UNIVERSITÄT PASSAU
Wirtschaftswissenschaftliche Fakultät

Local inequality and crime: New evidence from South Africa

Nicolas Büttner

Diskussionsbeitrag Nr. V-89-22

Volkswirtschaftliche Reihe ISSN 1435-3520

PASSAUER DISKUSSIONSPAPIERE

Herausgeber:
Die Gruppe der volkswirtschaftlichen Professoren
der Wirtschaftswissenschaftlichen Fakultät
der Universität Passau
94030 Passau

Local inequality and crime: New evidence from South Africa

Nicolas Büttner

Diskussionsbeitrag Nr. V-89-22

Volkswirtschaftliche Reihe ISSN 1435-3520

Adresse des Autors/der Autoren:

Nicolas Büttner
Wirtschaftswissenschaftliche Fakultät
Universität Passau
94030 Passau

Telefon: +49 851 509 5316
Telefax: +49 851 509 37 5316
E-Mail: nicolas.buettner@uni-passau.de

Local inequality and crime: New evidence from South Africa

Nicolas Büttner*

April 2022

Abstract: The relationship between inequality and crime has been of long-standing interest to social scientists of various disciplines. While theorical work from both economics and sociology postulates a positive link between the two, the empirical evidence is rather inconclusive and typically focuses on higher-income countries. In this study, I investigate the relationship between socio-economic inequalities of various dimensions and both violent and property crime at the local level in South Africa. For this, I created a novel panel dataset of police precincts that combines official crime records from the South African Police Service with socio-economic data from two population censuses and household surveys. For identification, I exploit the variation of inequality and crime across time and space, while controlling for socio-economic and demographic characteristics of police precincts, province-specific time trends, police cluster-fixed effects, and the spatial correlation of crime. I find strong and robust evidence for a significant, positive and linear relationship between income inequality within police precincts and local rates of violent crime and an inverted u-shaped relationship with property crime. Education inequality is more strongly related to violent crime, while housing inequality is only associated with property crime. In turn, cultural heterogeneity is positively correlated with all analyzed crimes. I also find suggestive evidence that inter-racial inequality contributes more to property crime, while intra-racial inequality contributes more to violent crime. Lastly, the results indicate that precincts which are relatively rich as compared to their neighbors suffer from higher rates of vehicle theft and aggravated robbery.

Keywords: Crime; Local inequality; Small Area Estimation; South Africa.

JEL-Codes: D31, D74, O12.

---

* Nicolas Büttner, University of Passau, Department of Economics, Innstraße 29, 94032 Passau, Germany (Phone +49 (0)851 509-5316, E-mail: nicolas.buettner@uni-passau.de).
Acknowledgements: This paper benefited greatly from the valuable advice provided by Michael Grimm.

1 Introduction

South Africa is among the countries with the highest income inequality worldwide, reporting a Gini Index of 0.67 in 2017 (UNU-WIDER). Strong disparities exist also with regards to other socio-economic domains, such as housing, health, and education (StatsSA, 2019a). At the same time, the country is known for persistently high levels of crime, especially violent crime. For 2019, the South African Police Service reported a total of 21,325 murders, corresponding to 58 murders a day and an annual murder rate of 36 murders per 100,000 inhabitants.[1] This makes South Africa one of the countries with the highest murder rates worldwide. The contemporary phenomenon of high inequality and a high crime rate is not exclusive to South Africa, but observable in many other countries, especially in Latin America and Southern Africa, raising the question whether socio-economic inequalities might be a driver of high crime rates (compare e.g. World Bank Group, 2021a; 2021b).

Indeed, well-established theories from economics and sociology postulate a link between inequality and crime. Economic theories, based on Becker (1968), propose that criminal behavior is a function of net utility from engaging in illegal versus legal activities, and it can be argued that inequality increases the net utility in favor of engaging in illegal activities. In sociology, strain theory (Merton, 1938) argues that inequality increases the share of relatively deprived individuals who are more likely to commit crime. Social disorganization theory (Shaw and McKay, 1942) sees inequality as a driver of crime through its weakening effect on communities' social cohesion and informal social control capacities.

There is also a large and growing body of empirical research on the relationship between income inequality and crime, including several meta-analyses. Hsiegh and Pugh (1993) conclude a positive association between inequality and violent crime at various geographic levels and Nivette (2011) finds inequality to be a strong predictor of cross-national homicide rates. Focusing on Europe, Kim, Seo and Hong (2020), find a positive correlation between inequality and crime only in Eastern and Northern Europe. Influential cross-country studies include Fajnzylber, Lederman and Loayza (2002a; 2002b), who, using a two decades country panel, find a positive effect of inequality on homicide and robbery. Yet, Neumayer (2003; 2005) questions their results as he finds no statistically significant effects with an extended sample. Pridemore (2011) argues that much of the cross-country evidence on the crime-inequality nexus is spurious and driven by poverty. Replicating previous studies, but controlling for poverty, he

[1] Own calculations based on national crime counts (https://www.saps.gov.za/services/crimestats.php) and mid-year-population estimates/projections (http://www.statssa.gov.za/?page_id=1854&PPN=P0302&SCH=72983)

finds no statistically significant effect of inequality on homicides, which is confirmed by Pare and Felson (2014), also for assault, robbery, burglary, and theft. Various studies analyze the crime-inequality nexus at the subnational level in the United States. The level of aggregation ranges from large areas like states (Kennedy et al., 1998; Choe, 2008; Chintrakarn and Herzer, 2012; Costantini, Meco and Paradio, 2018; Atems, 2020) to medium-sized areas such as Standard Metropolitan Statistical Areas (Blau and Blau, 1982), counties (Kelly, 2000; Brush, 2007; Kang, 2016) and cities (Baily, 1984; Stolzenberg, Eitle and D'Alessio, 2006; Hipp and Kane, 2017) to small areas like residential areas (Patterson, 1991), census tracts (Hipp, 2007; Stucky, Payton and Ottensmann, 2015) and census block groups (Metz and Burdina, 2018; Wenger, 2019). The majority of these studies focuses on the effects of income inequality within a geographical area and most, yet not all, present evidence of positive associations between inequality and both violent and property crime. Importantly, most studies are not representative of the entire US and many rely on cross-sectional data. In contrast to the extensive literature on inequality and crime in higher-income countries, especially the US, few studies have analyzed local inequality and crime in countries of the Global South. Enamorado et al. (2016) exploit a 20-year panel of Mexican municipalities, instrumenting local inequality with initial local inequality and national trends. They find a positive effect on homicide rates, yet, the results are mostly driven by drug-related homicides during the Mexican drug war (2007-2010). Further studies have been conducted in Brazil (Scorzafave and Soares, 2009; Menezes et al., 2013) and Colombia (Poveda, 2011; Buonanno and Vargas, 2017). To my knowledge, there is only one peer-reviewed study that examines the crime-inequality nexus in Africa. Demombynes and Özler (2005) analyze the effect of inequality on crime with a cross-section of South African police precincts in 1996. Their results suggest a strong positive association between inequality and property crime, while the evidence is less robust for violent crime. One important restriction is that, apart from correcting for crime misreporting, the authors do not address the issue of potential endogeneity between inequality and crime.

The present study builds on the work of Demombynes and Özler (2005) and extends it in various aspects to investigate whether South Africa's soaring crime rates can be explained by its high levels of socio-economic inequalities. For this, I construct a novel panel dataset of police precincts that combines official crime records from the South African Police Service (SAPS) with socio-economic data from the two population censuses in 2001 and 2011 and compatible household surveys. In absence of detailed income data in the census, Small Area Estimation (SAE) is employed to simulate household incomes necessary for income inequality calculations at the precinct level. The analysis comprises various types of property crime, namely residential burglary, vehicle theft, and aggravated robbery, as well as violent crime, namely aggravated assault, sexual offences, and murder. I do not restrict my analysis to the effects of income inequality, but also contribute to the rather scarce literature considering alternative inequality dimensions, namely education inequality (Kelly, 2000; Fajnzylber, Lederman and Loayza, 2002a), housing inequality (Manea, Piraino and Viarengoand, 2021) and cultural heterogeneity (Demombynes and Özler, 2005; Stolzenberg, Eitle and D'Alessio, 2006; Hipp, 2007; Hipp and Kane, 2017). For identification, I exploit the variation of inequality and crime rates across time and space, while controlling for the economic, social and demographic characteristics of police precincts. I also use the panel structure of my dataset to mitigate issues of omitted variable bias and account for the spatial correlation of crime rates.

The main findings of this paper can be summarized as follows. For income inequality within precincts, I find a positive linear association with violent crimes and an inverted u-shaped association with property crimes. This might indicate that income inequality within a local community drives crime in general, but that for property crimes, this effect is counteracted by protective measures employed in high-inequality communities. Education inequality is more strongly related to violent crime, while housing inequality is only associated with property crime. In turn, cultural heterogeneity is positively correlated with all analyzed crimes. I conduct various robustness checks, which generally confirm these results. Overall, the strongest and most robust associations are found between violent crime and income inequality as well as cultural heterogeneity, which is most in line with strain theory.

This paper also contributes to the discussion on the relative importance of intra- versus inter-racial income inequality (Blaun and Blau, 1982; Demombynes and Özler, 2005; Stolzenberg, Eitle and D'Alessio, 2006; Hipp, 2007). Suggestive evidence indicates that inter-racial inequality is more important for property crime, while intra-racial inequality contributes particularly to violent crime. Lastly, the paper adds to the scarce literature on the role of income inequality between local communities, sometimes referred to as 'economic segregation' (Stucky, Payton and Ottensmann, 2015; Kang, 2016; Metz and Burdina, 2018). While income inequality within precincts is related to all analyzed property and violent crimes, between-precinct inequality is positively correlated only with rates of aggravated robbery and, to a lesser extent, vehicle theft.

The remainder of this paper is structured as follows. Section 2 provides more context on inequality and crime in South Africa. In Section 3, I summarize the theoretical literature and derive hypotheses to be tested in the empirical analysis. Section 4 describes the various data sources and the construction of crime and inequality measures and the control variables, and Section 5 illustrates the empirical specification. In Section 6, I present and discuss the results and Section 7 concludes.

2 Inequality and crime in South Africa

Inequality in South Africa

Social and economic inequalities are deeply entrenched in South Africa's history. Their origins can be traced back to the centuries of exploitation of the African population under colonial rule and the policies of racial discrimination and segregation during Apartheid in the 20th century.[2]

Whereas the Apartheid system provided high quality education, health and housing services to the white minority population, the non-white majority population only had access to rudimentary services. In addition, the government introduced discriminating labor market policies favoring the white workforce and impeded the allocation of resources (including land, mining rights, and access to capital) among the non-white population. The system of racial discrimination denied a huge portion of the South African population to accumulate significant human or physical capital and marginalized them to the worst-paid sectors and jobs in the economy (Woolard, 2002). Even though racially discriminating labor market policies were gradually dismantled from the 1970's on, the combination of rising unemployment, destroyed subsistence agriculture and the racially biased distribution of human capital (and its intergenerational persistence) prevented the vast majority of the non-white population (and in particular of the black population) to escape poverty (Seekings and Natrass, 2005). The racially biased distribution of income over the course of the twentieth century is illustrated in Table 1 (based on Leibbrandt, Woolard and Woolard, 2009). From as early as 1917, i.e. pre-dating Apartheid, all non-white population groups had consistently lower average incomes than white South Africans, and this gap was the strongest at the peak of the Apartheid state in the 1950's and 1960's. The disparity of incomes was especially pronounced for black Africans, receiving only between 6.8% and 10.9% of an average white person's income between 1917 and 1993.

[2] See Wilson (2001) for a concise but thorough overview of the "Historical Roots of Inequality in South Africa" and Terreblanche (2004) for an extensive description of "A History of Inequality in South Africa 1652-2002".

Table 1: Average income of non-white population groups (% of white level), 1917-1993

	1917	1924	1936	1946	1956	1959	1960	1970	1975	1980	1987	1993
Indian/Asian	22.1	19.4	23.1	23.0	21.9	17.1	17.1	20.2	25.4	25.5	30.2	42.0
Coloured	22.0	20.0	15.6	16.3	16.9	15.7	15.9	17.3	19.4	19.1	20.9	19.3
Black African	9.1	7.9	7.6	8.9	8.6	7.7	8.1	6.8	8.6	8.5	8.5	10.9

Source: Leibbrandt, Woolard and Woolard (2009).

These huge income differences were reflected by a Gini coefficient of per capita income of around 0.67 towards the end of the Apartheid era in 1993. Despite the transition to a democratic, non-racial and inclusive society in 1994, inequality levels have remained very high until today. Income inequality even increased to a Gini coefficient of around 0.70 (temporarily up to 0.73) between 2001 and 2011, and afterwards reduced, reaching again its 'end-of-Apartheid'-level of around 0.67 in 2014.[3] The latest estimates of the World Income Inequality Database (UNU-WIDER) see South Africa at the same level in 2017, making it 'officially' and by far the most unequal country in the world.[4]

Table 2: Income inequality in 1993, 2008, and 2015; overall, by population group, decomposed

	GE(0)			GE(1)			Gini		
	1993	2008	2015	1993	2008	2015	1993	2008	2015
Overall	0.94	1.05	0.91	0.92	1.02	1.19	0.67	0.70	0.68
Black African	0.60	0.82	0.68	0.58	0.89	0.74	0.56	0.64	0.60
Coloured	0.53	0.81	0.56	0.76	0.70	0.66	0.52	0.59	0.56
Indian/Asian	0.40	0.96	0.70	0.46	0.69	0.64	0.46	0.61	0.60
White	0.36	0.51	0.83	0.37	0.46	1.18	0.43	0.51	0.65
Within (%)	58.5	74.6	75.1	51.5	66.8	74.9	-	-	-
Between (%)	41.5	25.4	24.9	48.8	33.2	25.1	-	-	-

Notes: GE(0) refers to the mean log deviation and GE(1) to the Theil index.
Source: Hino et al. (2018)

While inequality between races is still a major contributor to overall inequality, it has decreased significantly in the post-Apartheid era. Hino et al. (2018) use two measures of general entropy to decompose income inequality into inequality between and within racial groups. As illustrated in Table 2, the contribution of between-race inequality to overall inequality dropped from 41.5% resp. 48.8% in 1993 to 24.9% resp. 25.1% in 2015. This means that today around 75% of overall inequality is caused by inequalities in the income distributions within racial groups.

[3] This description is based on Gini estimates for various years in the period 1993-2015 from the following sources: Leibbrandt. et al. (2005, 2010), Yu (2010), Leibbrandt, Wegner and Finn (2011), Finn and Leibbrandt (2013), Leibbrandt, Finn and Woolard (2012), Hino et al. (2018), Hundenborn, Leibbrandt and Woolard (2018), StatsSA (2019a), UNU-WIDER.
[4] For comparison: Brazil 0.53, Mexico 0.45-0.50, USA 0.41-0.46.

Nevertheless, while income disparities between racial groups, especially between whites and black Africans, have strongly declined, they remain profound. In 2015, the median income of whites was still six times higher than that of black Africans and four times higher than that of coloureds.

Lastly, and not surprisingly, inequality in South Africa is not nearly restricted to the distribution of income. Huge disparities continue to exist also with regards to asset ownership and access to and quality of education, health and basic services. Table 3 illustrates the ongoing and strong discrepancies in living standards between population groups, with black Africans being by far the most disadvantaged. Black Africans own on average less assets, go less often to private schools and private health facilities, are less often covered by medical aid and have less access to electricity, piped water, improved sanitation, refuse removal and internet as compared to the other population groups, especially white South Africans. Similarly, one can see a strong urban-rural divide across all (except electricity access) these domains.

Table 3: Inequality in assets, education, health, basic services: overall, by population group, by location

	Overall	Black African	Coloured	Indian/ Asian	White	Urban	Rural
Asset ownership							
Number of assets owned (mean)	9.8	9.0	11.5	13.2	14.7	10.9	7.8
Education							
Attendance of private school, children 6-18 (%)	6.5	5.0	4.0	22.6	32.0	9.8	2.3
Health							
Access to medical aid (%)	16.9	10.1	20.2	48.9	72.4	23.4	5.4
Use of private health facility (%)	24.9	17.0	32.4	65.5	87.4	-	-
Basic services (access to…)							
… electricity (%)	84.4	81.5	92.0	98.1	98.6	83.7	85.8
… piped water (%)	74.2	69.1	95.4	97.5	94.8	88.8	40.4
… improved sanitation (%)	82.5	78.7	96.9	98.8	99.4	90.6	62.8
… refuse removal (%)	65.9	58.8	91.6	96.3	94.0	88.2	10.9
… internet (%)	62.2	58.1	64.1	78.4	90.3	70.5	42.9

Notes: Data on asset ownership are from LCS 2014/15 and data on education, health and basic services from GHS 2017. Unit of analysis is the household for basic services and the individual for all other categories.
Source: StatsSA (2019a)

Overall, while some progress has been made since the fall of Apartheid, especially with regards to basic services, a large portion of the population continues to be excluded from the promises of the 'new South Africa' (StatsSA, 2019a).

Crime in South Africa

South Africa is notorious for its high crime rates and in particular for its high murder rate. With 34 murders per 100,000 inhabitants (yearly average 2014-2018), it ranks 5th across the globe, after El Salvador (73), Venezuela (52), Honduras (52), and Jamaica (45). To put this into perspective, in the same period, the murder rate was 22 in Mexico, 11 in Uganda, 5 in the US, and 1 in Germany.[5] South Africa's crime problem isn't nearly restricted to murder – the country displays in general a remarkably high level of crime, especially violent crime. Figure 1 on the next page illustrates the reported crime rates at the national level for six serious crimes over the period 1994 to 2019.[6] Throughout the entire period, South Africa experienced very high rates of residential burglary and aggravated assault. Yet, after an increase in the first years after Apartheid, crime rates declined steadily and essentially halved from a maximum of 694 resp. 630 in 2000 to a minimum of 351 resp. 284 in 2019. Though still very high, rates of vehicle theft continuously declined from 274 at the end of Apartheid to 80, i.e. by 70%. Rates of aggravated robbery show strong fluctuations over this time period, and overall increased slightly, from 219 to 245. Rates of sexual offences are only reported from 2001 on, and the legal definition changed in 2007, making comparisons over time difficult. Focusing on the period 2008 to 2019, a remarkable decrease is visible, from 139 in 2008 to 91 in 2019, yet, the last years indicate a stall or even a slight reversal of the trend. To put this into perspective, in 2019 more than 53,000 sexual offences have been reported, of which around 43,000 cases were rapes, corresponding to around 188 reported rapes every single day, not considering a substantial amount of unreported rapes. Lastly, the murder rate declined significantly from 67 in 1994 to its minimum of 30 in 2011, but has since then steadily increased, reaching 36 in 2019. That means that in one year more than 21,000 people are murdered in South Africa, the equivalent of the population of Stellenbosch.[7]

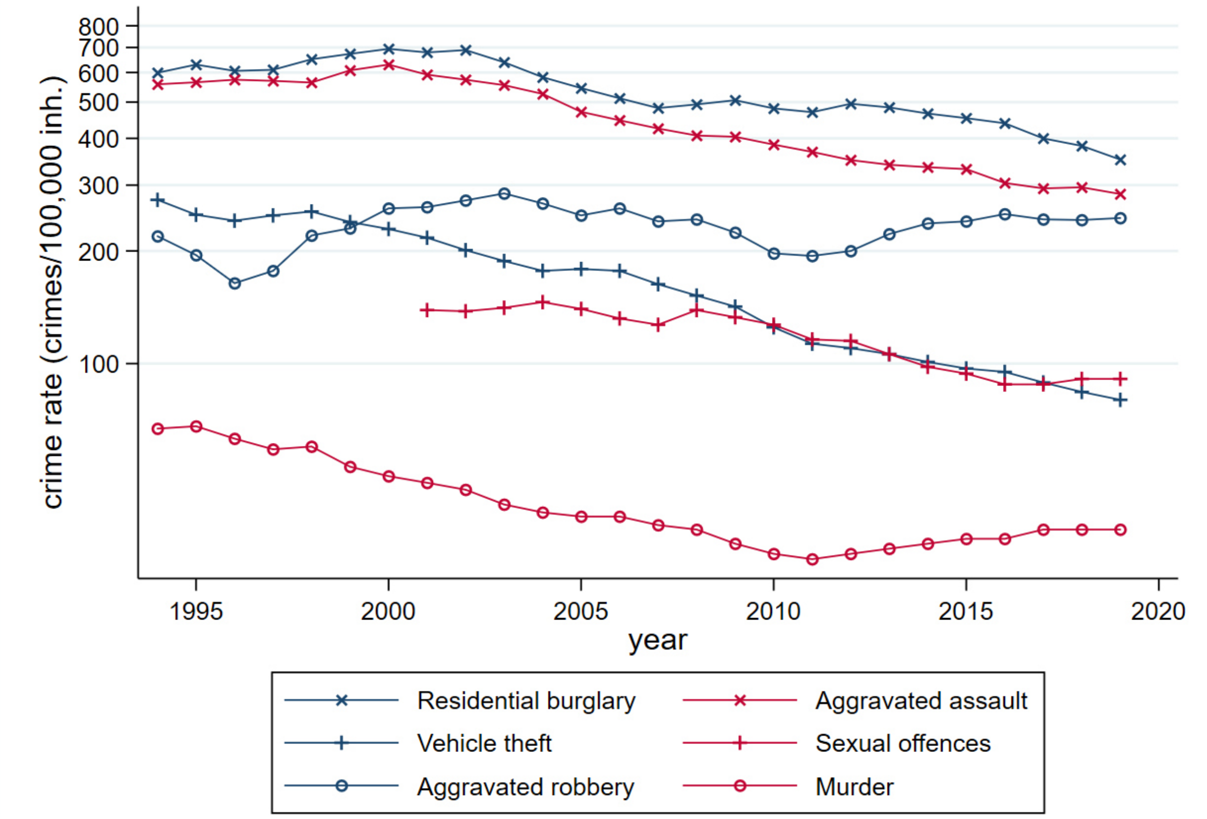
---

[5] Own calculations based on UNODC database "Victims of intentional homicide, 1990-2018". Data available under https://public.tableau.com/app/profile/unodc.rab/viz/Homiciderates_15826327950430/Homicide-rates. I only considered countries with data available from 2014 to 2018. Saint Kitts and Nevis, United States Virgin Islands, Lesotho and Nigeria displayed crime rates between 50 and 35, but were excluded due to incomplete data records.

[6] National crime counts:
1994-2003: https://issafrica.org/crimehub/facts-and-figures/crimehub/crime-trends-1994-to-2004/crimehub/national-crime-statistics-by-crime-type
2004-2019: https://www.saps.gov.za/services/crimestats.php and https://africaopendata.org/dataset
Mid-year-population estimates and projections:
http://www.statssa.gov.za/?page_id=1866&PPN=P0302&SCH=72983&page=1 and
http://www.statssa.gov.za/?page_id=1854&PPN=P0302&SCH=72983.

[7] Considering only the town itself, not Stellenbosch local municipality.

Figure 1: National crime rates from 1994 to 2019 for selected property and violent crimes



Notes: Y-axis in log-scale. Crime rates were calculated by combining annual national crime counts from ISS (1994-2003) and SAPS (2004-2019) and annual mid-year-population estimates and projections from StatsSA.
Source: Own calculations using data from and ISS, SAPS, and StatsSA.

Overall, despite substantial improvements in the post-Apartheid era, crime levels remain extraordinarily high, with detrimental effects on the South African population. Apart from the direct cost of crime for victims, high crime levels might also increase stress levels and worsen (mental) health for large parts of the population. This is illustrated by results from the Victims of Crime Survey (VOCS) 2019/2020, in which 58% of adults aged 16 or older reported to feel unsafe when walking alone in their area of residence when it was dark (StatsSA, 2020). 31% reported to have taken physical protection measures (e.g. burglar doors), 7% invested in private security (e.g. paid armed response), and 2.5% carried a gun to protect themselves from criminals, as recorded by the VOCS 2018/2019 (StatsSA, 2019b). These investments into protective measures might reduce investment into other (more productive) assets. Lastly, it has also been suggested that in particular the high levels of violent crime have been a major driver of emigration out of South Africa. This could be especially detrimental to the country's development as the majority of emigrants are highly educated, potentially causing, at least to some extent, a 'brain-drain' (Dodson, 2002; Marchetti-Mercer, 2012).

