** (0.2471)	3.9167*** (0.2897)
District FE	Yes	No	No
Year FE	Yes	Yes	Yes
Observations	6,448	6,448	6,448
Number of districts	39	39	39
R-squared	0.334	0.361	0.358

Note: Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

Table B9

Impact of crime on business performance (SESE): Total investment – Instrumental variable 2nd stage regression results.

Dependent variable	Total investment last year (dummy)		
	Recap: Fixed-effects	Instrumental variable: 2 nd stage	
		Contact crime	Sexual crime
	(1)	(2)	(3)
Crime: Robbery & burglary (ln)	-0.0080 (0.0167)	-0.0000 (0.0084)	-0.0125 (0.0109)
Male (dummy)	0.1065*** (0.0124)	0.1045*** (0.0097)	0.1052*** (0.0097)
Age	0.0038** (0.0017)	0.0034* (0.0018)	0.0034* (0.0018)
Age squared	-0.0001*** (0.0000)	-0.0001*** (0.0000)	-0.0001*** (0.0000)
White (dummy)	0.2064*** (0.0357)	0.1968*** (0.0355)	0.1993*** (0.0355)
Firm age < 3 years (dummy)	-0.0310*** (0.0112)	-0.0301*** (0.0091)	-0.0300*** (0.0091)
Single owner (dummy)	0.0140 (0.0196)	0.0046 (0.0155)	0.0039 (0.0156)
Business motivation (dummy)	0.0463** (0.0197)	0.0415*** (0.0127)	0.0419*** (0.0127)
Debt (dummy)	0.0879*** (0.0265)	0.0848*** (0.0259)	0.0849*** (0.0259)
Startup capital needed (dummy)	0.0175 (0.0108)	0.0122 (0.0099)	0.0121 (0.0098)
Electricity expenses	0.0000 (0.0000)	0.0000* (0.0000)	0.0000* (0.0000)
Employees (categorical)	0.1220*** (0.0133)	0.1194*** (0.0102)	0.1197*** (0.0102)
Constant	0.0746 (0.1295)	0.0386 (0.0765)	0.1284 (0.0902)
District FE	Yes	No	No
Year FE	Yes	Yes	Yes
Observations	6,714	6,714	6,714
Number of districts	39	39	39
R-squared	0.111	0.110	0.110

Note: Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.