The root causes of South Africa's extreme levels of violent crime have been subject to longstanding and controversial debates. In 2007, the South African government commissioned the Centre for the Study of Violence and Reconciliation (CSVR) to carry out a comprehensive study on the violent nature of crime. The CSVR (2009) identified two groups of factors contributing to South Africa's endemic violence. First, the legacy of Apartheid and colonialism, namely the evolution of a self-perpetuating 'culture of violence', the mass-separation of African families, the institutionalization of racism, a biased criminal justice system, and the proliferation of firearms. Second, factors in post-Apartheid South Africa which reinforce the legacy of Apartheid, namely weak state institutions, structural economic factors (high levels of poverty, structural unemployment, and social and political exclusion and marginalization), and a high level of inequality. Earlier, Schönteich and Louw (2001) had also identified the 'culture of violence', the proliferation of firearms and the weak criminal justice system as contributing factors to South Africa's high crime rates. In addition, they suggest that rapid urbanization, the relatively young population, organized crime syndicates and the period of political transition in the 1990's increased crime rates. This paper focuses on one of the above aspects; the contribution of inequality to South Africa's high crime rates.

3 Theoretical considerations

This section discusses the theoretical link between inequality and crime. It draws on various theories from economics and other social sciences and derives a set of hypotheses that will guide the empirical analysis below. All theories postulate a positive link between inequality and crime, i.e. they suggest that higher inequality leads to more crime. However, most approaches are not very specific regarding the type of crime and the concept of inequality they have in mind, i.e. inequality with regards to which aspect, on which level, and within or between which groups.

The economics of crime

Economic theories of crime can be traced back to Gary Becker's (1968) seminal work "Crime and Punishment: An Economic Approach". In his choice theoretic model, a person commits a crime if the expected utility from this offense exceeds the expected utility from alternative (legal or illegal) activities. Accordingly, Becker models the number of offences committed by an individual as a function of the probability of being convicted and the expected cost of punishment as well as the income that is accessible through legal and other illegal activities. Ehrlich (1973) contributes by explicitly incorporating the costs and benefits of legitimate and illegitimate activities. Hence, he augments Becker's model by adding returns from illegal activities, returns from legal activities, and the probability of being unemployed in the legal sector. Whereas an increase in returns from illegal activities and in the probability of unemployment leads to more offenses, higher returns from legal activities decrease the number of offenses. Several more authors have made theoretical contributions to the 'economics of crime' literature; including Stigler (1970), Sjoquist (1973), Block and Heinecke (1975), Chiu and Madden (1998), Freeman (1999), Imrohoroglu, Merlo and Rupert (2000), Bourguignon (2001), and Machin and Meghir (2004).

Although most economic theories of crime do not focus on inequality, it is straightforward to embed them in the concept of inequality. Since higher returns (income) from legal activities reduce the utility from illegal activities, one may argue that, conversely, lower returns from legal activities increase the utility from crime. Put differently, a community with a large share of the population being poor or excluded from accessing higher income from legal activities, e.g. due to high unemployment, is likely to be accompanied by higher crime levels. Thus, the theories do not make specific predictions regarding inequality, but rather have implications for poverty. However, if poverty is driven by excess inequality, one can at least derive an indirect link between inequality and crime. Yet, these theories cannot establish a direct link from inequality *per se* on crime; it is only through the implied assumption that a higher level of inequality is associated with a higher level of poverty. This is of course not necessarily true, as inequality could also be mainly driven by high levels of inequality in the upper tail of the income distribution.

The idea that higher income inequality is associated with more crime is also put forward by Demombynes and Özler (2005). They argue that in particular the distance from the poor's incomes to the mean household income is an important determinant of crime levels, as a higher mean household income should transfer into more valuable assets that, in turn, increase the returns from criminal activities. Hence, they establish a more direct link between inequality and crime by explicitly considering the poor's distance from the better off instead of solely looking at poverty. The above does of course only have relevance for property crime; it is not capable of making any predictions regarding violent crime. Hence, the following hypotheses can be derived:

H1: Higher within-community inequality as well as higher between-community inequality is associated with higher levels of property crime, but not violent crime.

H2: Property crime is higher in communities with a relatively high mean income as compared to their adjacent communities since they offer higher potential returns from property crime for their ‘poor neighbors’.

Social disorganization theory

In their seminal work “Juvenile Delinquency and Urban Areas”, Shaw and McKay (1942) introduce a theory of social disorganization as an explanation for the variation in crime rates amongst urban American communities. A central claim is that the crime rate is influenced by the economic, social and cultural structure of a community and that “neighborhood ecological conditions shape crime rates over and above the characteristics of individual residents” (Kubrin and Weitzer, 2003). They argue that adverse neighborhood “conditions eroded community norms and conventional values and prevented residents from responding to or preventing unwanted behavior in the neighborhood, and as a result, crime flourished” (Wickes, 2017). Social disorganization theory was further developed by, amongst others, Kornhauser (1978), Stark (1987), Bursik (1988), Sampson and Groves (1989), and Bursik and Grasmick (1999). The common ground of social disorganization theory is their focus on a community’s ability to maintain informal social control over its inhabitants. Kelly (2000) and Stucky, Paytona and Ottensmann (2015), for example, argue that, within this framework, inequality within a community is likely to cause crime through its adverse effects on family stability, social networks, supervision of youth and collective efficacy and thereby weakening social control. Yet, it is not so straightforward to derive a *direct* link between inequality and crime from social disorganization theory. The mentioned ‘adverse effects’ are unlikely to unfold as long as inequality does not come along with a substantial level of poverty. This implies that, similar as in the economic theories, one can rather establish an *indirect* link from inequality to crime via its effects on poverty. Importantly, however, social disorganization explicitly mentions cultural heterogeneities, which do not necessarily imply poverty, but could directly weaken social cohesion and impede crime prevention via social control. The described mechanisms hold for property as well as for violent crime as it is not plausible that a community is able to prevent one type of crime but not the other. Hence, the following hypothesis can be derived:

H3: Higher within-community inequality is associated with higher levels of property crime as well as violent crime.

Strain theory

In his seminal work "Social Structure and Anomie", Merton (1938) develops strain theory as an explanation for criminal behavior. According to Merton, criminal behavior emerges when a society overemphasizes the importance of symbols of success, while strongly restricting the access to these symbols through legal means. He states that "when a system of cultural values extols, virtually above all else, certain common symbols of success for the population at large while its social structure rigorously restricts or completely eliminates access to approved modes of acquiring these symbols for a considerable part of the same population, […] antisocial behavior ensues on a considerable scale." Merton argues that individuals of low social status are frustrated due to the lack of success or the symbols of success. He further suggests that frustration is increased when being confronted with more successful individuals and that this frustration leads to alienation from society and fosters criminal behavior (Kelly, 2000).[8] Strain theory has been developed further by Cohen (1955), Cloward and Ohlin (1960), Goode (1960), Agnew (1992), and Messner and Rosenfeld (1994), among others. The concept of inequality is relevant in strain theory as inequality increases the gap between social classes and thus feelings of relative deprivation and thereby leads to more crime. This mechanism is applicable to both property and violent crime. Property crime can be seen as a direct means to alleviate relative deprivation by accumulating resources through illegitimate activities. This reasoning also holds in the presence of legitimate alternatives. If the feelings of frustration and injustice are very strong, they might be used as justification to take away from those who have plenty of what is desirable. Violent crime can be explained rather as a product of constant feelings of frustration and social exclusion. These feelings could transmit into anger and result in violent behavior. Thus, communities with higher within-community inequality should display higher rates of both property and violent crime, as postulated in H3.

If peoples' reference group also includes individuals from adjacent communities, between-community inequality should likewise affect crime rates. Property crime is expected to be higher in better-off communities. Relatively deprived individuals from adjacent communities

[8] Although not mentioned explicitly, Merton makes implicit use of the concept of "relative deprivation". According to Runciman (1966), a person is relatively deprived by an object if 1) he doesn't possess it, 2) he knows of other persons who possess it, 3) he wants to possess it, and 4) he believes obtaining it is realistic.

might deliberately commit property crimes in better-off communities as they seem more justifiable when committed against the rich. Empirically, this leads to the same prediction as the one expressed by H2. However, violent crime is expected to be higher in communities that are worse-off in comparison to adjacent communities, as they are likely to inhabit more relatively deprived individuals that express their frustration (in most cases) close to their place of residence. Hence, the following hypothesis can be derived:

H4: Violent crime is higher in communities with a relatively low mean income as compared to their adjacent communities since they inhabit more relatively deprived individuals that are more likely to resort to violence.

4 Data

I create a novel dataset by combining official crime statistics from the South African Police Service (SAPS) with socio-economic and demographic data from the South African censuses of 2001 and 2011. As the censuses do not contain detailed information on income, I impute this data with the help of household surveys and the method of Small Area Estimation (SAE). In the following, I describe in detail the various data sources and how they were merged to create this novel dataset and provide a brief summary of the income imputation via SAE. Then, I explain the measurement of crime, the definitions of the various inequality indices, and the choice of control variables. Lastly, I present some summary statistics.

4.1 Data sources and preparation of dataset

Crime statistics (South African Police Service)

I use the official crime statistics for South Africa, which are published every year in September by the South African Police Service (SAPS). Crimes are recorded at the level of police stations for the period April to March (SAPS financial year).[9] Crimes are grouped into five broad categories: contact crimes, contact related crimes, property related crimes, other serious crimes (altogether labelled community reported crimes), and crimes detected as result of police action.

[9] The most current version is publicly available on the website of the SAPS and contains data for the past ten years (https://www.saps.gov.za/services/crimestats.php). Hence, crime data from more than ten years ago is not available on the SAPS website. However, older versions of the SAPS crime statistics are archived and available from openAFRICA including data back until the year 2003/2004 (https://africaopendata.org/dataset). In addition, I received crime data for earlier years (2000/2001 and 2001/2002) from the Institute of Security Studies (ISS).

In total, data is provided for 21 different offences, plus several aggregates and subcategories.[10] With the exception of sexual offences, there have been no changes with regards to the legal definitions of the recorded offences.[11] In addition, data availability for the various offences is very consistent over time.[12] A major drawback of the SAPS crime statistics results from the fact that since 2000/2001 more than 50 new police stations have been opened and therefore police precinct boundaries (and policing areas) of neighboring stations have changed over time. Crime rates of affected police precincts are thus not comparable over time. In order to create a panel of police precincts with consistent boundaries over time, one needs information about the opening date of all new stations established since 2000/2001 and the names of the stations where crimes in the affected areas where recorded prior to the opening. Since SAPS does not publish these data, I collected the information manually by requesting non-published SAPS data, scanning news reports about station openings, and contacting affected stations individually via e-mail and phone. In combination with a shapefile of all police precinct boundaries as of September 2014[13], the collected information enabled me to create a police precinct panel of crime statistics with consistent boundaries and policing areas from 2000/2001 until 2014/2015, covering 1,089 police precincts, organized in 179 police clusters. I drop four ‘special’ precincts, since their entire policing area is comprised of a major international airport or harbor. These precincts have a very small population, but extremely high crime levels, potentially biasing the results. In the end, I analyze 1,085 precincts from 176 clusters.

Censuses

Since the end of Apartheid, Statistics South Africa (StatsSA) has conducted three national censuses, in 1996, 2001, and 2011. Due to the limited comparability of the 1996 census with the newer censuses (Cronje and Budlender, 2004) and the unavailability of crime data from that era, I only use the censuses from 2001 and 2011. 10% samples of both censuses are publicly available. Apart from not covering the full population, the main obstacle of using these samples for this study is the lack of detailed information on the location of the surveyed households.

[10] Table A1 in Appendix A illustrates the SAPS crime categorization.
[11] Until 2007, sexual offences only included rape and indecent assault. Since the *Criminal Law (Sexual Offences and Related Matters) Amendment Act, Act 32 of 2007*, this definition has broadened considering over 50 crimes relating to sexual violence, sexual exploitation, sexual grooming, trafficking, and pornography. This makes comparisons over time problematic, yet, in a regression framework with time-fixed effects, this is less of a problem.
[12] Table A2 in Appendix A provides an overview of precinct-level data availability across SAPS financial years per offence.
[13] https://www.saps.gov.za/services/boundary.php (downloaded in July 2017)

Precisely, the lowest administrative unit for which information is included is the local/metropolitan municipality (2001: N=262, 2011: N=234). It is not feasible to precisely match census data and crime statistics based on these municipalities, since their boundaries are not aligned with police precinct boundaries.[14] Fortunately, I received customized population censuses for both years from StatsSA upon request. Most importantly, these datasets contain detailed geolocation data for each household in form of the Small Area Layer (SAL) code. The SAL is the second lowest aggregation level of the census; in fact, only the Enumeration Areas (EA) are more spatially disaggregated. For the Census 2001, South Africa was divided into 56,255 SALs and in 2011 into 84,907 SALs. The very small size of the SALs in comparison to the police precincts enabled me to adequately match census and crime data, which is illustrated in the next section. A further advantage of the customized census datasets is that they cover the full population; yet, for confidentiality reasons, they do not contain all variables (only those relevant for this study). The 2001 census covers around 9.5 million households and 36.5 million individuals, the 2011 census around 12.5 million household and 44.2 million individuals.[15]

Data merging

The lacking alignment of boundaries between police precincts and other administrative units in South Africa makes it impossible to allocate all of the census households precisely to ‘their’ police precinct. However, the availability of highly disaggregated geo-information data enables me to merge crime and census data in a sufficiently precise manner. For this, I use shapefiles of the police precinct boundaries (consistent over time), shapefiles of the SAL boundaries in 2001 and 2011, and raster data of population distribution estimates for 2001 and 2011.[16] In each census year, I start by overlaying the shapefiles of police precincts and SALs. For those cases where a SAL lies entirely within the boundaries of one police precinct, all households are allocated to this police precinct.[17] Whenever a SAL overlaps two or more police precincts, I additionally use the population distribution estimates to calculate a ‘SAL-to-precinct-weight’. This weight denotes for each SAL the share of its population that is located in every precinct it

[14] Administrative units like provinces, district municipalities and local/metropolitan municipalities are hierarchical. However, police precinct boundary demarcation is the sole responsibility of SAPS and is not harmonized with the boundaries of those administrative units.
[15] The census post-enumeration survey reports indicate an undercount of around 20.5% for households and 17.6% for individuals in 2001 and 14.3% respectively 14.6% in 2011.
[16] Shapefiles of census SALs are publicly available together with the 10%-Census samples. Population distribution estimates are provided from WorldPop (https://www.worldpop.org/doi/10.5258/SOTON/WP00645).
[17] 91.0% of SALs, 90.4% of households and 92.4% of individuals in 2001; 92.4% / 91.1% / 92.9% in 2011.

overlaps. This weight is later used when creating precinct level variables based on the census data. More details on the merging procedure, in particular the creation of the weights, is provided in Appendix B.

Small Area Estimation

As indicated above, I use Small Area Estimation (SAE) to impute income data for the census households. SAE was developed by Elbers, Lanjouw and Lanjouw (2003) and combines the advantages of census data (comprehensive coverage) and survey data (detailed income/expenditure data) to impute income or expenditure for households in the census data. The approach contains three main steps. First, selection of candidate variables that are available and measured in the same way in both census and survey and that could plausibly serve as explanatory variables for the outcome of interest (e.g. income). Second, modelling income based on selected candidate variables with the survey data, including a decomposition of the residual into a household and cluster component. Third, simulation of income for the census households based on the estimated model parameters, including the distributions of the household and cluster error. Importantly, all steps have to be conducted at the stratum level, i.e. separately for urban and rural households in each province, and separately for 2001 and 2011. I combine the Census 2001 with the Income and Expenditures Survey (IES) 2000 and the Labor Force Survey (LFS) 2000 and the Census 2011 with the IES 2011. For the large part of the implementation of SAE, I rely on the World Bank's poverty mapping Software, PovMap 2.0, and the corresponding user manual (Zhao and Lanjouw, n.d.).[18] A step-by-step description of the implementation is provided in Appendix C.

4.2 Measurement of crime

I measure criminal activity as the year- and precinct-specific crime rate, i.e. the incidence of a particular crime per 100,000 inhabitants. The required police precinct population estimates are calculated by combing the 'SAL-to-precinct-weights' with SAL-level population counts from the censuses.[19] I include six distinct crimes: burglary at residential premises ('residential burglary'), theft of motor vehicle and motor cycle ('vehicle theft'), assault with the intent to

[18] Software and manual available at: https://www.worldbank.org/en/research/brief/software-for-poverty-mapping.

[19] My approach is inspired by Kempe (2016), however, I use weights based on population distribution, while he uses area-based weights.

inflict grievous bodily harm (‘aggravated assault’), sexual offences, murder, and robbery with aggravating circumstances (‘aggravated robbery’).[20] I classify residential burglary and vehicle theft as ‘property crimes’ since for both offences the main motivation is material gain and violence need not necessarily be used. Aggravated assault and sexual offences are classified as ‘violent crimes’, since these offences by definition use violence against a person and material gain is, if at all, a secondary motive. Murder and aggravated robbery cannot exclusively be allocated into either category. While murder always inhibits a violence component, it is frequently committed with material gains as a primary motive. In case of aggravated robbery, the criminal uses force against the victim to obtain money or assets from them. Yet, considering which motives are more often of primary nature, I treat murder as ‘violent crime’ and aggravated robbery as ‘property crime’.

4.3 Measurement of inequality

I consider inequality with regards to the following four dimensions: income, education, housing, and culture. Inequality is measured alternatively within or between police precincts.[21] In addition, I decompose within-precinct income inequality into intra- and inter-racial inequality. In the following, I explain the construction of the various inequality indicators.

Inequality within police precincts

The basic idea of within-precinct inequality is to consider the population within a given precinct and calculate inequality indices based on this population’s distribution with regards to a variable of interest, such as income. This implies that individuals in a given precinct mainly compare themselves to others from the same precinct.

I measure income inequality with the Gini index of monthly per capita household income. Household income includes income from salaries/wages and net business profits, rental income, royalties, interests and dividends, pensions, welfare and grants, and allowances from family members or other individuals. The Gini index reaches its minimum value of zero (i.e. complete equality) when all households earn the same income, i.e. when there are no income differences, and its maximum of one (i.e. complete inequality) when one household earns all available

[20] An overview of the official definitions of crime types as employed by SAPS is provided by ISS under https://issafrica.s3.amazonaws.com/site/uploads/factsheet-sa-crime-stats-definitions.pdf.

[21] I treat the police precinct as a form of a local community, i.e. the term ‘police precinct’ is synonymous to the term ‘community’ employed in Section 3.

income and all other households earn no income. Alternatively, I also use two indices of Generalized Entropy, namely the mean logarithmic deviation (GE(0)) and the Theil index (GE(1)). While the Gini index is more sensitive to income differences at the middle of the distribution, GE(0) is more sensitive to inequality at the bottom and GE(1) more sensitive to inequality at the top. For both indices, the minimum value zero stands for perfect equality, while larger positive values indicate more inequality.

Education and housing are both measured as categorical variables, for which neither of the above presented indices are suitable. This is due to the ordinal nature of the ordering, i.e. it is possible to clearly state that category 3 is better than 2 and 2 is better than 1; however, it is not possible to interpret these differences quantitively. I thus rely on 'status'-based measures of inequality for categorical variables, introduced by Cowell and Flachaire (2017) and further developed by Jenkins (2020). Cowell and Flachaire (2017) define (peer-inclusive downward-looking) status, s_i, as the share of the population that is as well off or worse off than individual i, e.g. the share of individuals that have at most the education level of individual i. Maximum status is thus 1, i.e. no one is better off than individual i. Based on this concept of status, they developed a family of indices that, in essence, sum up the differences between individuals' statuses and the maximum attainable status. With $0 \leq \alpha < 1$, the indices are defined as:

$$I_\alpha = \begin{cases} \frac{1}{\alpha(\alpha-1)}\left[\frac{1}{n}\sum_{i=1}^{n} s_i^\alpha - 1\right], if\ \alpha \neq 0{,}1, \\ -\frac{1}{n}\sum_{i=1}^{n} \log s_i\,, if\ \alpha = 0. \end{cases} \qquad (1)$$

The parameter α determines the sensitivity of the index towards inequality at the bottom (smaller α) versus at the top (larger α) of the status distribution. The index reaches its minimum of zero when all individuals have the same status (are in the same category). Importantly, a uniform distribution across all categories yields higher values than polarization (half of the population in the lowest category and half in the highest). The basis of Jenkins' (2020) index are generalized Lorenz (GL) curves for status distributions. Graphically spoken, the larger the area between the 45-degree line (complete equality) and the GL curve, the higher the inequality. If two GL curves do not cross, inequality comparison is unambiguous, i.e. independent of the value of α. Mathematically, the index is defined as:

$$J = 1 - \sum_{j=0}^{K-1}(p_{j+1} - p_j)\big(GL_j + GL_{j+1}\big) = 1 - \sum_{j=0}^{K-1} f_{j+1}\big(GL_j + GL_{j+1}\big), \qquad (2)$$

where p are the vertices of the GL curve and $J = 0$ if all individuals have the same status. Again, a uniform distribution yields higher inequality values than polarization. Education inequality is based on the highest education level attained amongst all individuals aged 18 or older. For this, I grouped education levels into 17 categories, with ‘No education’ as the lowest and ‘Postgraduate degree’ as the highest.[22] Housing inequality is based on the type of dwelling amongst all households. Dwelling types are categorized in seven categories, with ‘Informal dwelling’ as the lowest and ‘Freestanding house’ as the highest.[23]

Lastly, I consider cultural heterogeneity. The basis for this measure is individuals’ affiliation with one of South Africa’s official population groups: Black African, Coloured, Indian or Asian, White. In contrast to education or housing, culture/race/ethnicity/population group is clearly of non-ordered nature (neither group is better than another); thus, the status-based concept of inequality illustrated above is not suitable here. Instead, I rely on an approach of generalized variance (GV), as summarized by Budescu and Budescu (2012). Specifically, I use the trace of the covariance, which is the sum of the variances of all categories, calculated as:

$$GV = \sum_{i=1}^{C} P_i(1 - P_i) = 1 - \sum_{i=1}^{C} P_i^2. \quad (3)$$

As its maximum value depends on the number of categories, making comparisons across populations with a varying number of categories difficult, the index is often normalized, such that it becomes bounded between zero and one:

$$NGV = \frac{GV}{Max(GV)} = \frac{C}{C-1}\left(1 - \sum_{i=1}^{C} P_i^2\right), \quad (4)$$

where C is the number of population groups and P_i is the share of people in population group i. The index reaches its minimum of zero (complete homogeneity) when all individuals belong to the same population group and its maximum of 1 (complete heterogeneity) when the population is equally distributed across all population groups. The index value can be interpreted as the probability that two randomly selected individuals don’t belong to the same population group.

Table 4 on the next page provides an overview of all measures of inequality presented in this section.

[22] I also use two alternative groupings with eight respectively five categories. For more details see ‘Further robustness tests’ in Section 6.7.

[23] I also use two alternative groupings with seven (but different ordering) respectively five categories. For more details see ‘Further robustness tests’ in Section 6.7.

Table 4: Overview of the measures of within-precinct inequality

Dimension	Proxy	Data	Level	Indices
Income	Monthly per capita household income	Income from: salaries/wages and net business profits rental income royalties, interests and dividends pensions, welfare and grants allowances (family or other individuals)	Households	Gini index Mean log deviation Theil index
Education	Highest level of education completed	17 ranked categories: No schooling Grade 1/Sub A Grade 2/Sub B Grade 3/Std 1 Grade 4/Std 2 Grade 5/Std 3 Grade 6/Std 4 Grade 7/Std 5 Grade 8/Std 6/Form 1 Grade 9/Std 7/Form 2 Grade 10/Std 8/Form 3/NTCI Grade 11/Std 9/Form 4/NTCII Grade 12/Std 10/Form 5/NTCIII/Matric Certificate/Diploma (w/o Grade 12) Certificate/Diploma (w/ Grade 12) Undergraduate degree Postgraduate degree	Individuals aged ≥18	Jenkins index Cowell-Flachaire-Indices (α=0/0.25/0.5/0.75/0.9)
Housing	Dwelling type	7 ranked categories: Informal dwelling Traditional dwelling Dwelling in backyard Flatlet Flat in block of flats Semi-detached / cluster / town house Freestanding house	Households	Jenkins index Cowell-Flachaire-Indices (α=0/0.25/0.5/0.75/0.9)
Culture	Population group	4 non-ranked categories: Black African Coloured Indian or Asian White	Individuals	Normalized generalized variance

Source: Author

Inequality between police precincts

The basic idea of between-precinct inequality is to compare the 'level' of an outcome of interest in one precinct, e.g. income, with the 'level' in all adjacent precincts. This approach implies that individuals identify themselves strongly with 'their' precinct and rather than comparing themselves with other individuals in this precinct, they compare themselves (and their peers) to individuals from adjacent precincts. The analysis of between-precinct inequality is restricted to the dimension of income, since the measurement is much more straightforward for continuous as compared to ordinal variables. I create three alternative measures: 1) The ratio of mean per capita household income in precinct i to mean per capita household income across all adjacent precincts, 2) a dummy that equals one if this ratio is larger than one, and 3) a dummy that equals one if precinct i has the highest mean income as compared to all adjacent precincts. All three measures capture how 'rich' (in terms of income) a precinct is compared to its neighbors.

Intra- and inter-racial inequality

Given South Africa's history of racial discrimination and the persistent racial disparities, I am interested whether inequality between (inter-racial) or within population groups (intra-racial) is a better predictor of local crime rates. Pyatt (1976) proposed a simple way of decomposing the Gini index into three elements, namely, between-, overlap-, and within- group inequality, i.e.:

$$G = G_b + G_o + G_w. \quad (5)$$

Between-group inequality (G_b) is defined as the Gini index if all members of one group had the same income, i.e. the group's mean income, and within group-inequality (G_w) depends on the Gini-index within the groups. G_o stems from overlapping income ranges between the groups, i.e. if income ranges do not overlap, it is zero. None of the inequality components can be negative. Since the suitability of the Gini index for inequality decompositions has been criticized, mainly due to the unclear interpretation of the overlap term, I also use inequality decompositions based on GE(0) and GE(1), that can be additively decomposed solely into a within- and between-component, as shown by Shorrocks (1984).

4.4 Control variables

In order to mitigate the issue of omitted variable bias, I include a range of precinct-level control variables that likely affect both local inequality and crime rates. The first group of control variables can be classified as 'economic' determinants of crime, following Becker (1968), while the second group comprises social and demographic characteristics, more relevant in the sociological literature. Similar to Demombynes and Özler (2005), I proxy the returns to crime with mean household income, the cost of crime with population density (inhabitants per square kilometer) and the returns from legal work of the poor with the unemployment rate (share of unemployed among working age population).[24] Precincts with higher mean household income are likely to have accumulated more assets which can be targets of criminals (higher returns to crime), a higher population density decreases conviction probability (and thus costs of crime), and a higher unemployment rate implies lower incomes for the poor (returns from legal work). Following Kelly (2000), I also control for the share of adolescents (individuals aged 15 to 30), as this group has been shown to be more prone to commit crimes, for the share of female-headed households, a common measure of family instability, and for the share of recently (past five years) moved individuals (from other provinces), a common measure of social cohesion. In addition, I control for urbanization (share of urban households), as crime levels are generally higher in urban areas. Lastly, at least in some regions of South Africa, especially in the provinces of Western Cape and Eastern Cape, street gangs are important sources of criminal activity. As there is no data on gang prevalence available, I proxy gang prevalence with the share of the coloured population. Petrus (2015) documents that the majority of gangs is found within coloured communitities, and, referring to Bangstad (2005, pp.196–197), he also states that coloured gangs are not only more numerous, but also more 'durable' and more 'pervasive' as compared to black African gangs.

4.5 Summary statistics

Table 5 on the next page provides summary statistics on crime rates, inequality indices, and control variables for 2001 and 2011. Both crime rates and inequality indices, especially income and education inequality, have very high mean values, but also a large variation across precincts as indicated by their respective minimum and maximum values. While mean crime rates have significantly decreased from 2001 to 2011, inequality levels have remained rather constant. Mean incomes have almost doubled, from 1,852 Rand to 3,400 Rands, while the unemployment rate dropped and urbanization increased.

[24] Notably, Demombynes and Özler use total expenditures instead of income and the unemployment rate in the 'catchment area' (i.e. they consider also unemployment in all adjacent precincts) instead of the precinct.

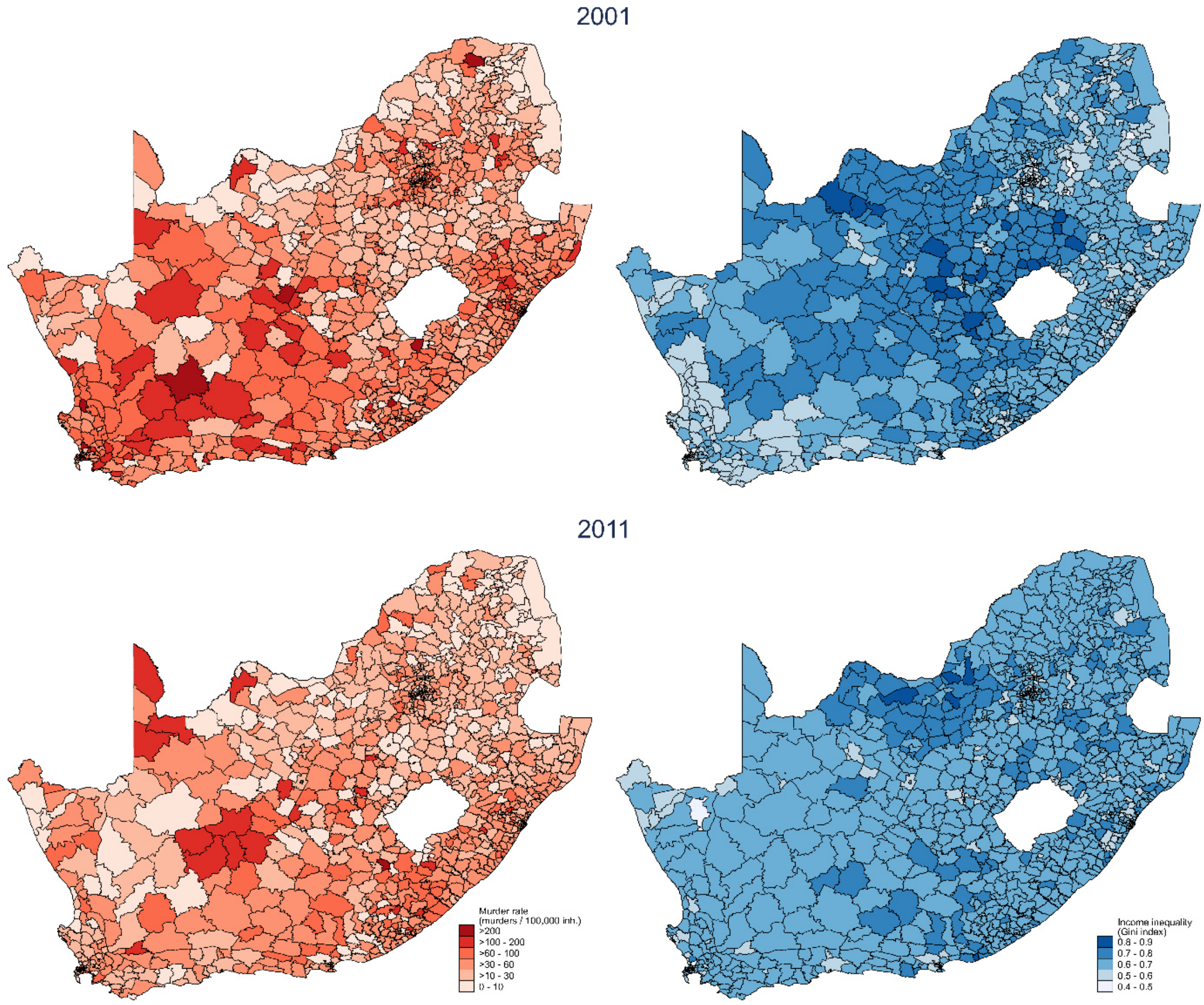
Figure 2 illustrates the spatial variation of murder rates and income inequality across South African police precincts in 2001 and 2011. Both show tremendous variation across space, with murder rates ranging from 0 to more than 200 and Gini indices from around 0.4 up to almost 0.9. Appendix D additionally provides maps for each of the three alternative inequality dimensions as well as for all analyzed property and violent crimes.

Table 5: Summary statistics on crime rates, inequality indices, and control variables

	2001			2011		
	Mean	Min	Max	Mean	Min	Max
	(1)	(2)	(3)	(4)	(5)	(6)
Crime rates						
Residential burglary	796	0.00	6,128	501	0.00	2,968
Vehicle theft	218	0.00	5,544	99	0.00	2,376
Aggravated robbery	202	0.00	5,155	148	0.00	2,264
Aggravated assault	716	0.00	3,510	444	0.00	4,036
Sexual offences	150	0.00	1,131	131	0.00	816
Murder	48	0.00	413	30	0.00	204
Inequality indices						
Income inequality (Gini)	0.63	0.41	0.85	0.64	0.46	0.82
Education inequality (Jenkins)	0.59	0.36	0.63	0.60	0.47	0.63
Housing inequality (Jenkins)	0.36	0.01	0.57	0.33	0.01	0.56
Cultural heterogeneity (NGV)	0.30	0.00	0.95	0.31	0.00	0.93
Control variables						
Mean p.c. household income	1,852	280	12,691	3,400	582	21785
Population density	691	0.10	22,018	802	0.08	23547
Unemployment rate	0.20	0.01	0.46	0.14	0.02	0.32
Share of adolescents	0.30	0.18	0.53	0.30	0.16	0.56
Share of recently moved	0.05	0.00	0.46	0.04	0.00	0.42
Share of female-headed households	0.40	0.09	0.70	0.41	0.09	0.68
Share of urban households	0.55	0.00	1.00	0.59	0.00	1.00
Share of coloureds	0.16	0.00	0.98	0.15	0.00	0.95

Notes: Income is adjusted for inflation between 2001 and 2011 and displayed in South African Rand, 2011 prices.
Source: Own calculations using data from ISS, SAPS, StatsSA, and WorldPop.

Figure 2: Spatial variation of murder rates and income inequality across South African police precincts in 2001 and 2011



Notes: The darker red, the higher the murder rate. The darker blue, the higher the Gini index.
Source: Own calculations using data from ISS, SAPS, StatsSA, and WorldPop.

5 Empirical specification

I rely on the following specification to explore the link between local inequality and crime rates:

$$crime_{icpt} = \rho W crime_{icpt} + \beta_1 \text{ineq}_{icpt} + X'_{icpt}\beta_2 + \beta_3(\gamma_p \times T_t) + \upsilon_{cp} + \varepsilon_{icpt}, \quad (6)$$

where i denotes the police precinct, c the police cluster, p the province, and t the year. The dependent variable $crime$ is the crime rate with regards to any of the six crimes introduced above. As the crime rates are to a considerable extent right-skewed and, some more than others, zero-inflated, they are transformed using the inverse hyperbolic sine (IHS) transformation (Burbidge, Magee and Robb, 1988). The main independent variable $ineq$ is any of the inequality indices introduced above. The vector X' includes all control variables introduced in the last section and the term $(I \times T)$ captures province-specific time-effects. I also control for the spatial correlation of crime, i.e. I add the spatially lagged crime rate, $Wcrime$, as a control variable. To do so, I first estimate a spatial weights matrix using the geographical information of latitude and longitude (Kondo, 2021). This matrix is then used to compute a precinct-specific spatially lagged variable that accounts for spatial dependencies across precincts.[25] In addition, I control for police cluster-fixed effects (υ) to capture any unobserved time-constant heterogeneity at the level of police clusters.[26] In principle, the panel structure also allows to control for police precinct-fixed effects, i.e. to rely only on the within-estimator and thus remove all time-constant unobserved heterogeneity. However, as the panel only includes two years and within-variation in inequality is rather low, I prefer police cluster-fixed effects; yet, I will also present results from using police precinct-fixed effects in Appendix E.

In conclusion, my identification strategy relies on the assumption that when comparing two police precincts located in the same police cluster, observed in the same year, with the same economic, social and demographic structure, any difference in inequality can be considered as random. Alternatively, anything that could cause inequalities in one police precinct and not in another is, conditional on these controls, orthogonal to the precinct's crime rates.

In this setup, two threats to identification remain. First, reverse causality, i.e. precincts with high crime rates are less attractive as place of residence so that households prefer to move out of / not to move into 'high-crime' precincts. As mobility is obviously positively correlated with income, richer households are more able to react to rising crime rates, which in turn affects the level of inequality. Second, omitted variable bias through unobserved (time-varying) precinct- or cluster-level characteristics, such as number of police officers or police expenditures. In order to strengthen identification, I will implement a whole range of robustness checks.

6 Results

6.1 The effect of inequality within precincts on crime rates

Tables 6 to 9 show the effects of inequality within police precincts with regards to income, education, housing, and culture, respectively. Each table comprises six columns, one for each of the six crime types introduced above. As the crime rates are IHS-transformed and the

---

[25] Kondo (2021) illustrates that a neighborhood might suffer from high crime rates, not just due to its own (socio-economic) characteristics, but also because nearby precincts are crime-riddled and criminals (and crimes) from these precincts spread over.

[26] The 1085 police precincts are organized in 176 police clusters, i.e. in one cluster comprises on average around six precincts.

inequality indices (and most controls) in levels, the corresponding coefficients can be interpreted similar to a log-level specification. Only mean income and population density are log-transformed and thus their coefficients are to be interpreted as elasticities.

As the set of control variables is the same for all tables shown below and their coefficients barely change in response to alternative inequality dimensions as main regressors, I first summarize the effects associated with the various control variables. The results suggest a strong, positive and highly significant elasticity between mean household income and all three property crimes, but a rather weak, negative and mostly insignificant association with violent crimes. This is consistent with one aspect of the Becker-model, namely that returns from crime are an important determinant of criminal activity. Mean household income is a good proxy for the returns from property crimes, but obviously not for violent crimes, which is confirmed by these regression results. Overall, though much smaller, the elasticity between population density and crime rates is generally positive and there is no clear difference between property and violent crime, yet, the coefficients are only significant for vehicle theft, aggravated robbery, and murder. One explanation might be that especially these crimes are facilitated when the perpetrator can quickly go into hiding. The unemployment rate is also positively correlated with crime rates and coefficients are sizable and mostly significant, suggesting that a decrease in the returns from legal work leads to more criminal activity. As unemployment captures not only low wages of the poor, but is also tied with frustration, drug abuse etc., it is not surprising that this mechanism holds for both property and violent crimes. Turning to the social and demographic control variables, also their coefficients generally point into the expected direction. The share of adolescents is (with the exception of burglary) positively and highly significantly correlated with crime rates, especially with assault and murder. Similarly, the share of recently moved individuals is positively associated with crimes rates across all crime types and this association is mostly significant, indicating the relevance of social cohesion. For the share of female-headed households, coefficients are generally positive, but (except for robbery) rather small in magnitude and often insignificant. The regressions also confirm that crimes are generally higher in urban areas, though for murder the coefficient is insignificant. Lastly, it is interesting to see that the share of coloureds has no meaningful correlation with property crimes, but is significantly and positively related with all types of violent crimes, supporting its usefulness as a proxy for gang prevalence. Most gangs engage in violence to defend their territory (‘turf wars’) and their primary source of income is drug trade, not burglary, theft, or robbery (Petrus and Kinnes, 2019).

Table 6: Regressions of local crime rates on income inequality within police precincts

Independent variables	Property crimes			Violent crimes		
	Residential burglary	Vehicle theft	Aggravated robbery	Aggravated assault	Sexual offences	Murder
	(1)	(2)	(3)	(4)	(5)	(6)
Income inequality	0.359	0.259	-0.351	1.671***	1.380***	2.491***
(Gini index)	(0.4286)	(0.7884)	(0.8194)	(0.4223)	(0.4960)	(0.7943)
Returns to crime	0.781***	1.481***	0.676***	-0.126**	-0.035	-0.039
(ln mean p.c. hh income)	(0.0628)	(0.1002)	(0.1053)	(0.0640)	(0.0729)	(0.1014)
Costs of crime	-0.000	0.136***	0.140***	0.003	0.033*	0.126***
(ln population density)	(0.0163)	(0.0312)	(0.0310)	(0.0140)	(0.0194)	(0.0287)
Returns from work	0.995***	0.745	1.401**	0.684**	0.775*	0.414
(unemployment rate)	(0.3382)	(0.6950)	(0.7007)	(0.3432)	(0.4133)	(0.6755)
Crime prone groups	-0.744	1.964**	2.609***	3.760***	1.861**	3.353***
(share of adolescents)	(0.7630)	(0.9999)	(0.9906)	(0.5333)	(0.8488)	(1.0113)
Social cohesion	1.211**	1.853*	2.013**	0.983***	1.674***	2.259**
(share of recently moved)	(0.4719)	(1.0088)	(0.8667)	(0.3773)	(0.5850)	(0.9374)
Family instability	0.701*	0.884	3.387***	-0.051	0.231	0.651
(share of female-headed hh)	(0.3936)	(0.5697)	(0.5425)	(0.3599)	(0.3993)	(0.5562)
Urbanization	0.595***	0.905***	0.863***	0.605***	0.516***	0.248
(share of urban hh)	(0.1139)	(0.1791)	(0.1752)	(0.1090)	(0.1269)	(0.1664)
Gang prevalence	0.155	-0.016	-0.131	0.867***	0.496***	0.616**
(share of coloureds)	(0.1229)	(0.2621)	(0.2396)	(0.1191)	(0.1508)	(0.2406)
Province-specific time-effects	yes	yes	yes	yes	yes	yes
Spatial lag	yes	yes	yes	yes	yes	yes
Cluster-fixed effects	yes	yes	yes	yes	yes	yes
R-Squared	0.602	0.660	0.604	0.481	0.274	0.318
Observations	2,170	2,170	2,170	2,170	2,170	2,170

Notes: Each column refers to one regression. Dependent variables are IHS-transformed crime rates (crimes per 100,000 inhabitants). The spatial lag of the crime rate is instrumented with the spatial lags of first, second and third order of all independent variables. Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.
Source: Own calculations using data from ISS, SAPS, StatsSA, and WorldPop.

The first row of Table 6 presents the coefficients associated with income inequality as measured by the Gini index and reveals an interesting pattern. For all property crimes, the coefficients are small and statistically insignificant. This suggests that, conditional on the returns and costs of crime and the returns from work as well as on social and demographic characteristics, income inequality does not contribute to higher rates of property crimes. Importantly, this does not invalidate Becker's theory, as he focused on the differential returns from crime versus work, for which the coefficients are as expected. One could infer from this pattern that only one specific type of income inequality, i.e. the difference between the mean income (how 'rich' is an area on average, i.e. how high are the returns from crime) and the bottom incomes (how much do the poor earn, i.e. how high are their returns from legal work) is an adequate predictor of property crime rates. In contrast to this, income inequality seems to be an important determinant of violent crime as the coefficients are positive, large and highly significant for assault, sexual offences, and murder. An increase in the Gini index by 0.1 (i.e. by 1.4 SD) is associated with an increase in the crime rate by 17% for assault, 14% for sexual violence, and even 25% for murder. To illustrate this magnitude, a 0.1 increase in the Gini index is comparable to a shift in income inequality from the level as of 2011 in Camps Bay (0.47) to that in Mamelodi (0.57), Mamelodi to East London (0.67), or East London to Hazyview (0.77). Comparing two (otherwise equal) police precincts, one with an inequality level similar to Camps Bay and the other similar to Hazyview, murder rates are predicted to be 75% higher in the latter as compared to the former. Overall, the results provide the most support for one aspect of strain theory – that relative deprivation causes frustration and alienation from society, which culminate in violent behavior.

I now turn to the results concerning education inequality, proxied by the Jenkins index. The first row in Table 7 shows positive and statistically significant coefficients across all columns, indicating a positive relationship between education inequality and all types of crimes (as postulated in H3). However, the coefficients are more precisely estimated for violent crimes and almost double in magnitude. While a 0.1 increase in education inequality increases property crime rates by about 23% to 25%, it increases violent crime rates by about 42% to 45%. As both the Jenkins index and the Gini index are bounded between 0 and 1, it is possible and meaningful to compare these effects to those obtained from income inequality. This comparison suggests that education inequality might contribute more to crime rates than income inequality. Yet, the two patterns are not too different as in both cases violent crime rates respond much stronger to increases in inequality. Given the high returns to education on the South African labor market, this similarity is not too surprising, as inequality in educational attainment contributes considerably to income inequality. To sum up, the results regarding education inequality support social disorganization theory as well as strain theory. All types of crimes are affected, and it is not possible to infer whether this is caused by weakened social control (as suggested by social disorganization theory) or by increased relative deprivation (as suggested by strain theory) or both.

Table 7: Regressions of local crime rates on education inequality within police precincts

Independent variables	Property crimes			Violent crimes		
	Residential burglary	Vehicle theft	Aggravated robbery	Aggravated assault	Sexual offences	Murder
	(1)	(2)	(3)	(4)	(5)	(6)
Education inequality	2.253***	2.478*	2.480**	4.190***	4.512***	4.210***
(Jenkins index)	(0.7125)	(1.3398)	(1.1982)	(0.8689)	(0.9596)	(1.3541)
Returns to crime	0.779***	1.475***	0.649***	-0.095	-0.016	0.020
(ln mean p.c. hh income)	(0.0561)	(0.1000)	(0.0999)	(0.0595)	(0.0664)	(0.0951)
Costs of crime	-0.007	0.133***	0.143***	-0.025*	0.007	0.087***
(ln population density)	(0.0150)	(0.0286)	(0.0294)	(0.0133)	(0.0179)	(0.0269)
Returns from work	0.886***	0.590	1.098	0.725**	0.707*	0.651
(unemployment rate)	(0.3125)	(0.6780)	(0.6724)	(0.3166)	(0.3981)	(0.6322)
Crime prone groups	-0.765	1.911**	2.363**	4.008***	1.958**	3.845***
(share of adolescents)	(0.7268)	(0.9710)	(0.9535)	(0.5368)	(0.8289)	(1.0057)
Social cohesion	1.174***	1.846*	2.127**	0.697*	1.430***	1.803**
(share of recently moved)	(0.4447)	(0.9944)	(0.8586)	(0.3601)	(0.5531)	(0.9200)
Family instability	0.728*	0.892	3.387***	0.050	0.365	0.785
(Share of female-headed hh)	(0.3900)	(0.5742)	(0.5452)	(0.3658)	(0.3959)	(0.5534)
Urbanization	0.586***	0.898***	0.871***	0.563***	0.479***	0.185
(share of urban hh)	(0.1101)	(0.1798)	(0.1726)	(0.1043)	(0.1244)	(0.1656)
Gang prevalence	0.101	-0.073	-0.180	0.751***	0.372**	0.489**
(share of coloureds)	(0.1221)	(0.2615)	(0.2392)	(0.1171)	(0.1492)	(0.2399)
Province-specific time-effects	yes	yes	yes	yes	yes	yes
Spatial lag	yes	yes	yes	yes	yes	yes
Cluster-fixed effects	yes	yes	yes	yes	yes	yes
R-Squared	0.604	0.660	0.605	0.487	0.281	0.319
Observations	2,170	2,170	2,170	2,170	2,170	2,170

Notes: Each column refers to one regression. Dependent variables are IHS-transformed crime rates (crimes per 100,000 inhabitants). The spatial lag of the crime rate is instrumented with the spatial lags of first, second and third order of all independent variables. Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.
Source: Own calculations using data from ISS, SAPS, StatsSA, and WorldPop.

Turning to housing inequality, we again see a differential pattern between property and violent crimes. While the coefficients are positive for all crimes, they are much smaller for violent crimes (null for assault) and statistically significant only for property crimes. This pattern indicates that housing inequality is an important determinant of property crime rates, but not of violent crime rates (equivalent to H1). Intuitively, this seems plausible as housing inequality is manifested in plainly visible differences with regards to property (e.g. informal shacks close to formal apartments or houses) and might thus trigger crimes specifically targeted at property. The returns from property crimes are 'omnipresent' and the path on how to obtain them is clear. At the same time, housing inequality (as measured here via different dwelling types) can by definition not be as stark as income inequality, i.e. its potential for generating serious frustration and hence violence is somewhat restricted. The results thus support Becker's theory and also one aspect of strain theory, namely the alleviation of relative deprivation through asset accumulation by illegal means. Overall, the effect sizes are rather modest, ranging from a 6% increase in burglary rates to a 15% increase in robbery rates for a 0.1 increase in housing inequality (see Table 8).

Table 8: Regressions of local crime rates on housing inequality within police precincts

Independent variables	Property crimes			Violent crimes		
	Residential burglary	Vehicle theft	Aggravated robbery	Aggravated assault	Sexual offences	Murder
	(1)	(2)	(3)	(4)	(5)	(6)
Housing inequality	0.612***	1.236***	1.506***	0.009	0.329	0.495
(Jenkins index)	(0.2091)	(0.3716)	(0.3779)	(0.1727)	(0.2579)	(0.3507)
Returns to crime	0.775***	1.454***	0.622***	-0.070	0.003	0.031
(ln mean p.c. hh income)	(0.0606)	(0.1001)	(0.1008)	(0.0591)	(0.0697)	(0.0962)
Costs of crime	-0.008	0.129***	0.142***	-0.023*	0.010	0.089***
(ln population density)	(0.0151)	(0.0287)	(0.0292)	(0.0138)	(0.0183)	(0.0269)
Returns from work	1.025***	0.694	1.178*	1.095***	1.077***	0.973
(unemployment rate)	(0.3148)	(0.6539)	(0.6534)	(0.3108)	(0.3895)	(0.6132)
Crime prone groups	-0.867	1.593	1.961**	4.206***	2.125**	3.905***
(share of adolescents)	(0.7474)	(0.9777)	(0.9602)	(0.5425)	(0.8497)	(1.0190)
Social cohesion	1.177***	1.894*	2.212***	0.612*	1.394**	1.776*
(share of recently moved)	(0.4393)	(0.9702)	(0.8453)	(0.3540)	(0.5518)	(0.9063)
Family instability	0.776**	0.993*	3.495***	0.039	0.339	0.797
(Share of female-headed hh)	(0.3957)	(0.5739)	(0.5456)	(0.3638)	(0.4043)	(0.5567)
Urbanization	0.582***	0.889***	0.858***	0.565***	0.479***	0.182
(share of urban hh)	(0.1091)	(0.1780)	(0.1717)	(0.1063)	(0.1242)	(0.1658)
Gang prevalence	0.180	0.043	-0.050	0.836***	0.491***	0.608**
(share of coloureds)	(0.1202)	(0.2558)	(0.2359)	(0.1178)	(0.1474)	(0.2422)
Province-specific time-effects	yes	yes	yes	yes	yes	yes
Spatial lag	yes	yes	yes	yes	yes	yes
Cluster-fixed effects	yes	yes	yes	yes	yes	yes
R-Squared	0.605	0.662	0.609	0.472	0.271	0.316
Observations	2,170	2,170	2,170	2,170	2,170	2,170

Notes: Each column refers to one regression. Dependent variables are IHS-transformed crime rates (crimes per 100,000 inhabitants). The spatial lag of the crime rate is instrumented with the spatial lags of first, second and third order of all independent variables. Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.
Source: Own calculations using data from ISS, SAPS, StatsSA, and WorldPop.

Lastly, Table 9 presents the effects associated with cultural heterogeneity. The coefficients are positive and statistically highly significant across all six crimes (equivalent to H3). Precincts with a higher diversity in terms of population groups thus experience a generally higher level of criminality as compared to (otherwise equal) precincts with a more homogenous population. The effects are the largest for theft and robbery (14% and 11%) and quite a bit smaller for all other crimes (from 3.7% for burglary to 4.8% for murder). These results strongly support social disorganization theory, specifically the claim that cultural heterogeneities impede social cohesion and thus the 'collective efficacy' of a community and its ability to prevent crime through social control.

Table 9: Regressions of local crime rates on cultural heterogeneity within police precincts

Independent variables	Property crimes			Violent crimes		
	Residential burglary	Vehicle theft	Aggravated robbery	Aggravated assault	Sexual offences	Murder
	(1)	(2)	(3)	(4)	(5)	(6)
Cultural heterogeneity	0.365***	1.360***	1.132***	0.419***	0.381***	0.479**
(Generalized variance, norm.)	(0.1106)	(0.2117)	(0.2132)	(0.1073)	(0.1267)	(0.1996)
Returns to crime	0.720***	1.218***	0.447***	-0.153**	-0.065	-0.050
(ln mean p.c. hh income)	(0.0629)	(0.1111)	(0.1116)	(0.0674)	(0.0760)	(0.1037)
Costs of crime	-0.000	0.160***	0.171***	-0.017	0.013	0.096***
(ln population density)	(0.0147)	(0.0282)	(0.0286)	(0.0138)	(0.0175)	(0.0268)
Returns from work	1.185***	1.188*	1.644**	1.213***	1.201***	1.156*
(unemployment rate)	(0.3160)	(0.6527)	(0.6505)	(0.3097)	(0.3944)	(0.6119)
Crime prone groups	-0.613	2.183**	2.639***	4.250***	2.192***	4.118***
(share of adolescents)	(0.7229)	(0.9669)	(0.9552)	(0.5393)	(0.8273)	(1.0178)
Social cohesion	0.834*	0.703	1.240	0.271	1.014*	1.329
(share of recently moved)	(0.4417)	(1.0018)	(0.8663)	(0.3444)	(0.5442)	(0.9322)
Family instability	0.789**	1.093*	3.552***	0.115	0.454	0.857
(Share of female-headed hh)	(0.3955)	(0.5647)	(0.5417)	(0.3664)	(0.4133)	(0.5572)
Urbanization	0.558***	0.793***	0.770***	0.531***	0.448***	0.149
(share of urban hh)	(0.1095)	(0.1744)	(0.1700)	(0.1032)	(0.1242)	(0.1658)
Gang prevalence	0.080	-0.277	-0.343	0.757***	0.391***	0.487**
(share of coloureds)	(0.1259)	(0.2692)	(0.2455)	(0.1187)	(0.1513)	(0.2471)
Province-specific time-effects	yes	yes	yes	yes	yes	yes
Spatial lag	yes	yes	yes	yes	yes	yes
Cluster-fixed effects	yes	yes	yes	yes	yes	yes
R-Squared	0.604	0.664	0.615	0.477	0.265	0.315
Observations	2,170	2,170	2,170	2,170	2,170	2,170

Notes: Each column refers to one regression. Dependent variables are IHS-transformed crime rates (crimes per 100,000 inhabitants). The spatial lag of the crime rate is instrumented with the spatial lags of first, second and third order of all independent variables. Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.
Source: Own calculations using data from ISS, SAPS, StatsSA, and WorldPop.

6.2 Nonlinearity of within-precinct inequality effects

Section 6.1 presented the effects of socio-economic inequalities within precincts on local crime rates. In order to explore a potential nonlinearity of these effects, I re-ran the main regressions including the squared term of the respective inequality measure. Table 10 summarizes the results from these regressions. For the sake of space, the coefficients of the control variables are not shown. Columns indicate the crime category and rows the inequality dimension. In addition, Figure 3 illustrates the nonlinearity of the effects graphically. The various graphs plot predicted crime rates against the different inequality measures, based on the coefficients in Table 10, with inequality values ranging from the sample minimum to the sample maximum and the values of all control variables held constant at the sample mean.

The linear regressions in Section 6.1 indicated a positive and significant effect of income inequality on violent, but not property crimes. Looking at the first two rows in Table 10 reveals an interesting pattern. For all property crimes, the coefficients are large, positive and significant for the linear term of the Gini index and large, negative and significant for its squared term. This indicates that there are indeed significant positive effects of income inequality on property crimes, yet, these effects are strongly nonlinear and diminishing with increasing levels of income inequality. As shown in Figure 3, the relationship between income inequality and property crimes follows an inverted u-shape. A potential explanation for this might be that when precincts exceed a certain level of income inequality, their (richer) inhabitants invest increasingly in protective measures, such as neighborhood watches, armed response services, and anti-burglary protection. Possibly, also public resources, such as police patrols, might be directed towards high-inequality precincts. Overall, this means that property crime rises when inequality increases, but at some point, decreases again in response to prevention through protective investments. To a lesser extent, the results also suggest some curvilinearity for aggravated assault, while there is no evidence for a meaningful nonlinearity in the cases of sexual offences and murder. This is not surprising, as a large share of sexual offences and murders are committed within the extended family or circle of acquaintances and can thus generally not be prevented by the above-described protective investments.

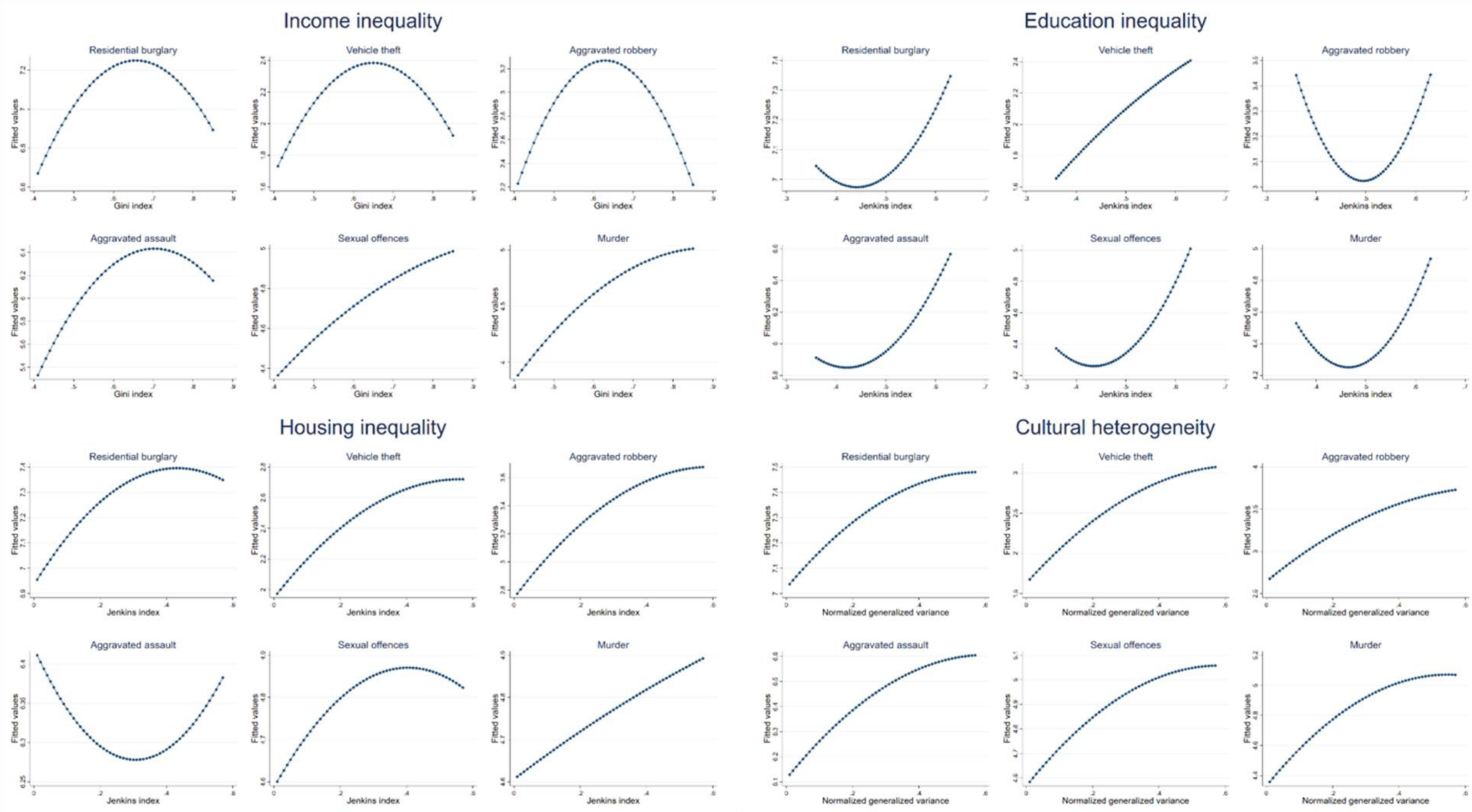
For housing inequality, the linear regressions showed positive and significant effects only on property crimes. The nonlinear regressions suggest that these are also diminishing, yet, not as strong as in the case of income inequality, and to a meaningful extent only for residential burglary. The mechanism behind might again be increasing protective investments at higher levels of housing inequality, potentially more targeted to prevent burglaries, as these might be the most obvious targets in such settings.

Table 10: Regressions of local crime rates on inequality within police precincts, quadratic specification

Independent variables	Property crimes			Violent crimes		
	Residential burglary	Vehicle theft	Aggravated robbery	Aggravated assault	Sexual offences	Murder
	(1)	(2)	(3)	(4)	(5)	(6)
Income inequality						
Gini index	12.533*** (3.3088)	14.801** (7.4130)	27.172*** (7.5962)	18.084*** (5.4061)	3.458 (4.3813)	8.801 (7.4232)
Gini index, squared	-9.544*** (2.6238)	-11.397* (5.9506)	-21.581*** (6.0974)	-12.862*** (4.3195)	-1.625 (3.4610)	-4.953 (5.9455)
Education inequality						
Jenkins index	-9.339 (6.4484)	5.641 (20.2711)	-22.703 (14.3255)	-13.926* (8.2815)	-17.274 (11.4806)	-23.455 (19.2121)
Jenkins index, squared	10.562* (5.9995)	-2.884 (18.1716)	22.938* (13.1210)	16.515** (7.4012)	19.834* (10.3088)	25.214 (17.2058)
Housing inequality						
Jenkins index	2.131** (0.8900)	2.760* (1.6309)	3.129* (1.6662)	-0.927 (0.7453)	1.402 (1.2167)	0.599 (1.5801)
Jenkins index, squared	-2.463* (1.3194)	-2.464 (2.4010)	-2.621 (2.4347)	1.510 (1.1186)	-1.732 (1.7597)	-0.170 (2.3183)
Cultural heterogeneity						
Generalized variance (norm.)	1.565*** (0.3039)	4.559*** (0.5721)	3.234*** (0.5843)	1.635*** (0.2976)	1.685*** (0.3523)	2.705*** (0.5859)
Generalized variance (norm.), squared	-1.336*** (0.3179)	-3.562*** (0.5740)	-2.338*** (0.5956)	-1.354*** (0.3184)	-1.447*** (0.3746)	-2.473*** (0.5891)
Province-specific time-effects	yes	yes	yes	yes	yes	yes
Spatial lag	yes	yes	yes	yes	yes	yes
Cluster-fixed effects	yes	yes	yes	yes	yes	yes
Full set of controls	yes	yes	yes	yes	yes	yes
Observations	2,170	2,170	2,170	2,170	2,170	2,170

Notes: Each column contains four regressions; one for each inequality dimension and including the linear and squared term of the respective inequality index. The full set of controls includes ln mean p.c. household income, ln population density, unemployment rate, share of adolescents, share of recently moved, share of female-headed hh, share of urban hh, share of coloureds. Dependent variables are IHS-transformed crime rates (crimes per 100,000 inhabitants). The spatial lag of the crime rate is instrumented with the spatial lags of first, second and third order of all independent variables. Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.
Source: Own calculations using data from ISS, SAPS, StatsSA, and WorldPop.

Figure 3: Illustrations of the nonlinear relationships between socio-economic inequalities and crime rates



Notes: Predicted crimes rates (IHS-transformed) are on the y-axis and inequality measures (within sample range) on the x-axis. Control variables were held constant at the sample mean for crime rate prediction. The plots should be interpreted with simultaneous consideration of the regressions results presented in Table 10.
Source: Own calculations using data from ISS, SAPS, StatsSA, and WorldPop.

The results for education inequality are at a first glance rather surprising, with the linear term being generally negative (but insignificant) and the squared term positive (significant only at 10%-level). Yet, in all cases, the turning point is within the first percentile of education inequality, meaning that for 99% of the observations the effects are positive and exhibit no meaningful nonlinearity.

Section 6.1 also revealed positive and significant effects of cultural heterogeneity on all violent and property crimes. The last two rows of Table 10 present strong evidence that these effects are diminishing with increasing levels of cultural heterogeneity, which is also clearly visible in Figure 3. An intuitive explanation for the curvilinearity along the lines of social disorganization and strain theory might be as follows. Culturally completely homogenous precincts (i.e. only one population group) have high levels of social control and thus low crime rates. Having another population group of minority status in a precinct might strongly reduce social control capacities if the minority group is not well-integrated into the community or if inter-group trust is low. In addition, members of the minority might experience some form of relative deprivation due to social exclusion. As a consequence, crime rates are higher, and they rise further if the minority groups are bigger or if more minority groups are added. Yet, as precincts approach even higher values of cultural heterogeneity, i.e. a uniform distribution across all population groups, this effect diminishes. First, because minorities are not confronted with one dominant majority, leaving less scope for exclusion and relative deprivation. Second, because communities with such high levels of cultural heterogeneity might be more tolerant, maybe because they have historically been more diverse, and are able to exert social control across population groups.

6.3 Intra- or inter-racial inequality?

This section serves to assess whether it is inequality within population groups, i.e. intra-racial inequality, or inequality between population groups, i.e. inter-racial inequality, or both, that drives the positive association between local inequality and crime rates. For this, I decomposed the indices of income inequality (within police precincts) into their within- and between-group components and again regressed local crime rates on income inequality, yet, replacing overall inequality with intra- and inter-racial inequality.[27] Since I did not find a significant linear relationship between income inequality and property crimes, but evidence for an inverted u-

[27] The decomposed Gini index additionally comprises a residual called 'overlap inequality'.

shape, I run the 'decomposed regressions' with the linear and squared terms of all inequality components for property crimes. For violent crimes, I adhere to the linear specification. Table 11 on the next page presents the results, subdivided by crime types (columns) and inequality indices (rows).

Starting with property crimes, the coefficients for the decomposed Gini index suggest that the inverted u-shape found in Section 6.2 is driven primarily by inter-racial inequality and that intra-racial inequality does not affect rates of property crimes significantly. Looking at the two alternative measures of generalized entropy, also intra-racial inequality is significantly associated with property crime, yet, only with two out of three crimes, namely, residential burglary and aggravated robbery. Overall, while the role of intra-racial inequality can't be neglected, it seems like inequality between population groups is the major contributor to the inverted u-shaped association between local inequality and rates of property crimes. The idea put forward in the last section, that very high levels of inequality are accompanied by protection efforts of richer residents, might apply particularly to the case of inter-racial inequality. High inter-racial inequality is often concomitant with geographical segregation of population groups, e.g. a small rich (predominantly white) neighborhood located next to a populous very poor (predominantly black) township in the same precinct. While the richer neighborhood provides high returns to property crime for potential offenders from the township, the returns are difficult to realize, as the rich neighborhood is completely sealed off from the township inhabitants.

Turning to violent crimes, the coefficients for the decomposed Gini index suggest that both intra-and inter-racial inequality are significantly related with rates of violent crimes. However, when considering the measures of generalized entropy instead, the coefficients for inter-racial inequality are insignificant for all three violent crimes. These results suggest that, overall, the positive association between local inequality and rates of violent crimes is driven primarily by inequality within population groups and less by inequality between them. Potentially, this could partly be explained by South Africa's long history of inter-racial inequality and the more recent surge in intra-racial inequality. Put simply, the country's historically oppressed population groups might have grown accustomed to high inter-racial disparities and see them, to some extent, as given. In addition, people might compare their living standard more with that of members of their own population group (with whom they interact more frequently) and thus rising inequality within that group could be especially frustrating, causing feelings of relative deprivation and triggering violent behavior.

Table 11: Regressions of local crime rates on income inequality within police precincts, decomposed into intra-racial and inter-racial inequality

Independent variables	Property crimes			Violent crimes		
	Residential burglary	Vehicle theft	Aggravated robbery	Aggravated assault	Sexual offences	Murder
	(1)	(2)	(3)	(4)	(5)	(6)
Gini index						
Intra-racial	-0.666 (1.0946)	-2.844 (2.3399)	0.841 (2.3503)	1.880*** (0.6582)	1.281* (0.6706)	2.807** (1.0950)
Intra-racial, squared	0.690 (1.0036)	1.566 (1.9931)	0.282 (2.0511)			
Inter-racial	1.395** (0.6798)	1.868 (1.2404)	3.217*** (1.2330)	1.801*** (0.4822)	1.420*** (0.5104)	2.657*** (0.8233)
Inter-racial, squared	-2.388** (0.9407)	-4.577** (2.1370)	-5.895*** (2.1197)			
Residual	2.760* (1.6629)	9.537*** (3.1471)	11.592*** (3.1494)	3.426*** (0.9994)	2.949*** (1.0057)	5.142*** (1.5152)
Residual, squared	-9.005* (5.3687)	-34.093*** (11.5320)	-27.558** (11.4455)			
Mean log deviation						
Intra-racial	2.146*** (0.7386)	1.065 (1.4742)	5.186*** (1.3492)	0.442 (0.2694)	0.359* (0.2083)	0.907*** (0.3450)
Intra-racial, squared	-1.434*** (0.4172)	-0.844 (0.8959)	-3.133*** (0.7826)			
Inter-racial	0.531* (0.3192)	1.556** (0.6504)	0.141 (0.6352)	0.233 (0.1648)	0.370 (0.2306)	0.414 (0.3545)
Inter-racial, squared	-0.874** (0.3563)	-2.384** (0.9274)	-2.017** (0.9183)			
Theil index						
Intra-racial	0.414 (0.3545)	0.198 (0.9866)	2.252** (0.9381)	0.376** (0.1889)	0.267 (0.1874)	0.697** (0.2959)
Intra-racial, squared	-0.498** (0.2479)	-0.218 (0.4988)	-1.040** (0.4446)			
Inter-racial	0.535** (0.2630)	1.523*** (0.5308)	0.949* (0.5215)	0.116 (0.1313)	0.169 (0.1848)	0.300 (0.2915)
Inter-racial, squared	-0.702*** (0.2578)	-2.046*** (0.6863)	-2.200*** (0.6903)			
Province-specific time-effects	yes	yes	yes	yes	yes	yes
Spatial lag	yes	yes	yes	yes	yes	yes
Cluster-fixed effects	yes	yes	yes	yes	yes	yes
Full set of controls	yes	yes	yes	yes	yes	yes
Observations	2,170	2,170	2,170	2,170	2,170	2,170

Notes: The table contains in total 18 regressions, corresponding to the six crimes and three alternative indices of income inequality. Inequality indices are decomposed into their intra-and interracial component (plus the residual for the Gini index). For all property crimes, the squared term of all inequality components is additionally included. The full set of controls includes ln mean p.c. household income, ln population density, unemployment rate, share of adolescents, share of recently moved, share of female-headed hh, share of urban hh, share of coloureds. Dependent variables are IHS-transformed crime rates (crimes per 100,000 inhabitants). The spatial lag of the crime rate is instrumented with the spatial lags of first, second and third order of all independent variables. Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

Source: Own calculations using data from ISS, SAPS, StatsSA, and WorldPop.

6.4 The effect of inequality between precincts on crime rates

I now turn to the effects of income inequality between precincts on local crime rates. The results are displayed in Table 12, for all crimes (columns) and for the different measures of between-precinct inequality (rows).

Table 12: Regressions of local crime rates on income inequality between police precincts

Independent variables	Property crimes: Residential burglary (1)	Property crimes: Vehicle theft (2)	Property crimes: Aggravated robbery (3)	Violent crimes: Aggravated assault (4)	Violent crimes: Sexual offences (5)	Violent crimes: Murder (6)
Income ratio						
Mean income own precinct / mean income across adj. prec.	-0.051 (0.0362)	-0.001 (0.0726)	0.172** (0.0763)	-0.057 (0.0387)	-0.031 (0.0523)	0.026 (0.0675)
Income ratio, squared spec.						
Mean income own precinct / mean income across adj. prec.	-0.132 (0.0916)	0.271* (0.1593)	0.490*** (0.1685)	-0.043 (0.1005)	-0.007 (0.0978)	0.213 (0.1656)
Mean income own precinct / mean inc. acr. adj. prec., sq.	0.019 (0.0163)	-0.063* (0.0324)	-0.077** (0.0366)	-0.003 (0.0250)	-0.006 (0.0234)	-0.045 (0.0382)
Richer precinct						
Mean income own precinct > mean inc. acr. adj. prec. (=1)	0.013 (0.0500)	0.089 (0.0921)	0.180* (0.0954)	0.038 (0.0523)	-0.016 (0.0506)	0.138 (0.0845)
Richest precinct						
Mean income own precinct > mean inc. each adj. prec. (=1)	-0.154** (0.0638)	-0.209* (0.1086)	-0.150 (0.1157)	-0.084 (0.0590)	-0.145* (0.0828)	-0.164 (0.1070)
Province-specific time-effects	yes	yes	yes	yes	yes	yes
Spatial lag	yes	yes	yes	yes	yes	yes
Cluster-fixed effects	yes	yes	yes	yes	yes	yes
Full set of controls	yes	yes	yes	yes	yes	yes
Observations	2,170	2,170	2,170	2,170	2,170	2,170

Notes: The table contains in total 24 regressions, corresponding to the six crimes and four alternative specifications of between-precinct inequality. Dependent variables are IHS-transformed crime rates (crimes per 100,000 inhabitants). The spatial lag of the crime rate is instrumented with the spatial lags of first, second and third order of all independent variables. Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.
Source: Own calculations using data from ISS, SAPS, StatsSA, and WorldPop.

Starting with the income ratio (ratio of mean income of own precinct to mean income across all adjacent precincts) as measure of between-precinct inequality, across five of the six crimes, coefficients are small and insignificant. The only exception is aggravated robbery, for which the coefficient is positive and statistically significant, suggesting that precincts that are relatively ‘rich’ as compared to their neighbors suffer from higher rates of aggravated robbery. To illustrate, going from a precinct that is ‘as rich’ as its neighbors to a precinct that is ‘twice as rich' as its neighbors rises aggravated robbery rates by around 17%. The effect holds when using a dummy indicating that the ratio is larger than 1, yet, there is no significant effect for being the richest among all precincts. On the contrary, being the richest among all adjacent precincts is generally associated with lower crime rates, however, the coefficient is often insignificant. This might be explained by those 'richest' precincts sealing themselves off from their poor neighbors, e.g. through gated communities or using alternative protective measures. Re-running the regressions with the income ratio plus its squared term (Col. 2 and 3) supports this claim – while the linear term is positive and significant, the squared term is negative and significant. In this specification, the inequality coefficients now also show the same pattern for vehicle theft (also significant, but smaller). Overall, inequality between precincts seems to play a smaller role for local crime rates in comparison to inequality within precincts. Yet, it confirms to some extent the idea put forward in H2, i.e. that criminals (from poorer neighborhoods) are attracted by high returns to property crimes in relatively rich precincts.

6.5 The role of catchment area conditions

So far, the analysis assumed that the crime level in a given precinct is determined by socio-economic and demographic conditions in this specific precinct. Yet, it can't be ruled out that not all crimes are committed in the 'home precinct' of the offender, but also in neighboring precincts. Furthermore, the police precinct is a somewhat 'artificial construct' as it doesn't necessarily coincide with the boundaries of neighborhoods, towns or similar local communities. Hence, the reference group of the precinct inhabitants might expand beyond the borders of their precinct and include inhabitants of neighboring precincts. To account for these circumstances, I re-calculated for each precinct all inequality indices and control variables, considering all inhabitants in the precinct itself and in all adjacent precincts. This does not change the level of observation, which is still the precinct; it simply extends the borders of the precinct to what Demombynes and Özler (2005) refer to as the 'catchment area'. I then re-ran the main regressions with catchment area-based inequality indices and control variables.[28] Importantly, local crime rates were not adapted to include crimes in the catchment area.

As visible in Table 13 on the next page, the associations between precinct-level crime rates and inequalities in the corresponding catchment area are rather weak and mostly insignificant. With

[28] Demombynes and Özler (2005) only calculate inequality based on the catchment area, but do not consider the catchment area for the control variables (except unemployment, which is calculated for the catchment area already in their main specification).

the exception of education inequality, which displays positive and highly significant coefficients for all violent crimes, almost all other coefficients are statistically insignificant. This suggests that, overall, inequality in the precinct itself is a far more important contributor to precinct-level crime rates than inequality in the catchment area. One explanation in the context of South Africa might be that many police precincts cover relatively large areas (median precinct area: 560 km²). This means that in many cases the precinct already includes more than just one's 'own' neighborhood or, put differently, goes beyond the area of daily live. Extending these large precincts further to include all adjacent precincts leads to very large catchment areas (median catchment area: 5,424 km²) to which the arguments made above might not apply. It thus could be more insightful to conduct the analysis at a lower geographic level than the police precinct, and thereby come closer to the concept of a neighborhood (in urban areas) or village (in rural areas), yet, the publicly available SAPS crime statistics do not allow for this.

Table 13: Regressions of local crime rates on inequality within the catchment area

Independent variables	Property crimes			Violent crimes		
	Residential burglary	Vehicle theft	Aggravated robbery	Aggravated assault	Sexual offences	Murder
	(1)	(2)	(3)	(4)	(5)	(6)
Income inequality						
Gini index	0.443 (0.7965)	0.371 (1.5333)	0.338 (1.3699)	0.289 (0.7644)	-0.444 (0.8126)	-0.291 (1.3200)
Education inequality						
Jenkins index	0.829 (1.2634)	1.007 (2.4453)	2.040 (2.1244)	3.881*** (1.2884)	8.240*** (1.7207)	4.309** (2.1322)
Housing inequality						
Jenkins index	0.034 (0.3987)	0.153 (0.7866)	-0.027 (0.7468)	-0.671* (0.4070)	0.092 (0.4352)	0.147 (0.6552)
Cultural heterogeneity						
Generalized variance (norm.)	0.161 (0.2050)	-0.366 (0.4271)	0.306 (0.4095)	0.330 (0.2112)	0.415* (0.2440)	0.427 (0.3827)
Province-specific time-effects	yes	yes	yes	yes	yes	yes
Spatial lag	yes	yes	yes	yes	yes	yes
Cluster-fixed effects	yes	yes	yes	yes	yes	yes
Full set of controls	yes	yes	yes	yes	yes	yes
Observations	2,170	2,170	2,170	2,170	2,170	2,170

Notes: Each column contains four regressions; one for each inequality dimension. The full set of controls includes ln mean p.c. household income, ln population density, unemployment rate, share of adolescents, share of recently moved, share of female-headed hh, share of urban hh, share of coloureds. Dependent variables are IHS-transformed crime rates (crimes per 100,000 inhabitants). The spatial lag of the crime rate is instrumented with the spatial lags of first, second and third order of all independent variables. Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.
Source: Own calculations using data from ISS, SAPS, StatsSA, and WorldPop.

6.6 Heterogeneity analysis

South African police precincts display substantial heterogeneities in many aspects. While the median precinct covers an area of around 560 km², the smallest precincts cover less than 5 km² and the largest ones more than 15,000 km². Precincts also differ significantly in terms of urbanization, with 49% being highly urbanized (>75% urban households), 20% semi-urban (25%-75% urban households) and 31% rural (<25% urban households). Further potential for heterogeneity analysis lies in the complex colonial history of South Africa, which has been marked by both the British Empire and the Dutch Colonial Empire. In the decades before the formation of the Union of South Africa in 1910 (the predecessor state of the Republic of South Africa), parts of the country were under British colonial and parts under Dutch colonial rule (Britannica, 2007; 2012; 2013; 2020).[29] Lastly, within South Africa's administrative structure, there are eight municipalities of special status, so-called 'metropolitan municipalities', which are strongly urbanized conurbations of high economic importance (Mokoena, 2020).[30] In order to explore potential heterogeneities along the lines of the above-mentioned aspects, I re-ran the main analysis using the following subsamples: SMALL – precincts with an area not larger than the median, URBAN – precincts with an urbanization rate higher than the median, METRO – precincts in metropolitan municipalities, BRITISH – precincts in former British colonies.

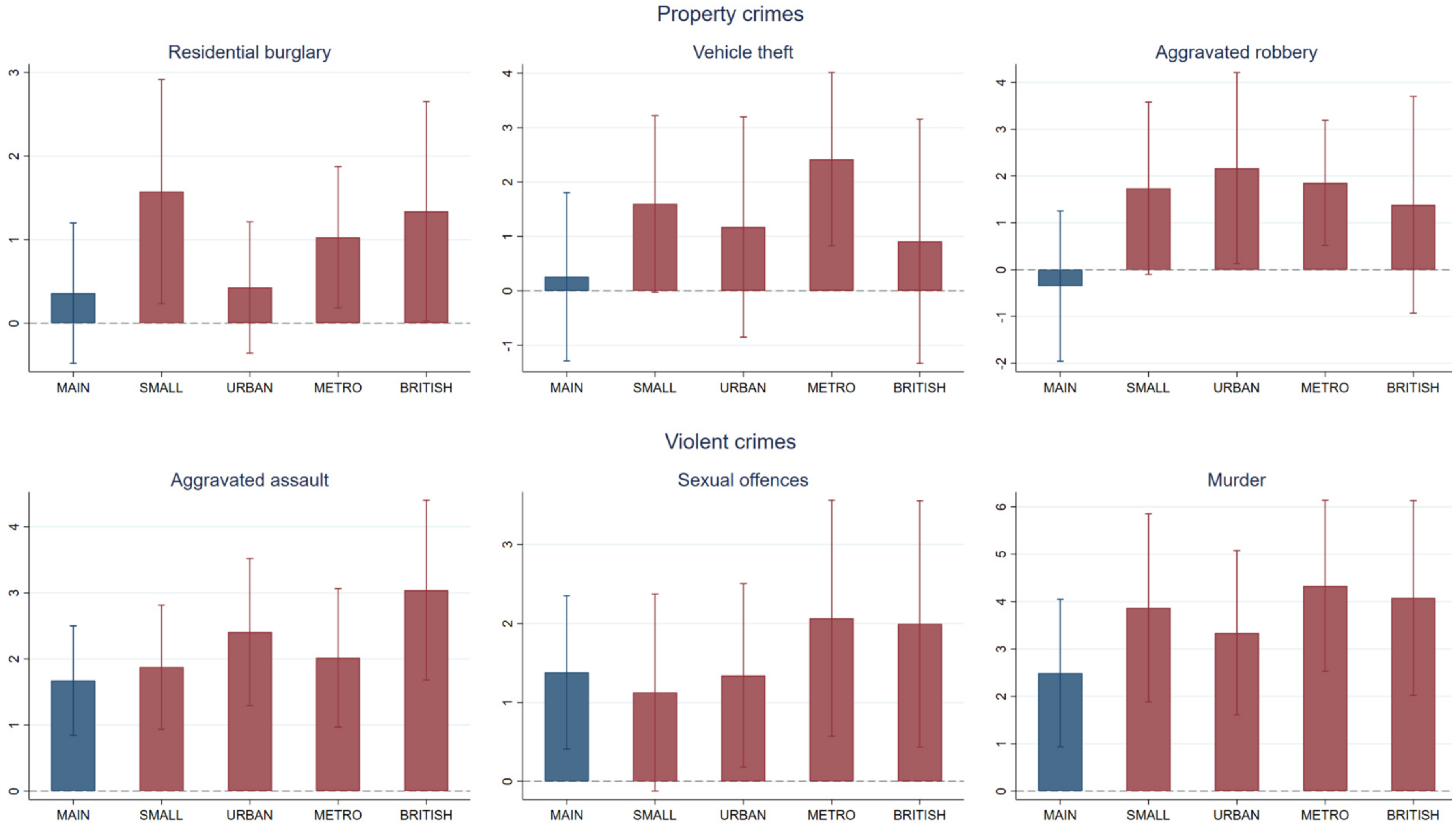
The coefficient plots in Figure 4 present the results from this analysis for each crime type and compare the effects obtained from the various sub-samples with the effects from the main sample. I focus on the results for income inequality, as this inequality dimension reveals the clearest pattern of heterogeneity. Several deductions can be made from the results. First, across all crimes, the effects of income inequality on crime rates are larger in smaller precincts, in more urbanized areas, in metropolitan municipalities and in former British colonies as compared to the main sample.[31] Second, while the main analysis revealed no statistically significant effect of income inequality on property crimes, coefficients are positive, statistically significant and sizable in smaller precincts and metropolitan municipalities for all property crimes. I also find significant effects on rates of residential burglary in former British colonies

[29] British: North West (western part), Northern Cape, Western Cape, Eastern Cape, KwaZulu-Natal; Dutch: North West (eastern part), Limpopo, Mpumalanga, Free State, Gauteng. See Figure A6 in Appendix F for a graphical illustration.

[30] Buffalo City (East London), City of Cape Town, Ekurhuleni Metropolitan Municipality (East Rand), City of eThekwini (Durban), City of Johannesburg, Mangaung Municipality (Bloemfontein), Nelson Mandela Metropolitan Municipality (Port Elizabeth), City of Tshwane (Pretoria). See Figure A7 in Appendix F for a graphical illustration.

[31] Only exception: The effect of income inequality on the rate of sexual offences is smaller in small precincts.

Figure 4: Heterogenous effects of income inequality on crimes rates by precincts' size, urbanization, metropolitan status, and former colonial power



Notes: Bars indicates point estimates and capped lines indicate 95% confidence intervals of the effects of income inequality (Gini index) on IHS-transformed crime rates. n(MAIN)=2,170; n(SMALL)=1,086; n(URBAN)=1,085; n(METRO)=482; n(BRITISH)=1,224.
Source: Own calculations using data from ISS, SAPS, StatsSA, and WorldPop.

and on rates of aggravated robbery in more urbanized areas. Third, turning to violent crimes, the effects are especially large in former British colonies, metropolitan municipalities and (to a lesser extent) in more urbanized precincts. To illustrate, the effect of income inequality on the murder rate is 74% larger in metropolitan municipalities and the effect on the rate of aggravated assault 82% larger in former British colonies. Overall, across all six analyzed crimes, precincts in metropolitan municipalities seem to be the biggest drivers of the positive correlation between income inequality and crime rates. In this subsample, the coefficients are large and significant across all crimes and for vehicle theft, sexual offences and murder, they are larger than in any other sub-sample. This is an important finding, as these eight metropolitan municipalities cover only 2% of the country's surface area, but a large share of its population. They include almost a quarter of all police precincts and more than a third of South Africa's inhabitants. Furthermore, precincts in metropolitan municipalities display markedly higher crime rates, especially in terms of property crimes, as compared to precincts outside of metropolitan municipalities. Looking at absolute crime counts in 2001 and 2011, around half of all reported crimes were committed in metropolitan municipalities; 48% of residential burglaries, 72% of vehicle thefts, 65% of aggravated robberies, 37% of aggravated assaults, 40% of sexual offences, and 45% of murders. Considering that metropolitan municipalities cover only 2% of the country's surface area, but more than a third of the population, are responsible for half of all reported crimes and display an extraordinarily strong correlation between income inequality and crime rates, these areas provide significant scope for crime prevention via inequality reduction.

6.7 Robustness checks

A test of unobservable selection and coefficient stability

Even though I control for a large number of precinct-level confounders, in addition to accounting for spatial correlations, cluster-fixed effects and province-specific time-effects, one might still be worried about remaining omitted variable bias. In order to get some sense of the magnitude of potential omitted variable bias, I follow Altonji, Elder and Taber (2005) and Oster (2019) and calculate bias-adjusted coefficients for the inequality indices using the following formula:

$$\rho^* \approx \tilde{\rho} - \delta \frac{R_{max} - \tilde{R}}{\tilde{R} - \dot{R}} (\dot{\rho} - \tilde{\rho}), \tag{7}$$

where ρ^* is the bias-adjusted coefficient, $\dot{\rho}$ and $\dot{R}$ ($\tilde{\rho}$ and $\tilde{R}$) are the coefficient and the R-squared from the regressions without any controls, except the spatial lag, cluster-fixed effects and province-specific time-effects (with the full list of controls), and δ measures the importance of unobserved variables relative to observed variables.[32] Analog to Kung and Zhou (2021), I follow Oster (2019) and set $\delta = 1$ and $R_{max} = 1.3\tilde{R}$. Alternatively, following Nunn and Wantchekon (2011), I set $R_{max} = \tilde{R} + (\tilde{R} - \dot{R})$. In short, the assumption is that the closer ρ^* to $\tilde{\rho}$, the smaller the extent of unobserved variable bias. According to Oster (2019), a coefficient can be considered as robust if the bias-adjusted coefficient has the same sign as the unadjusted coefficient. In addition, I calculate the value for δ that would be necessary to 'explain away' the effect of instability on my outcomes of interest, i.e. an estimate of how much more important the unobservables need to be than the observables (Calvo et al., 2020). Table A11 in Appendix G presents values for ρ^* and δ for each regression from my main specification, for both assumptions on R_{max}.

Overall, the results from this robustness test are rather reassuring. The effects of income inequality on violent crimes were positive and statistically significant in the main specification and they survive this test for both assumptions on R_{max}. Bias-adjusted coefficients are of the same sign as the unadjusted coefficients and δ is either larger than 1 (implying that the unobservables would need to be more important than the observed control variables in order to nullify the estimated effects) or negative (implying that controlling for unobservables further increases the coefficient size). For education inequality, five out of six coefficients survive the test under $R_{max} = \tilde{R} + (\tilde{R} - \dot{R})$ and two out of six under $R_{max} = 1.3\tilde{R}$, while for housing inequality, the coefficients are generally not robust to this test (yet, δ is not too far away from one). Lastly, for cultural heterogeneity, four out of six coefficients pass the test, independent of the assumption on R_{max}. I can thus conclude that, while omitted variable bias might distort the coefficient of interest to some extent, it does not invalidate the main findings.

Randomization inference

As a further robustness check, I applied randomization inference, i.e. I randomly re-arranged 500 times the inequality indices across years and police precincts and re-run the regressions for

[32] The spatial lag of the crime rate was not instrumented as the Stata command used for the calculation of β and δ, psacalc, does not support instrumental variable regressions. However, this should not influence the results in a meaningful way, as the results from the main regressions barely differ between instrumenting or not instrumenting the spatial lag of the crime rate.

each generated sample. The coefficients obtained with the permuted samples should approximately follow a normal distribution with a high density around zero as the permuted values of the inequality indices should be orthogonal to the actual inequality indices and hence have no explanatory power. Hence, the 'true' regression coefficient should be significantly larger in magnitude and be significantly different from the distribution of coefficients obtained with the randomized samples. If, however, this is not the case, it could be a hint that the estimated effect is in fact driven by unobserved third variables and hence falsely attributed to the true inequality indices. For this, I first permuted all actual inequality indices across all police precincts and both years 500 times, leading to 500 different samples. Then, I re-ran the main regressions using each of the 500 permuted samples. The kernel density plots in Figure A8 in Appendix G plot the coefficients obtained with the permuted samples and compare them to the coefficient from the true sample. In all cases where the main regressions in Section 6.1 yielded statistically significant coefficients for the different inequality indices, the 'true' coefficient is located to the extreme ends of the distribution. In addition, I calculated a one-sided p-value that equals the share of permuted coefficients that exceed the magnitude of the true coefficient. In all above-mentioned cases, this p-value is well below 0.01, often exactly 0. Overall, the permutation test strongly supports the main results and suggests that the estimated effects are not driven by third variables. The results are very similar when, alternatively, the inequality indices are permuted across police precincts within years (results not shown).

Robustness to reduction and expansion of the set of controls

Even though the set of control variables was carefully selected and guided by the existing literature on crime and inequality, it would be informative to explore to what extent the results change in response to excluding some existing or introducing some additional controls.

As the first approach, I exclude the following three control variables as these might be particularly prone to endogeneity issues: unemployment rate (returns from work), share of female-headed households (family instability), and share of coloureds (gang prevalence). As explained above, all three factors should theoretically increase the crime rate, yet, it is easily conceivable that the relationships (also) go in the other direction or that third factors drive the relationship. To illustrate, higher local crime rates could increase the unemployment rate (e.g. through an exodus of firms) and cause family instability (e.g. through incarceration of predominantly male delinquents), while adverse local conditions (e.g. weak police presence) could simultaneously increase gang prevalence and criminal activity in general. I re-ran the main regressions without these three control variables and the coefficients of the various inequality indices barely change (see Table A12 in Appendix G).

As the second approach, I add the following five control variables: a dummy that indicates whether the precinct lies at the coast, a dummy that indicates whether the precinct borders a neighboring country, the distance to the nearest metropolitan municipality (log), the size of the precinct (log), and a dummy that equals one if the precinct was already split and thus actually contains two precincts. The maps presented in Appendix D are indicative of generally higher rates of (especially property) crimes in coastal areas and rather less crimes in border areas, while the heterogeneity analysis revealed higher crime rates in metropolitan municipalities. The size of the precinct could be negatively correlated with police presence and thus facilitate crimes. Lastly, as illustrated in Section 4.1, several new precincts have been formed since 2000/2001 and thus some precincts covered in the dataset actually contain two precincts (two in 2001/2002 and 34 in 2011/2012). These additional policing capacities could help to prevent crimes. When re-running the main regressions with these additional five control variables, the coefficients generally shrink a bit in size and for education inequality they lose their significance with respect to vehicle theft and aggravated robbery (yet, the sign doesn't change). However, overall, they do not deviate considerably from the main results (see Table A13 in Appendix G).

Further robustness checks

I also ran some further, rather simple, robustness check regarding the specification of crime rates and inequality indices.

First, I re-ran the main regressions using log-transformed instead of IHS-transformed crime rates and, while coefficients slightly change in magnitude, the results are very robust to this exercise (see Table A14 in Appendix G).

Second, I used alternative inequality indices, i.e. mean log deviation and the Theil index for income inequality, and Cowell-Flachaire-indices with $\alpha=0$, $\alpha=0.25$, $\alpha=0.5$, $\alpha=0.75$ and $\alpha=0.9$ for education and housing inequality. Table A15 in Appendix G shows that this yields very similar results (though with changing coefficient sizes) as those presented in Section 6.1.

Lastly, I applied alternative grouping and/or ordering to the underlying categories of education and housing inequality. For education inequality, instead of the 17 categories presented in Section 4.3, I group them into eight (No schooling, Grade 1 – 3, Grade 4 – 6, Grade 7 – 9, Grade

10 – 11, Matric, Certificate/Diploma, Degree) and five categories (No schooling, Grade 1 – 6, Grade 7 –11, Matric, Certificate/Diploma/Degree), respectively. For housing inequality, I again use seven categories, but rank traditional dwelling as the 'worst' and informal dwelling as the 'second worst' category and, alternatively, I group them into five categories (Informal / traditional dwelling, Dwelling in backyard / flatlet, Flat in block of flats, Semi-detached / cluster / town house, Freestanding house). As Table A16 in Appendix G shows, the results do not change meaningfully when using these alternative groupings/orderings of the underlying categories.

7 Conclusion

This article contributes to the relatively scarce evidence on the relationship between crime and inequality in countries of the Global South. Specifically, I addressed the question whether South Africa's soaring rates of property and violent crime can be explained by its high levels of socio-economic inequalities. For this purpose, I built a novel panel dataset of South African police precincts by matching official crime records to socio-economic data from two population censuses and compatible household surveys. To the best of my knowledge, this is the first study that uses a panel of police precincts to examine the crime-inequality-nexus in South Africa.

Overall, the results from this study support the notion that socio-economic disparities contribute to higher crime rates. Regression analyses revealed a positive, statistically significant and linear relationship between income inequality within police precincts and local rates of violent crime. Specifically, an increase in the Gini index of income inequality by 0.1 is associated with an increase in the crime rate by 17% for aggravated assault, 14% for sexual violence and 25% for murder, conditional on otherwise equal socio-economic and demographic precinct characteristics. The results further indicate an inverted u-shaped association between local income inequality and property crime rates, suggesting that rising income inequality within precincts also increases property crime rates, but that this effect is counteracted by protective measures employed in high-inequality precincts. These protective investments could potentially hamper investments in other, more productive, assets. In addition, I find evidence that other types of socio-economic inequalities, such as education and housing inequality, as well as cultural heterogeneity, are positively correlated with crime rates.

While I don't claim causality for these results, given the observational nature of the data, I carefully address potential issues of endogeneity between inequality and crime by accounting for province-specific time trends, police cluster-fixed effects, and a large range of controls as suggested by theoretical and empirical work from economists and sociologists. In addition, I use alternative specifications of inequality measures and crime rates, vary the range of controls, employ randomization inference and test for unobservable selection and coefficient instability. The results are largely robust to the various tests, yet, the evidence is strongest for the positive association between violent crimes and income inequality as well as cultural heterogeneity. This finding underlines especially one important mechanism of strain theory: relative deprivation leads to frustration and ultimately culminates in violent behavior.

In light of South Africa's Apartheid past, enduring racial disparities and associated racial tensions, another finding is particularly interesting. Suggestive evidence indicates that inter-racial inequality, i.e. income inequality between the country's four official population groups, contributes more to property crimes, while intra-racial inequality, i.e. inequality within each of these groups, contributes more to violent crime. A reduction of both property and violent crime could thus be fostered by increasing the incomes at the bottom end of the distribution, especially among the largest and poorest population group, black South Africans. This would reduce intra-racial inequality directly and, indirectly, also inter-racial inequality, through a convergence of the mean incomes of the different population groups.

A further important finding, considering South Africa's distinct geographic segregation, is that not only inequality within police precincts, but also inequality between precincts contributes to local crime rates. In particular, I found the ratio of a precinct's mean income to that of its adjacent precincts to be positively correlated with rates of aggravated robbery and, to a lesser extent, vehicle theft. This indicates that inhabitants of poorer areas with low returns from legal work engage in property crime in nearby areas that are relatively rich and thus offer higher returns from these crimes.

In addition, it is worth mentioning that the associations between income inequality and the various types of crime are particularly strong across precincts in South Africa's eight metropolitan municipalities. In only 2% of the country's surface area, these municipalities inhabit more than a third of the population and account for around half of all reported crimes. Efforts directed at preventing crimes through inequality reductions would presumably be most effective and efficient in these areas.

An impediment for the analysis of local crime rates in South Africa is that these are not available at levels lower than the police precinct. Many of these precincts cover rather large areas, making especially the analysis of the role of inequality between local communities difficult. The variance with respect to crime rates and income differences is likely larger when comparing local communities of smaller size, like neighborhoods. Future research could thus greatly benefit from crime data published at lower geographic levels, or ideally, precisely geo-coded data. This would allow for an investigation at more meaningful geographic aggregations, such as suburbs in urban areas and villages in rural areas, and even at varying levels of aggregation.

References

Agnew, R. (1992), Foundation for a general strain theory of crime and delinquency, *Criminology*, 30(1): 47-87.

Altonji, J. G., T. E. Elder and C. R. Taber (2005), Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools, *Journal of Political Economy*, 113(1): 151-184.

Atems, B. (2020), Identifying the dynamic effects of income inequality on crime, *Oxford Bulletin of Economics and Statistics*, 82(4): 751-782.

Bailey, W. C. (1984), Poverty, inequality, and city homicide rates: Some not so unexpected findings, *Criminology*, 22(4): 531-550.

Bangstad, S. (2005), Hydra's heads: PAGAD and responses to the PAGAD phenomenon in a Cape Muslim community, *Journal of Southern African Studies*, 31(1): 187-208.

Becker, G. S. (1968), Crime and punishment: An economic approach, *Journal of Political Economy*, 76(2): 169-217.

Blau, J. R. and P. M. Blau (1982), The cost of inequality: Metropolitan structure and violent crime, *American Sociological Review*, 47(1): 114-129.

Block, M. K. and J. M. Heineke (1982), A labor theoretic analysis of the criminal choice, *American Economic Review*, 65(3): 314-325.

Bourguignon, F. (2001), Chapter 10: Crime as a social cost of poverty and inequality: A review focusing on developing countries, in S. Yusuf, S. Evenett and W. Wu (eds), *Facets of Globalization: International and Local Dimensions of Development, World Bank Discussion Paper No. 415*: 171-191, World Bank, Washington, DC.

Britannica, The Editors of Encyclopaedia (2007), *Natal*, available at https://www.britannica.com/place/Natal-historical-province-South-Africa.

Britannica, The Editors of Encyclopaedia (2012), *Orange Free State*, available at https://www.britannica.com/place/Orange-Free-State.

Britannica, The Editors of Encyclopaedia (2013), *Cape Province*, available at https://www.britannica.com/place/Cape-Province.

Britannica, The Editors of Encyclopaedia (2020), *Transvaal*, available at https://www.britannica.com/place/Transvaal.

Brush, J. (2007), Does income inequality lead to more crime? A comparison of cross-sectional and time-series analyses of United States counties, *Economics Letters*, 96(2): 264-268.

Budescu, D. V. and M. Budescu (2012), How to measure diversity when you must, *Psychological Methods*, 17(2): 215-227.

Buonannoa, P. and J. F. Vargas (2017), Inequality, crime, and the long run legacy of slavery, *Journal of Economic Behavior & Organization*, 159: 539-552.

Burbidge, J. B., L. Magee and A. L. Robb (1988), Alternative transformations to handle extreme values of the dependent variable, *Journal of the American Statistical Association*, 83(401): 123-127.

Bursik, R. J. (1988), Social disorganization and theories of crime and delinquency: Problems and prospects, *Criminology*, 26(4): 519-551.

Bursik, R. J. and H. G. Grasmick (1999), *Neighborhoods and crime: The dimensions of effective community control*, Lexington Books, Maryland.

Calvo, T., E. Lavallée, M. Razafindrakoto and F. Roubaud (2020), Fear not for man? Armed conflict and social capital in Mali, *Journal of Comparative Economics*, 48(2): 251-276.

Centre for the Study of Violence and Reconciliation (2009), *Why does South Africa have such high rates of violent crime? Supplement to the final report of the study on the violent nature of crime in South Africa*, available at https://www.csvr.org.za/why-does-south-africa-have-such-high-rates-of-violent-crime/.

Chintrakarn, P. and D. Herzer (2012), More inequality, more crime? A panel cointegration analysis for the United States, *Economics Letters*, 116(3): 389-391.

Chiu, W. H. and P. Madden (1998), Burglary and income inequality, *Journal of Public Economics*, 69(1): 123-141.

Choe, J. (2008), Income inequality and crime in the United States, *Economics Letters*, 101(1): 31-33.

Cloward, R. A. and L. E. Ohlin (1960), *Delinquency and opportunity: A theory of delinquent gangs*, Free Press, Glencoe.

Cohen, A. K. (1955), *Delinquent boys: The culture of the gang*, Free Press, Glencoe.

Costantini, M., I. Meco and A. Paradiso (2018), Do inequality, unemployment and deterrence affect crime over the long run?, *Regional Studies*, 52(4): 558-571.

Cowell, F. A. and E. Flachaire (2017), Inequality with ordinal data, *Economica*, 84(334): 290-321.

Cronje, M. and D. Budlender (2004), Comparing Census 1996 with Census 2001: An operational perspective, *Southern African Journal of Demography*, 9(1): 67-89.

Demombynes, G. and B. Özler (2005), Crime and local inequality in South Africa, *Journal of Development Economics*, 76(2): 265-292.

Dodson, B. (2002), Gender and the brain drain from South Africa, *SAMP Migration Policy Series No. 23*, Southern African Migration Programme, Waterloo.

Ehrlich, I. (1973), Participation in illegitimate activities: A theoretical and empirical investigation, *Journal of Political Economy*, 81(3): 521-565.

Elbers, C., J. O. Lanjouw and P. Lanjouw (2003), Micro-level estimation of poverty and inequality, *Econometrica,*, 71(1): 355-364.

Enamorado, T., L. F. López-Calva, C. Rodríguez-Castelán and H. Winkler (2016), Income inequality and violent crime: Evidence from Mexico's drug war, *Journal of Development Economics*, 120: 128-143.

Fajnzylber, P., D. Lederman and N. Loayza (2002a), Inequality and violent crime, *The Journal of Law & Economics*, 45(1): 1-39.

Fajnzylber, P., D. Lederman and N. Loayza (2002b), What causes violent crime?, *European Economic Review*, 46(7): 1323-1357.

Finn, A. and M. V. Leibbrandt (2013), Mobility and Inequality in the First Three Waves of NIDS, *SALDRU Working Paper No. 120/ NIDS Discussion Paper 2013/2*, Southern Africa Labour and Development Research Unit, University of Cape Town, Cape Town.

Freeman, R. (1999), Chapter 52: The economics of crime, in O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*: 3529-3571, Elsevier, Amsterdam.

Goode, W. J. (1960), A theory of role strain, *American Sociological Review*, 25(4): 483-496.

Hino, H., M. V. Leibbrandt, R. Machema, M. Shifa, and C. Soudien (2018), Identity, inequality and social contestation in the post-Apartheid South Africa, *SALDRU Working Paper No. 233*, Southern Africa Labour and Development Research Unit, University of Cape Town, Cape Town.

Hipp, J. R. (2007), Income inequality, race, and place: Does the distribution of race and class within neighborhoods affect crime rates?, *Criminology*, 45(3): 665-697.

Hipp, J. R. and K. Kane (2017), Cities and the larger context: What explains changing levels of crime?, *Journal of Criminal Justice*, 49: 32-44.

Hsieh, C.-C. and M. D. Pugh (1993), Poverty, income inequality, and violent crime: A meta-analysis of recent aggregate data studies, *Criminal Justice Review*, 18(2): 182-202.

Hundenborn, J., M. V. Leibbrandt, and I. Woolard (2018), Drivers of inequality in South Africa, *WIDER Working Paper, No. 2018/162*, United Nations University World Institute for Development Economics Research, Helsinki.

İmrohoroğlu, A., A. Merlo and P. Rupert (2000), On the political economy of income redistribution and crime, *International Economic Review*, 41(1): 1-25.

Jenkins, S. P. (2020), Inequality comparisons with ordinal data, *The Review of Income and Wealth*, 67(3): 547-563.

Kang, S. (2016), Inequality and crime revisited: Effects of local inequality and economic segregation on crime, *Journal of Population Economics*, 29: 593-626.

Kelly, M. (2000), Inequality and crime, *The Review of Economics and Statistics*, 82(4): 530-539.

Kempe, G. (2016), *Re-mapping crime statistics*, available at https://github.com/OpenUpSA/crime-stats-demystifed.

Kennedy, B. P., I. Kawachi, D. Prothrow-Stith, K. Lochner and V. Gupta (1998), Social capital, income inequality, and firearm violent crime, *Social Science and Medicine*, 47(1): 7-17.

Kim, B., C. Seo and Y.-O. Hong (2020), A systematic review and meta-analysis of income inequality and crime in Europe: Do places matter?, *European Journal on Criminal Policy and Research*.

Kondo, K. (2021), SPGEN: Stata module to generate spatially lagged variables, available at https://EconPapers.repec.org/RePEc:boc:bocode:s458105.

Kornhauser, R. R. (1978), *Social sources of delinquency: An appraisal of analytic models*, University of Chicago Press, Chicago.

Kubrin, C. E. and R. Weitzer (2003), New directions in social disorganization theory, *Journal of Research in Crime and Delinquency*, 40(4): 374-402.

Kung, J. K.-s. and T. Zhou (2021), Political elites and hometown favoritism in famine-stricken China, *Journal of Comparative Economics,* 49(1): 22-37.

Leibbrandt, M. V., A. Finn and I. Woolard (2012), Describing and decomposing post-apartheid income inequality in South Africa, *Development Southern Africa*, 29(1): 19-34.

Leibbrandt, M. V., L. Poswell, P. Naidoo, M. Welch and I. Woolard (2005), Measuring recent changes in South African inequality and poverty using 1996 and 2001 census data, *DPRU Working Paper 05/94*, Development Policy Research Unit, University of Cape Town, Cape Town.

Leibbrandt, M. V., E. Wegner and A. Finn (2011), The policies for reducing income inequality and poverty in South Africa, *SALDRU Working Paper No. 64*, Southern Africa Labour and Development Research Unit, University of Cape Town, Cape Town.

Leibbrandt, M. V., I. Woolard and C. Woolard (2009), Chapter 10: A long-run perspective on contemporary poverty and inequality dynamics, in J. Aron, B. Kahn, and G. Kingdon (eds), *South African Economic Policy under Democracy*: 270-299, Oxford University Press, Oxford.

Leibbrandt, M. V., I. Woolard, A. Finn and J. Argent (2010), Trends in South African income distribution and poverty since the fall of Apartheid, *OECD Social, Employment and Migration Working Papers No. 101*, OECD Publishing, Paris.

Machin, S. and C. Meghir (2004), Crime and Economic Incentives, *The Journal of Human Resources*, 39(4): 958-979.

Manea, R., P. Piraino and M. Viarengo (2021), Crime, inequality and subsidized housing: evidence from South Africa, *Research Paper 66*, The Graduate Institute Geneva, Centre for International Environmental Studies.

Marchetti-Mercer, M. C. (2012), Is it just about the crime? A psychological perspective on South African emigration, *South African Journal of Psychology*, 42(2): 243-254.

Menezes, T., R. Silveira-Neto, C. Monteiro and J. L. Ratton (2013), Spatial correlation between homicide rates and inequality: Evidence from urban neighborhoods, *Economics Letters*, 120(1): 97-99.

Merton, R. K. (1938), Social structure and anomie, *American Sociological Review*, 3(5): 672-682.

Messner, S. F. and R. Rosenfeld (1994), *Crime and the American dream*, Wadsworth, Belmont.

Metz, N. and M. Burdina (2018), Neighbourhood income inequality and property crime, *Urban Studies*, 55(1): 133-150.

Mokoena, M. (ed) (2020), *South Africa Yearbook 2019/20*, Government Communications (GCIS), Pretoria, available at https://www.gcis.gov.za/content/resourcecentre/sa-info/south-africa-yearbook-201920.

Neumayer, E. (2003), Good policy can lower violent crime: Evidence from a cross-national panel of homicide rates, 1980–97, *Journal of Peace Research*, 40(6): 619-640.

Neumayer, E. (2005), Inequality and violent crime: Evidence from data on robbery and violent theft, 1980–97, *Journal of Peace Research*, 42(1): 101-112.

Nivette, A. E. (2011), Cross-national predictors of crime: A meta-analysis, *Homicide Studies*, 15(2): 103-131.

Nunn, N. and L. Wantchekon (2011), The slave trade and the origins of mistrust in Africa, *American Economic Review*, 101(7): 3221-3252.

Oster, E. (2019), Unobservable selection and coefficient stability: Theory and evidence, *Journal of Business & Economic Statistics*, 37(2): 187-204.

Pare, P.-P. and R. Felson (2014), Income inequality, poverty and crime across nations, *The British Journal of Sociology*, 65(3): 434-458.

Patterson, E. B. (1991), Poverty, income inequality, and community crime rates, *Criminology*, 29(4): 755-776.

Petrus, T. S. (2015), Enemies of the “state”: Vigilantism and the street gang as symbols of resistance in South Africa, *Aggression and Violent Behavior*, 22: 26-32.

Petrus, T. S. and I. Kinnes (2019), New social bandits? A comparative analysis of gangsterism in the Western and Eastern Cape provinces of South Africa, *Criminology & Criminal Justice*, 19(2): 179-196.

Poveda, A. C. (2015), Economic development, inequality and poverty: An analysis of urban violence in Colombia, *Oxford Development Studies*, 39(4): 453-468.

Pridemore, W. A. (2011), Poverty matters: A reassessment of the inequality–homicide relationship in cross-national studies, *The British Journal of Criminology*, 51(5): 739-772.

Pyatt, G. (1976), On the interpretation and disaggregation of Gini coefficients, *The Economic Journal*, 86(342): 243-255.

Runciman, W. G. (1966), *Relative deprivation and social justice: A study of attitudes to social inequality in twentieth-century England*, University of California Press, Berkeley.

Sampson, R. J. and W. B. Groves (1989), Community structure and crime: Testing social-disorganization theory, *American Journal of Sociology*, 94(4): 774-802.

Schönteich, M. and A. Louw (2001), Crime in South Africa: A country and cities profile, *Occasional Paper No 49*, Institute for Security Studies, Pretoria.

Scorzafave, L. G. and M. K. Soares (2009), Income inequality and pecuniary crimes, *Economics Letters*, 104(1): 40-42.

Seekings, J. and N. Natrass (2005), *Class, race, and inequality in South Africa*, Yale University Press, New Haven and London.

Shaw, C. and H. McKay (1942), *Juvenile delinquency and urban areas*, University of Chicago Press, Chicago.

Shorrocks, A. F. (1984), Inequality decomposition by population subgroups, *Econometrica*, 52(6): 1369-1385.

Sjoquist, D. L. (1973), Property crime and economic behavior: Some empirical results, *American Economic Review*, 63(3): 439-446.

Stark, R. (1987), Deviant places: a theory of the ecology of crime, *Criminology*, 25(4): 893-909.

Statistics South Africa (2019a), *Report-03-10-19 - Inequality trends in South Africa: A multidimensional diagnostic of inequality*, available at http://www.statssa.gov.za/?page_id=1854&PPN=Report-03-10-19&SCH=7680.

Statistics South Africa (2019b), *P0341 - Victims of Crime Survey, 2018*, available at http://www.statssa.gov.za/?page_id=1854&PPN=P0341&SCH=7678.

Statistics South Africa (2020), *P0341 - Victims of Crime Survey, 2019*, available at http://www.statssa.gov.za/?page_id=1854&PPN=P0341&SCH=72635.

Stigler, G. J. (1970), The optimum enforcement of laws, *Journal of Political Economy*, 78(3): 526-536.

Stolzenberg, L., D. Eitle and S. J. D'Alessio (2006), Race, economic inequality, and violent crime, *Journal of Criminal Justice*, 34(3): 303-316.

Stucky, T. D., S. B. Payton and J. R. Ottensmann (2015), Intra- and inter-neighborhood income inequality and crime, *Journal of Crime and Justice*, 39(3): 345-362.

Terreblanche, S. J. (2002), *A history of inequality in South Africa, 1652 to 2002*, University of Natal Press, Pietermaritzburg.

United Nations University World Institute for Development Economics Research, *World Income Inequality Database (WIID), Version 31 May 2021*, available at https://doi.org/10.35188/UNU-WIDER/WIID-310521.

Wenger, M. R. (2019), Separating the influence of inequality on burglary and robbery by level of analysis: A multilevel approach, *Social Science Research*, 81: 42-60.

Wickes, R. (2017), Part II, Chapter 6: Social disorganization theory: Its history and relevance to crime prevention, in B. Teasdale and M. Bradley (eds), *Preventing Crime and Violence. Advances in Prevention Science*: 57-66, Springer International Publishing, Cham.

Wilson, F. (2011), Historical roots of inequality in South Africa, *Economic History of Developing Regions*, 26(1): 1-15.

Woolard, I. (2002), An overview of poverty and inequality in South Africa, *Working Paper prepared for DFID (SA), July 2002*.

World Bank Group (2021a), *Gini index (World Bank estimate)*, available at https://data.worldbank.org/indicator/SI.POV.GINI?view=map&year=2014.

World Bank Group (2021b), *Intentional homicides (per 100,000 people)*, available at https://data.worldbank.org/indicator/VC.IHR.PSRC.P5?view=map&year=2014.

Yu, D. (2010), Poverty and inequality estimates of National Income Dynamics Study revisited, *Stellenbosch Economic Working Papers: 05/13*, Stellenbosch University, Stellenbosch.

Zhao, Q. and P. Lanjouw (n.d.), *Using PovMap2: A user's guide*, World Bank, Washington, DC.

Appendix

Appendix A: Additional information on SAPS crime statistics

This section provides further details on the SAPS crime statistics, which were used to generate a panel of annual crime rates at the police precinct level from 2000/2001 to 2014/2015 for 1,089 precincts.

Table A1 on the next page presents the categorization of crimes as provided in the official annual SAPS crime statistics and Table A2 provides an overview of the availability of crime statistics for each crime per SAPS year.

Table A3 and Figure A1 illustrate the creation of new police precincts and the concomitant changes of police precinct boundaries between 2000/2001 and 2014/2015. In total, 51 new police stations were established during this period, affecting the policing areas and precinct boundaries of 45 existing precincts. Table A3 provides for each newly established precinct the existing precinct that previously policed the affected area and thus recorded all crime in this area. It also contains information on the SAPS year in which official crime statistics for the new station were first published and (whenever available) the official opening date of the new precinct. Lastly, it provides the source of information, i.e. whether the information was received from the SAPS Deputy Information Officer, via internet research (archived speeches, press releases, parliamentary monitoring group, twitter) or from the affected stations via e-mail or personal phone call. The information summarized in this table was used together with a shapefile of all police precincts as of September 2014 to create a shapefile of all police precincts as of SAPS financial year 2000/2001. Figure A1 shows the original shapefile from 2014, in which all newly created precincts are colored red and all affected existing stations are colored green. Merging all green precincts with their corresponding red precinct(s) yielded the police precinct boundaries as of 2000/2001.

Table A1: Official crime categorization of annual SAPS crime statistics

Community reported rimes				
Contact crimes	Contact related crimes	Property related crimes	Other serious crimes	Crimes detected as result of police action
Murder	Arson	Burglary at residential premises	Other theft	Illegal possession of firearm and ammunition
Attempted murder	Malicious damage to property	Burglary at nonresidential premises	Commercial crime	Drug related crimes
Assault with the intent to inflict grievous bodily harm		Theft of motor vehicle and motor cycle	Shoplifting	Driving under the influence of alcohol and drugs
Common assault		Theft out of or from motor vehicle		Sexual offences detected as a result of police action
Robbery with aggravating circumstances[1]		Stock theft		
- Carjacking (TRIO)				
- Robbery – residential (TRIO)				
- Robbery – nonresidential (TRIO)				
- Truck hijacking				
- Robbery of cash-in-transit				
- Bank robbery				
Common robbery				
Sexual offences[2]				
- Rape				
- Sexual assault				
- Contact sexual assault				
- Attempted sexual offences				

Notes: [1] Sub-categories of 'robbery with aggravating circumstances' are non-exhaustive. The first three subcategories are together labelled as so-called TRIO-crimes.
[2] Sub-categories of 'sexual offences' are exhaustive. Until 2007, there were only two sub-categories, 'rape' and 'indecent assault'.
Source: Author

Table A2: Availability of SAPS crime statistics, by SAPS year and crime

Crime (category)	2000/01	01/02	02/03	03/04	04/05	05/06	06/07	07/08	08/09	09/10	10/11	11/12	12/13	13/14	14/15
Comm. rep. crimes	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Contact crimes	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Murder	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Attempted murder	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Assault GBH	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Common assault	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Robbery aggravated	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
- Carjacking	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
- Robbery res.				✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
- Robbery nonres.				✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
➢ TRIO				✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
- Truck hijacking	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
- Robbery of CIT	✓	✓				✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
- Bank robbery	✓	✓				✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Common robbery	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sexual offences		✓		✓	✓				✓	✓	✓	✓	✓	✓	✓
- Rape									✓	✓	✓	✓	✓	✓	✓
- Sexual assault									✓	✓	✓	✓	✓	✓	✓
- Cont. sex. assault									✓	✓	✓	✓	✓	✓	✓
- Att. sex. offences									✓	✓	✓	✓	✓	✓	✓
Contact rel. crimes	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Arson	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mal. dam. to prop.	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Property rel. crimes	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Burglary – res.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Burglary – nonres.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Theft of vehicle	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Theft out of vehicle	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Stock theft	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Other ser. crimes	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Other theft	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Commercial crime	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Shoplifting	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Crimes det. police	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Illegal poss. firearm	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Drug related crimes	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Driving alc./drugs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sex. off. det. police											✓	✓	✓	✓

Notes: ✓ indicates availability of crime statistics.
Source: Author.

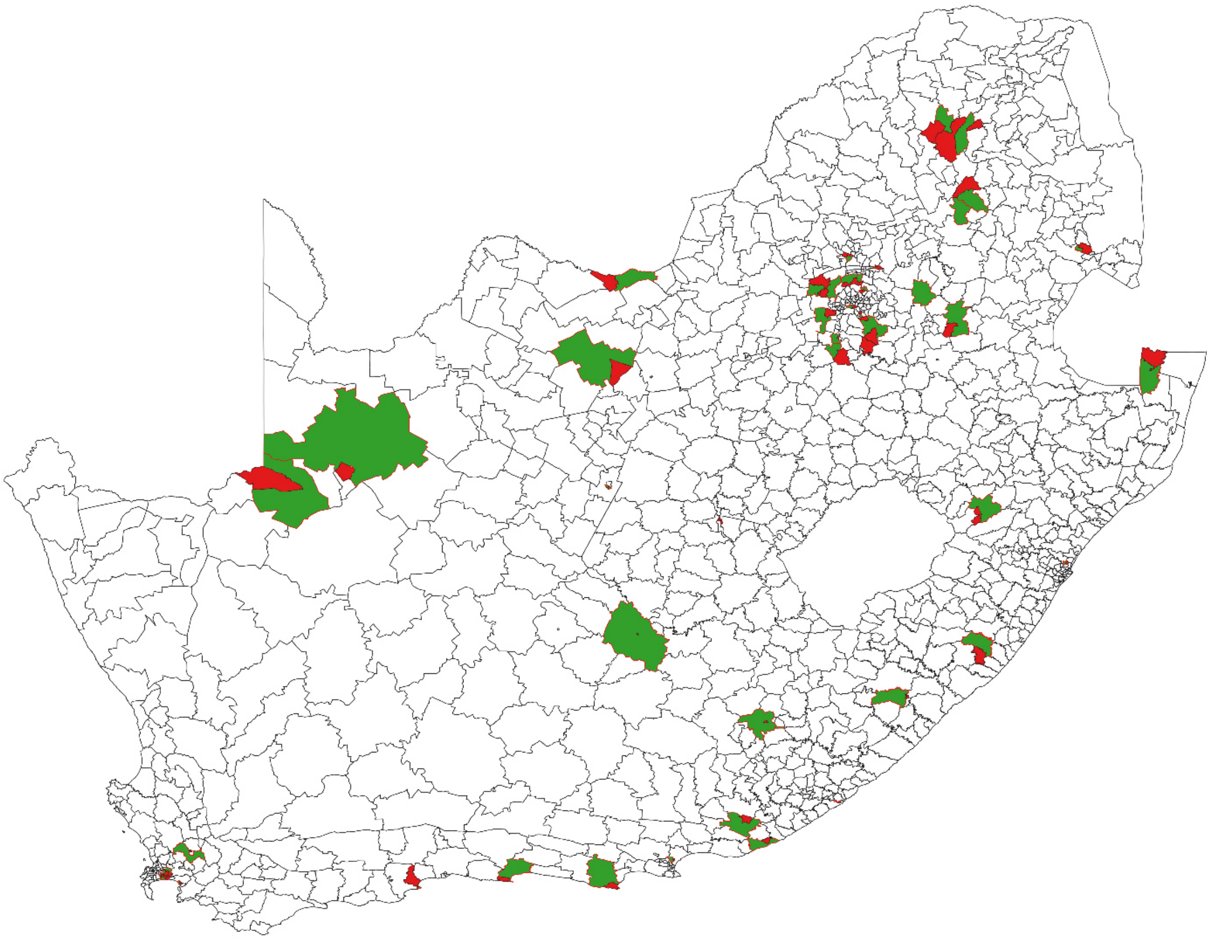
Table A3: Establishment of police precincts between 2000/2001 and 2014/2015

New station	Prior station of crime records	SAPS year of first statistics	Month/year of official opening	Source
Emzinoni	Bethal	2001/2002	10/2001	Phone
Zonkizizwe	Heidelberg (Gp)		09/2000	E-Mail
Boipatong	Vanderbijlpark	2003/2004	08/2003	Archived speech[1]
Rosedale	Upington		04/2004	Phone
Tshidilamolomo	Makgobistad		08/2004	Phone
Wembezi	Estcourt		08/2003	E-Mail
Da Gamaskop	Mossel Bay	2004/2005	?	Phone
Harare	Khayelitsha		-/2004	Phone
Lingelethu-West	Khayelitsha		08/2004	Phone
Kleinvlei	Kuilsriver		07/2004	E-Mail
Kopanong	Batho		08/2004	E-Mail
Kuyasa	Colesberg		04/2004	E-Mail
Mangaung	Batho		06/2004	Archived speech[2]
Mbewkweni	Paarl		?	Phone
Philippi East	Nyanga		08/2004	E-Mail
Rabie Ridge	Ivory Park		?	Phone
Kwanokuthula	Plettenberg Bay	2005/2006	05/2004	Phone
Mfuleni	Kuilsriver		?	Phone
Huhudi	Vryburg	2006/2007	12/2006	Phone
Ikamvelihle	Motherwell		12/2006	Phone
Mamelodi East	Mamelodi		-/2007	Parl. monit. group[3]
Ndumo	Ingwavuma		-/2006	Phone
Belhar	Delft	2007/2008	12/2007	E-Mail
Hekpoort	Magaliesburg		10/2006	Phone
Nemato	Port Alfred		10/2007	Phone
Ratanda	Heidelberg (Gp)		06/2007	Phone
Kagisho	Galeshewe	2008/2009	-/2009	Phone
Lwandle	Strand		?	Phone
Augrabies	Kakamas	2009/2010	04/2009	Phone
Tarlton	Krugersdorp		09/2009	E-Mail
Diepsloot	Erasmia	2010/2011	-/2010	Dep. inform. officer
Kwamashu E	Ntuzuma		-/2010	Dep. inform. officer
St Francis Bay	Humansdorp		-/2011	Dep. inform. officer
Hebron	Ga-Rankuwa	2011/2012	-/2011	Press release[4]
Masemola	Nebo		11/2011	Tweet by station[5]
Olievenhoutbosch	Wierdabrug		-/2011	Dep. inform. officer

Pienaar	Kanyamazane		01/2011	Phone
Zamdela	Sasolburg		-/2011	Dep. inform. officer
Bekkersdal	Westonaria		-/2012	Dep. inform. officer
Katlehong North	Katlehong		-/2012	Dep. inform. officer
Mashashane	Seshego		-/2012	Dep. inform. officer
Mlungisi	Queenstown	2012/2013	-/2012	Dep. inform. officer
Moffatview	Booysens		04/2012	Phone
Vaal Marina	Heidelberg (Gp)		-/2012	Dep. inform. officer
Westenburg	Polokwane		-/2012	Dep. inform. officer
Joza	Grahamstown		-/2014	Dep. inform. officer
Kwandengane	Bizana		-/2013	Dep. inform. officer
Lentegeur	Mitchells Plain	2013/2014	-/2013	Dep. inform. officer
Madeira	Mthatha		-/2013	Dep. inform. officer
Sebayeng	Polokwane		-/2013	Dep. inform. officer
Tembisa South	Tembisa	2014/2015	-/2013	Dep. inform. officer

Notes: [1] www.polity.org.za/article/shilowa-opening-of-police-station-in-boipatong-21082003-2003-08-21.
[2] www.gov.za/n-kganyago-hand-over-meloding-police-station.
[3] www.pmg.org.za/committee-question/4701/.
[4] www.gov.za/minister-police-nathi-mthethwa-officially-open-klipgat-and-hebron-police-stations-and-launch-safer.
[5] www.twitter.com/1060facts/status/578902503602937856?lang=de.
Small discrepancies between month/year of official station opening and SAPS year of first published crime statistics are possible. In this case the latter was used for the formation of the crime statistics panel.
Source: Author

Figure A1: Creation of shapefile of police precinct boundaries as of 2000/2001



Notes: Newly established precincts are colored red and outlined with a dotted black line. Affected existing precincts are colored green and outlined with a solid red line. Since new precincts lie above the affected old precincts, the area of an old precinct includes both the green and the red area.
Source: Own calculations using data from ISS and SAPS.

Appendix B: Additional information on the merging of census and crime data

This section provides additional information on how the census data was allocated to police precincts using ‘SAL-to-precinct-weights’. Figure A2 below illustrates the merging procedure exemplarily for the two police precincts (red dashed lines) ‘Pinelands’ (left) and ‘Elsies River’ (right) in 2001. SAL boundaries are denoted with solid green lines and the pixels in varying red tones indicate population distribution, with darker red corresponding to a higher population. Most of the SALs are contained fully in either of the two precincts, hence, all the census households from these SALs can be allocated precisely to their respective precincts (‘SAL-to-precinct-weight=1). One SAL though (orange-striped, henceforth SAL X) intersects with both precincts. Based on the census data, it is thus impossible to infer whether the households located in SAL X live in ‘Pinesland’ or ‘Elsies River’. Yet, in combination with the population distribution data from WorldPop, it is possible to estimate the probability with which the households in SAL X live in each of the two precincts. For this, I sum up the population data of all pixels in SAL X as well as in the intersecting area of SAL X with ‘Pinelands’ and in the intersecting area of SAL X with ‘Elsies River’. Dividing the estimated population in each intersecting area by the estimated total population of SAL X yields the probability of SAL X inhabitants living in ‘Pinelands’ and ‘Elsies River’, respectively. These probabilities serve as the ‘SAL-to-precinct-weights’. All households from SAL X are then allocated to both precincts, yet, all further calculations of precinct-level variables are executed using these weights.

Figure A2: Creation of ‘SAL-to-precinct-weights’ for merging of census and crime data

Notes: Red dashed lines indicate police precinct boundaries, green solid lines SAL boundaries, pixels in varying red tones population distribution (darker red: higher population), orange stripes intersection of SAL X with both police precincts. Population data was altered in this example for reasons of clarity and comprehensibility.
Source: Own calculations using data from SAPS, StatsSA, and WorldPop.

Appendix C: Small Area Estimation (SAE) of household income

This section provides a detailed description of how the method of Small Area Estimation (SAE) was implemented in order to simulate incomes for the census households. In the following, I will summarize how I proceeded in each of the three main steps of SAE, in particular, how the underlying data was created, how variables were selected and models specified and how simulations were run. For a detailed description of the econometric approach of the method, please refer to Elbers, Lanjouw and Lanjouw (2003). Alternatively, Demombynes and Özler (2005) provide a brief summary of the econometrics. The user manual of the software PovMap 2.0 (Zhao and Lanjouw, n.d.) contains comprehensive guidelines for each step of the practical implementation. Whenever I describe recommendations, rules, guidelines or the like, these are taken from Zhao and Lanjouw and not my original contributions. This section is not to be understood as a comprehensive guideline on how to implement SAE in general, but rather as an effort to be transparent about the implementation in this study.

General set-up

All stages of SAE have to be conducted separately for each year, for each province, and for rural and urban households (i.e. by year and stratum). The combination of province and location status (rural/urban) is referred to as 'domain' by Zhao and Lanjouw (n.d.). Importantly, the size of the sub-samples from the surveys should not fall below 300 households, such that in some cases adjacent provinces had to be merged to form one domain. Accordingly, I created the following 15/17 sub-samples in 2001/2011, as shown in Table A4 on the next page.

1) Selection of candidate variables

In principle, all variables that can plausibly serve as explanatory variables for household income, are available in both the census and the survey, have the same definition and a very similar distribution in both data sources (at the domain level), can serve as candidate variables for the income model. My pool of candidate variables covered four broad classes: demographics, education, occupation (including variables for the household head and for all household members), and housing. Table A5 summarizes all potential candidate variables for 2001 and 2011. Whenever variables had slightly different definitions in the census versus the survey, I harmonized them across datasets (e.g. finer division of highest education level). Speaking more generally, it is crucial that all included candidate variables are coded in the exact

Table A4: Overview of survey and census sub-samples based on domains, by year

2001					2011				
Domain	Included provinces	Urban / rural	n(survey)	n(census)	Domain	Included provinces	Urban / rural	n(survey)	n(census)
1	Western Cape	urban	1,920	892,080	1	Western Cape	urban	2,779	1,180,490
2	Eastern Cape	urban	1,572	552,996	2	Eastern Cape	urban	1,657	718,211
3	Northern Cape	urban	856	142,443	3	Northern Cape	urban	1,089	177,443
4	Free State	urban	1,347	471,805	4	Free State	urban	1,831	582,887
5	KwaZulu-Natal	urban	2,131	832,698	5	KwaZulu-Natal	urban	1,828	1,053,949
6	North West	urban	1,271	326,880	6	North West	urban	1,142	481,697
7	Gauteng	urban	3,546	2,002,697	7	Gauteng	urban	3,824	2,820,224
8	Mpumalanga	urban	1,000	299,583	8	Mpumalanga	urban	1,277	408,709
9	Limpopo	urban	848	124,448	9	Limpopo	urban	779	226,712
10	Western Cape	rural	627	83,133	10	Western Cape & Northern Cape & Free State	rural	595	193,136
11	Eastern Cape	rural	1,883	717,614	11	Eastern Cape	rural	1,654	724,939
12	Northern Cape	rural	451	28,293	12	KwaZulu-Natal	rural	1,791	886,349
13	Free State	rural	672	109,391	13	North West & Gauteng	rural	1,403	540,486
14	KwaZulu-Natal	rural	2,256	718,710	14	Mpumalanga	rural	1,021	419,907
15	North West & Gauteng	rural	1,595	455,558	15	Limpopo	rural	2,502	1,051,139
16	Mpumalanga	rural	1,188	308,178					
17	Limpopo	rural	2,218	852,680					

Notes: Domain 15 in 2001 and domains 10 and 13 in 2011 were created by combining all rural households from adjacent provinces, since the number of rural households sampled in the surveys in these provinces were less than 300.
Source: Author

Table A5: Overview of potential candidate variables for income model, by year

Demographics		Education		Occupation		Dwelling
Head	Members	Head	Members	Head	Members	
			2001			
Sex Age Population group Marital status	Household size Average age Share/number of children Share/number of children Share/number of adults Share/number of elderlies Share/number of males Share/number of male adults	Highest education level Total years of schooling	Years of schooling per capita Total years of schooling Share/number of members with matric	Employment status	Share/number of members with work Share/number of males with work Share/number of members with matric and work	Dwelling type Toilet type Refuse removal type Water access type
			2011			
Sex Age Population group Partner present?	Household size Average age Share/number of children Share/number of children Share/number of adults Share/number of elderlies Share/number of males Share/number of male adults	Highest education level Total years of schooling	Years of schooling per capita Total years of schooling Share/number of members with matric	Employment status	Share/number of members with work Share/number of males with work Share/number of members with matric and work	Dwelling type Toilet type Water access type

Notes: In case of a continuous variable, the square term was added to the Beta model if this increased the fit of the model. All categorical variables, such as 'highest education level' or 'dwelling type' were created in multiple versions (i.e. with varying grouping of categories) to provide flexibility for the specification of the Beta model. Lastly, the interaction of two candidate variables was also allowed to enter the Beta model, if this increased the fit of the Beta model and if both linear terms and the interaction term were statistically significant.
Source: Author

same manner in both datasets. Households in the survey have to be representative of the census households at the domain level for each included candidate variable (household weights are used). For continuous variables, this means that the mean and percentiles should be very similar across datasets. Categorical variables should exhibit a very similar distribution across categories in both datasets. For more flexibility, I applied several alternative groupings to all categorical variables (i.e. if the original variable 'dwelling type' contained eight categories, I included additional variables with six, four, and two categories by grouping adjacent categories). All households with missing data on one of the candidate variables were excluded in both the census and survey datasets. The final products of this stage are household level datasets for each domain in 2001 and 2011, containing all candidate variables that passed the above explained criteria, for the census and for the survey. Importantly, the range of included candidate variables vary across domains as representativeness is not guaranteed for all variables in all domains.

2) Income model

This stage serves to specify the econometric model of household income, i.e. the model that best explains the variation in per capita household income in the survey, for each domain. It can be divided into the following sub-stages: A.1) Selection of household level explanatory variables (Beta model), A.2) Selection of cluster level variables (Beta model), A.3) Correction for outliers (Beta model), B) Modelling of heteroskedasticity (Alpha model).

A.1) Beta model – selection of household level explanatory variables

In each survey sub-sample, I started with simple (weighted) regressions of log per capita household income on each included candidate variable. The candidate variable with the highest R^2 and $p<0.15$ (for categorical variables all categories have to pass $p<0.15$) is selected as the first explanatory variable. From here on, further explanatory variables are added iteratively along the lines of the following criteria: 1) Additional regressors (substantially) increase R^2, 2) additional regressors all have $p<0.15$, 3) if feasible, inclusion of regressors from all five classes.[33] The inclusion of an additional regressor could lead to insignificance of a previously included regressor. In this case, the regressor that contributes more to the R^2 stays in the model and the other one is dropped. Overall, the goal is to maximize the R^2, and it should not be smaller than 0.35. In order to avoid overfitting, the number of included regressors is bounded by the square root of the number of surveyed households. Table A6 presents, for each subsample, the R^2 after this stage.

[33] Explanatory variables could also enter the model together with their squared term (only for continuous variables) or in the form of interaction terms with other variables (possible for continuous and categorical variables).

A.2) Beta model – selection of cluster level variables

For the simulation of the error term, the error is split into a household component and a cluster component. This stage serves to minimize the cluster error by including cluster level variables.

For this, identifying common geographic clusters in the census and the survey is the first step. Often, household surveys and censuses of the same or nearby years have common enumeration areas (EAs) which can readily serve as clusters. In the case of South Africa, this is not possible, as the primary sampling units (PSUs) of the surveys are formed on the basis of the EA of the preceding census (i.e. the PSUs of the IES 2011 are based on the EAs of the Census 2001) and the census does not contain information on the EA, only on the SAL. Fortunately, I received detailed information from StatsSA on how exactly the survey PSUs were formed from the EAs of preceding censuses. Combining this information with shapefiles of census EAs, I created shapefiles of survey PSUs and overlaid them with the SAL shapefiles. This enabled me to identify (approximately) common geographic clusters between surveys and censuses. Generally, a cluster corresponds to one SAL; sometimes two or three SALs were merged to fit one PSU.

The second step comprises creating cluster level variables from the census data, again from all four classes introduced above. For education, this could be for example the share of adults with at least grade 12 (Matric) education in the cluster. In addition, I calculated mean night time light intensity as a proxy for local economic development using data from Li et al. (2021).

The next step serves to select those cluster level variables that most effectively reduce the cluster component of the error term. For this, cluster level regressions are run with the cluster component of the residual on the left-hand side and potential cluster level variables on the right-hand side. Again, selection criteria are maximizing R^2, including only significant regressors and including regressors from several classes. Hereby, the number of cluster level variables should not exceed the square root of the number of clusters in a domain.

After selection of cluster level variables, these were added to the income model. In case one or more previously significant regressors have now turned insignificant, the model needed to be adjusted accordingly. Table A6 presents, for each subsample, the R^2 after adding cluster level variables to the Beta model.

Table A6: Overview of R^2 in Beta model, in different stages, by year

2001				2011			
Domain	R^2 w/o cluster var.	R^2 w/ cluster var.	R^2 after outlier excl.	Domain	R^2 w/o cluster var.	R^2 w/ cluster var.	R^2 after outlier excl.
1	0.3060	0.4132	0.7023	1	0.5111	0.5336	0.6568
2	0.4644	0.4843	0.6667	2	0.4652	0.4934	0.6200
3	0.5138	0.5399	0.6162	3	0.5752	0.5991	0.6403
4	0.4330	0.4410	0.5786	4	0.5381	0.5653	0.6206
5	0.5198	0.5089	0.6400	5	0.4577	0.4647	0.5630
6	0.3881	0.3903	0.5576	6	0.4952	0.5122	0.6083
7	0.4238	0.4430	0.6447	7	0.4203	0.4383	0.5464
8	0.4496	0 5156	0.6645	8	0.4773	0.5020	0.6044
9	0.4528	0.5149	0.5687	9	0.4476	0.5101	0.5429
10	0.3644	0.3732	0.6774	10	0.3144	0.3809	0.5458
11	0.3446	0.3539	0.3934	11	0.3547	0.3932	0.3932
12	0.5424	0.5952	0.7506	12	0.5505	0.5609	0.5725
13	0.4525	0.4655	0.5193	13	0.5641	0.5802	0.5802
14	0.3010	0.3134	0.4102	14	0.3878	0.4394	0.4669
15	0.3896	0.4007	0.4960	15	0.3740	0.3879	0.4096
16	0.4188	0.4334	0.5853				
17	0.4528	0.4318	0.4843				

Source: Author

A.3) Beta model – correction for outliers

After completion of the Beta model, I identified potential outliers, i.e. households with suspiciously large residuals. Large residual might be an indication of income misreporting (mostly underreporting) in the household surveys. In most cases, outliers had very low (sometimes zero) incomes. All identified households underwent a plausibility check, i.e. it was assessed whether their low (or zero) income was plausible given their location, demographic composition, educational and occupational structure, and housing situation. If the low incomes seemed implausible and thus misreporting of income was concluded, the household was excluded. After exclusion, the Beta model (including cluster level variables) was run again and, if necessary, adjusted. Table A6 presents, for each subsample, the R^2 after excluding outliers.

B) Alpha model – modelling of heteroskedasticity

Up to this point, for the sake of more flexibility, all steps were conducted in Stata. Then, the data was loaded into the PovMap software and the Beta model specification set as described above. The next step, the specification of the Alpha model, serves to model the heteroskedasticity. In this model, the dependent variable is the household component of the residual. Due to the enormous amount of potential regressors for the Alpha model (income, income squared, all regressors of the Beta model, their interaction with income, and their interactions with income squared), I applied the automated selection model of the PovMap software, specifically "Stepwise Selection" and "Forward Selection" of regressors. In the end, all included regressors should be statistically significant and maximize together the R^2 of the household component of the residual.

3) Simulation

The last step of SAE is the simulation of household incomes based on the previously specified Beta and Alpha model. In addition, some choices regarding the simulation set-up have to be made, e.g. regarding the distributional form of the cluster and household error, trimming of simulated household incomes, and the simulation method. I specified normal distributions for both the cluster and the household error. Simulated incomes were trimmed at twice the maximum of the incomes of households in the survey (per domain) in order to avoid extreme outliers. This means that if a simulated income exceeds this trimming threshold, it is set to missing. Across both years and all domains, the share of simulated incomes affected by trimming ranged from 0.002% to 0.51%, with an average of 0.081%. The simulation method was 'Simultaneous drawing' and the number of replications was set to 100.

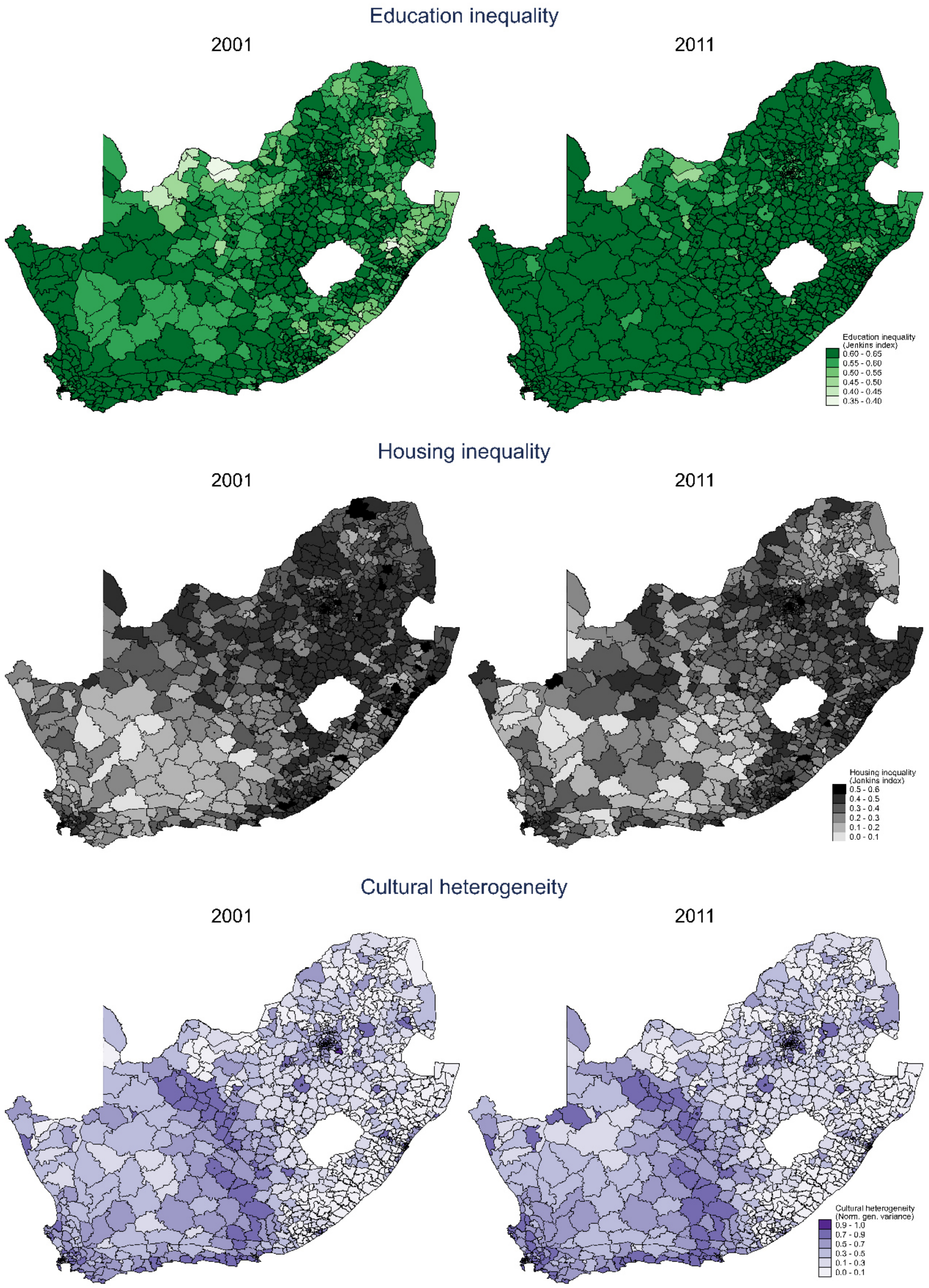
Processing of simulated income variables

The final products of SAE are 100 simulated income distributions for each domain in 2001 and 2011. These datafiles were then appended in order to cover the entire country. Next, variables of interest, i.e. mean per capita household income, the Gini index, mean log deviation and the Theil index were calculated in each police precinct for each of the 100 simulated income distributions. Lastly, for each variable of interest, the mean value was calculated from all 100 'simulated' variables, e.g. the mean of the Gini index for all 100 simulated income distributions. This mean value was then used in the econometric analysis of the study. As a plausibility check, the mean, median, 25th and 75th percentile and Gini index of the simulated household incomes were compared with the weighted equivalents in the survey for each domain. Generally, these statistics were very similar for the simulated census incomes in comparison to reported survey incomes.

Appendix D: Additional inequality and crime maps

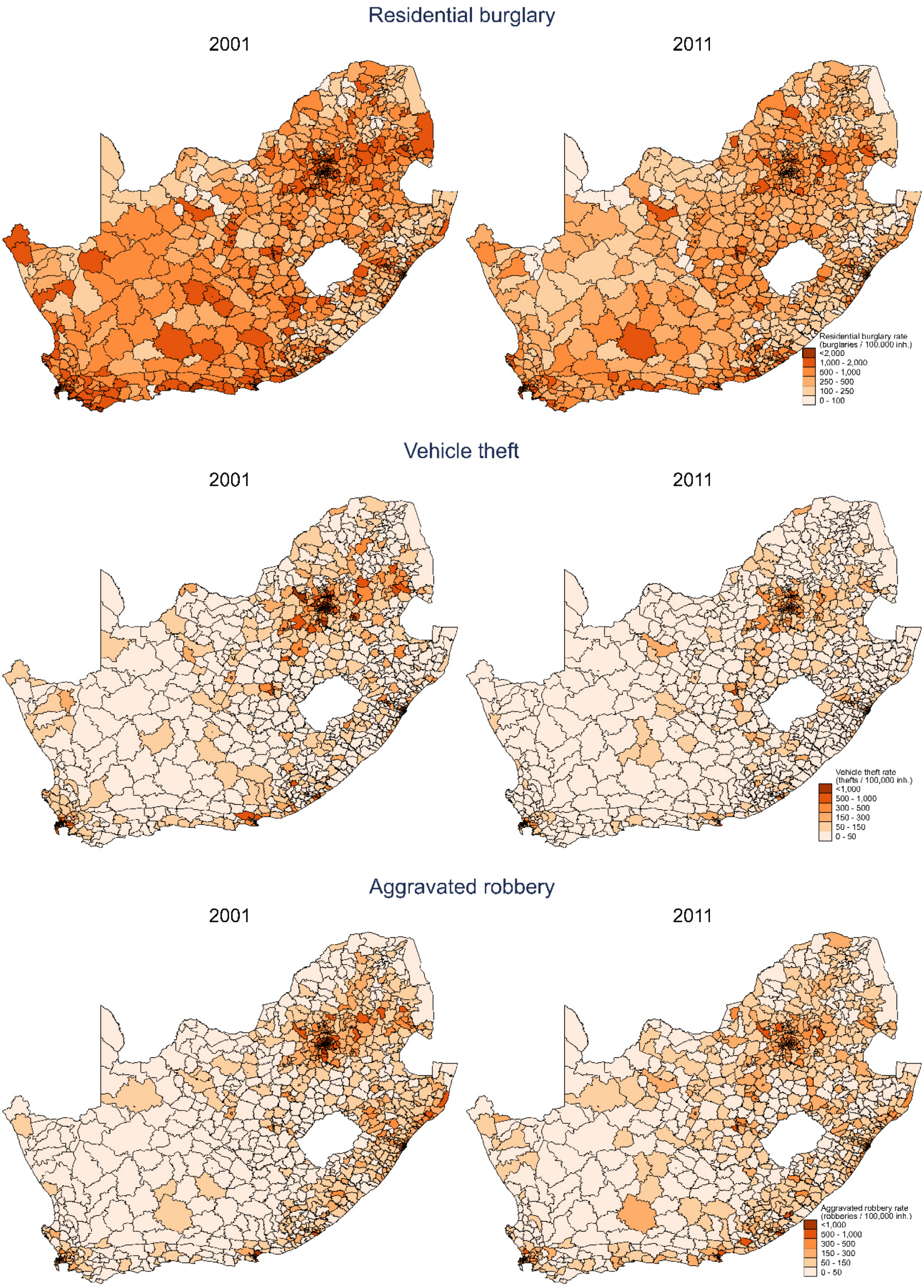
Figure A3 below maps education and housing inequality and cultural heterogeneity across South African police precincts in 2001 and 2011. Figures A4 and A5 on the following pages map rates of property crimes respectively violent crimes across South Africa.

Figure A3: Education inequality, housing inequality and cultural heterogeneity across South Africa



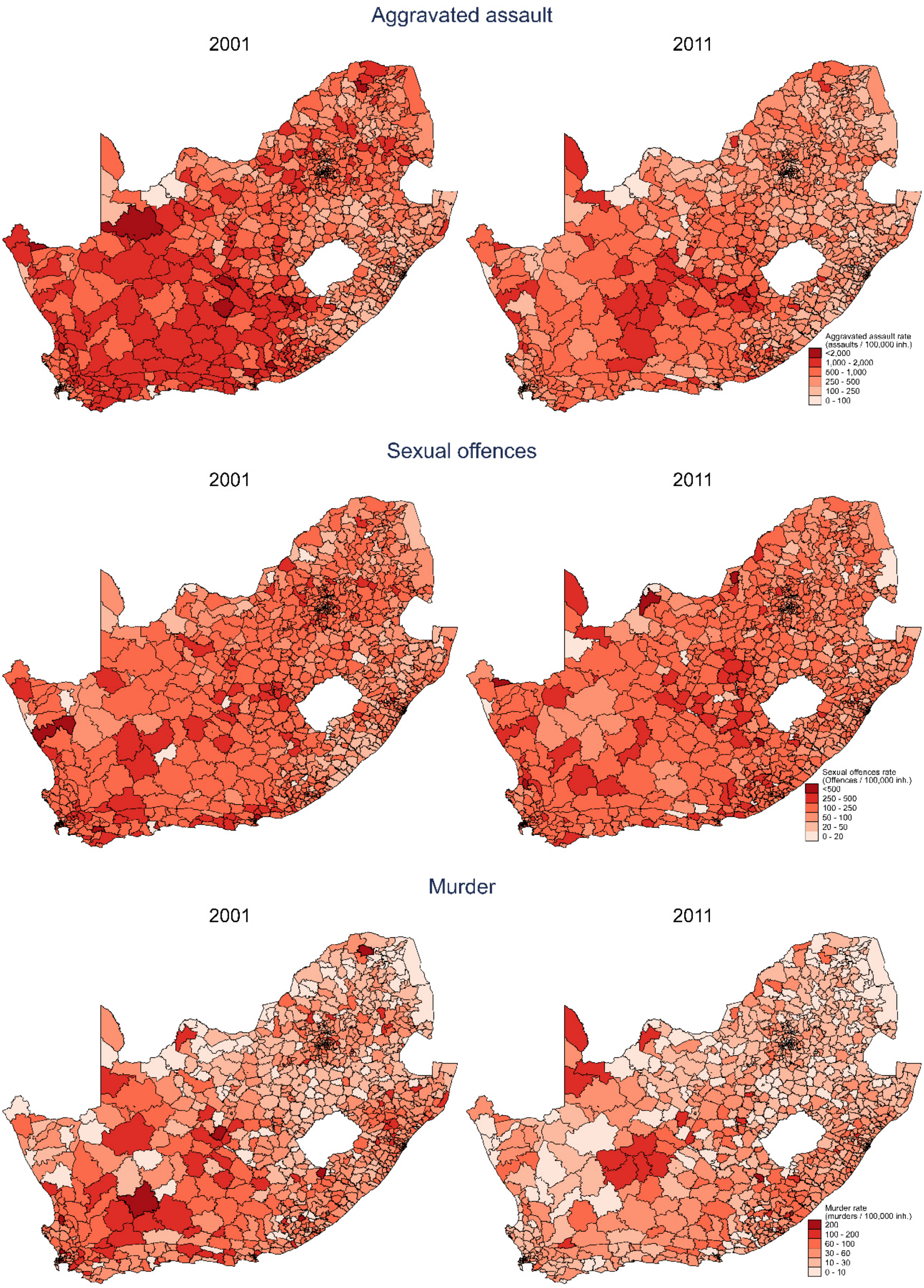
Notes: The darker green/grey/purple, the more education/housing/cultural inequality/heterogeneity.
Source: Own calculations using data from SAPS, StatsSA, and WorldPop.

Figure A4: Residential burglary, vehicle theft and aggravated robbery across South Africa



Notes: The darker orange, the higher the rate of residential burglary / vehicle theft / aggravated robbery.
Source: Own calculations using data from ISS and SAPS.

Figure A5: Aggravated assault, sexual offences and murder across South Africa



Notes: The darker red, the higher the rate of aggravated assault / sexual offences / murder.
Source: Own calculations using data from ISS and SAPS.

Appendix E: Within-estimator approach: regressions with police precinct-fixed effects

This section presents the results from regressions with police precinct- instead of police cluster-fixed effects, i.e. these regressions fully exploit the panel structure of the data set. Yet, due to the small within-variation of the inequality indices, this specification removes most of the variation in inequality and is thus not the preferred specification. As expected, the coefficients on inequality are generally insignificant, independent of the inequality dimension.

Table A7: Regressions of local crime rates on income inequality within police precincts

Independent variables	Property crimes			Violent crimes		
	Residential burglary	Vehicle theft	Aggravated robbery	Aggravated assault	Sexual offences	Murder
	(2)	(4)	(6)	(8)	(10)	(12)
Income inequality	0.167	-1.574	-1.260	1.151	1.600	-0.406
(Gini index)	(0.8400)	(1.2495)	(1.3017)	(0.8228)	(1.0623)	(1.6263)
Returns to crime	0.292	0.256	-0.351	-0.497**	-0.184	0.232
(ln mean p.c. hh income)	(0.2159)	(0.2328)	(0.2576)	(0.2382)	(0.2076)	(0.2716)
Costs of crime	-0.424**	0.218	-0.134	-0.412**	-0.455*	-0.072
(ln population density)	(0.1823)	(0.2519)	(0.3001)	(0.1981)	(0.2383)	(0.3171)
Returns from work	-0.638	0.805	0.852	-1.084*	-0.858	-0.271
(unemployment rate)	(0.4822)	(0.9721)	(1.0466)	(0.5712)	(0.7043)	(1.1445)
Crime prone groups	-2.339	0.927	0.272	3.511*	-0.272	1.080
(share of adolescents)	(1.9240)	(2.5356)	(2.8270)	(1.8116)	(2.5978)	(2.8260)
Social cohesion	-0.458	1.211	2.203	-0.002	2.100*	-0.074
(share of recently moved)	(0.6908)	(1.9168)	(1.8099)	(0.7917)	(1.1552)	(1.5918)
Family instability	2.172***	0.116	2.010	-0.642	-1.947**	-0.173
(Share of female-headed hh)	(0.7569)	(1.3081)	(1.4525)	(1.1325)	(0.9301)	(1.6215)
Urbanization	-0.057	0.084	0.380	0.292	0.591*	-1.084***
(share of urban hh)	(0.1741)	(0.3474)	(0.3311)	(0.2418)	(0.3183)	(0.3925)
Gang prevalence	0.836	-3.522*	1.465	1.099	2.497*	-0.489
(share of coloureds)	(0.6944)	(1.9581)	(1.8418)	(1.0246)	(1.3731)	(1.7949)
Province-specific time-effects	yes	yes	yes	yes	yes	yes
Spatial lag	yes	yes	yes	yes	yes	yes
Precinct-fixed effects	yes	yes	yes	yes	yes	yes
R-Squared	0.268	0.167	0.131	0.328	0.081	0.060
Observations	2,170	2,170	2,170	2,170	2,170	2,170

Notes: Each column refers to one regression. Dependent variables are IHS-transformed crime rates (crimes per 100,000 inhabitants). The spatial lag of the crime rate is instrumented with the spatial lags of first, second and third order of all independent variables. Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.
Source: Own calculations using data from ISS, SAPS, StatsSA, and WorldPop.

Table A8: Regressions of local crime rates on education inequality within police precincts

Independent variables	Property crimes			Violent crimes		
	Residential burglary	Vehicle theft	Aggravated robbery	Aggravated assault	Sexual offences	Murder
	(2)	(4)	(6)	(8)	(10)	(12)
Education inequality	-0.451	0.041	-0.875	0.740	3.036**	2.799
(Jenkins index)	(1.2701)	(1.8316)	(1.7910)	(1.7315)	(1.2930)	(1.9748)
Returns to crime	0.309	0.152	-0.407	-0.439**	-0.120	0.160
(ln mean p.c. hh income)	(0.1926)	(0.2367)	(0.2590)	(0.2142)	(0.1940)	(0.2550)
Costs of crime	-0.425**	0.141	-0.239	-0.345*	-0.345	-0.021
(ln population density)	(0.1805)	(0.2500)	(0.3022)	(0.1894)	(0.2266)	(0.2988)
Returns from work	-0.619	0.672	0.754	-0.949*	-0.610	-0.398
(unemployment rate)	(0.4493)	(0.9527)	(1.0348)	(0.5464)	(0.6796)	(1.0796)
Crime prone groups	-2.342	1.129	0.428	3.437*	-0.573	0.720
(share of adolescents)	(1.9407)	(2.5456)	(2.7887)	(1.8436)	(2.6162)	(2.7200)
Social cohesion	-0.448	1.519	2.521	-0.216	1.749	-0.236
(share of recently moved)	(0.7144)	(1.9468)	(1.8799)	(0.7747)	(1.1138)	(1.5485)
Family instability	2.185***	-0.096	2.011	-0.581	-1.840*	-0.317
(Share of female-headed hh)	(0.7561)	(1.3282)	(1.4645)	(1.1285)	(0.9393)	(1.5749)
Urbanization	-0.069	0.066	0.397	0.269	0.590*	-1.028***
(share of urban hh)	(0.1703)	(0.3560)	(0.3327)	(0.2394)	(0.3152)	(0.3725)
Gang prevalence	0.858	-3.411*	1.517	1.075	2.290*	-0.661
(share of coloureds)	(0.7064)	(1.9668)	(1.8437)	(1.0554)	(1.3537)	(1.7310)
Province-specific time-effects	yes	yes	yes	yes	yes	yes
Spatial lag	yes	yes	yes	yes	yes	yes
Precinct-fixed effects	yes	yes	yes	yes	yes	yes
R-Squared	0.265	0.175	0.125	0.317	0.092	0.116
Observations	2,170	2,170	2,170	2,170	2,170	2,170

Notes: Each column refers to one regression. Dependent variables are IHS-transformed crime rates (crimes per 100,000 inhabitants). The spatial lag of the crime rate is instrumented with the spatial lags of first, second and third order of all independent variables. Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.
Source: Own calculations using data from ISS, SAPS, StatsSA, and WorldPop.

Table A9: Regressions of local crime rates on housing inequality within police precincts

Independent variables	Property crimes			Violent crimes		
	Residential burglary	Vehicle theft	Aggravated robbery	Aggravated assault	Sexual offences	Murder
	(2)	(4)	(6)	(8)	(10)	(12)
Housing inequality	-0.154	-0.418	-0.928	0.173	0.147	-0.411
(Jenkins index)	(0.2480)	(0.6418)	(0.6494)	(0.2612)	(0.3938)	(0.6656)
Returns to crime	0.308	0.180	-0.399	-0.438**	-0.101	0.225
(ln mean p.c. hh income)	(0.1920)	(0.2393)	(0.2550)	(0.2234)	(0.1956)	(0.2578)
Costs of crime	-0.419**	0.128	-0.207	-0.352*	-0.370*	-0.104
(ln population density)	(0.1758)	(0.2465)	(0.2994)	(0.1805)	(0.2210)	(0.3080)
Returns from work	-0.618	0.632	0.697	-0.952*	-0.670	-0.299
(unemployment rate)	(0.4495)	(0.9513)	(1.0221)	(0.5454)	(0.6804)	(1.1077)
Crime prone groups	-2.308	1.286	0.863	3.369*	-0.476	1.341
(share of adolescents)	(1.9764)	(2.5785)	(2.8728)	(1.7900)	(2.6494)	(2.9105)
Social cohesion	-0.476	1.426	2.371	-0.151	1.921*	-0.010
(share of recently moved)	(0.6730)	(1.9688)	(1.8606)	(0.7598)	(1.1404)	(1.6376)
Family instability	2.183***	-0.061	1.806	-0.550	-1.796*	-0.202
(Share of female-headed hh)	(0.7486)	(1.3175)	(1.4559)	(1.1275)	(0.9454)	(1.6136)
Urbanization	-0.056	0.125	0.448	0.262	0.563*	-1.062***
(share of urban hh)	(0.1739)	(0.3551)	(0.3293)	(0.2468)	(0.3214)	(0.3988)
Gang prevalence	0.850	-3.435*	1.613	1.078	2.457*	-0.430
(share of coloureds)	(0.7001)	(1.9606)	(1.8438)	(1.0141)	(1.3624)	(1.7966)
Province-specific time-effects	yes	yes	yes	yes	yes	yes
Spatial lag	yes	yes	yes	yes	yes	yes
Precinct-fixed effects	yes	yes	yes	yes	yes	yes
R-Squared	0.266	0.167	0.133	0.323	0.080	0.046
Observations	2,170	2,170	2,170	2,170	2,170	2,170

Notes: Each column refers to one regression. Dependent variables are IHS-transformed crime rates (crimes per 100,000 inhabitants). The spatial lag of the crime rate is instrumented with the spatial lags of first, second and third order of all independent variables. Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.
Source: Own calculations using data from ISS, SAPS, StatsSA, and WorldPop.

Table A10: Regressions of local crime rates on cultural heterogeneity within police precincts

Independent variables	Property crimes			Violent crimes		
	Residential burglary	Vehicle theft	Aggravated robbery	Aggravated assault	Sexual offences	Murder
	(2)	(4)	(6)	(8)	(10)	(12)
Cultural heterogeneity	-0.743	0.619	-0.790	0.343	-0.995	-1.263
(Generalized variance, norm.)	(0.4824)	(0.7712)	(0.8609)	(0.4278)	(0.7442)	(0.8429)
Returns to crime	0.358*	0.112	-0.361	-0.458**	-0.025	0.304
(ln mean p.c. hh income)	(0.1874)	(0.2540)	(0.2706)	(0.2316)	(0.2026)	(0.2603)
Costs of crime	-0.452***	0.168	-0.245	-0.339*	-0.427*	-0.156
(ln population density)	(0.1747)	(0.2444)	(0.2984)	(0.1852)	(0.2401)	(0.3109)
Returns from work	-0.502	0.566	0.847	-1.005*	-0.498	-0.107
(unemployment rate)	(0.4380)	(0.9694)	(1.0650)	(0.5520)	(0.7148)	(1.1227)
Crime prone groups	-2.633	1.322	0.117	3.574*	-0.681	0.726
(share of adolescents)	(1.9525)	(2.5159)	(2.7727)	(1.8324)	(2.5652)	(2.8023)
Social cohesion	-0.220	1.274	2.661	-0.266	2.311**	0.421
(share of recently moved)	(0.6927)	(1.9466)	(1.8805)	(0.7285)	(1.1608)	(1.6735)
Family instability	2.104***	-0.012	1.839	-0.523	-1.943**	-0.338
(Share of female-headed hh)	(0.7391)	(1.3066)	(1.4432)	(1.1100)	(0.9431)	(1.6038)
Urbanization	-0.074	0.080	0.390	0.274	0.562*	-1.096***
(share of urban hh)	(0.1670)	(0.3583)	(0.3354)	(0.2407)	(0.3104)	(0.3920)
Gang prevalence	0.870	-3.461*	1.520	1.086	2.535*	-0.414
(share of coloureds)	(0.7006)	(1.9636)	(1.8510)	(1.0122)	(1.3734)	(1.8564)
Province-specific time-effects	yes	yes	yes	yes	yes	yes
Spatial lag	yes	yes	yes	yes	yes	yes
Precinct-fixed effects	yes	yes	yes	yes	yes	yes
R-Squared	0.268	0.174	0.129	0.322	0.086	0.052
Observations	2,170	2,170	2,170	2,170	2,170	2,170

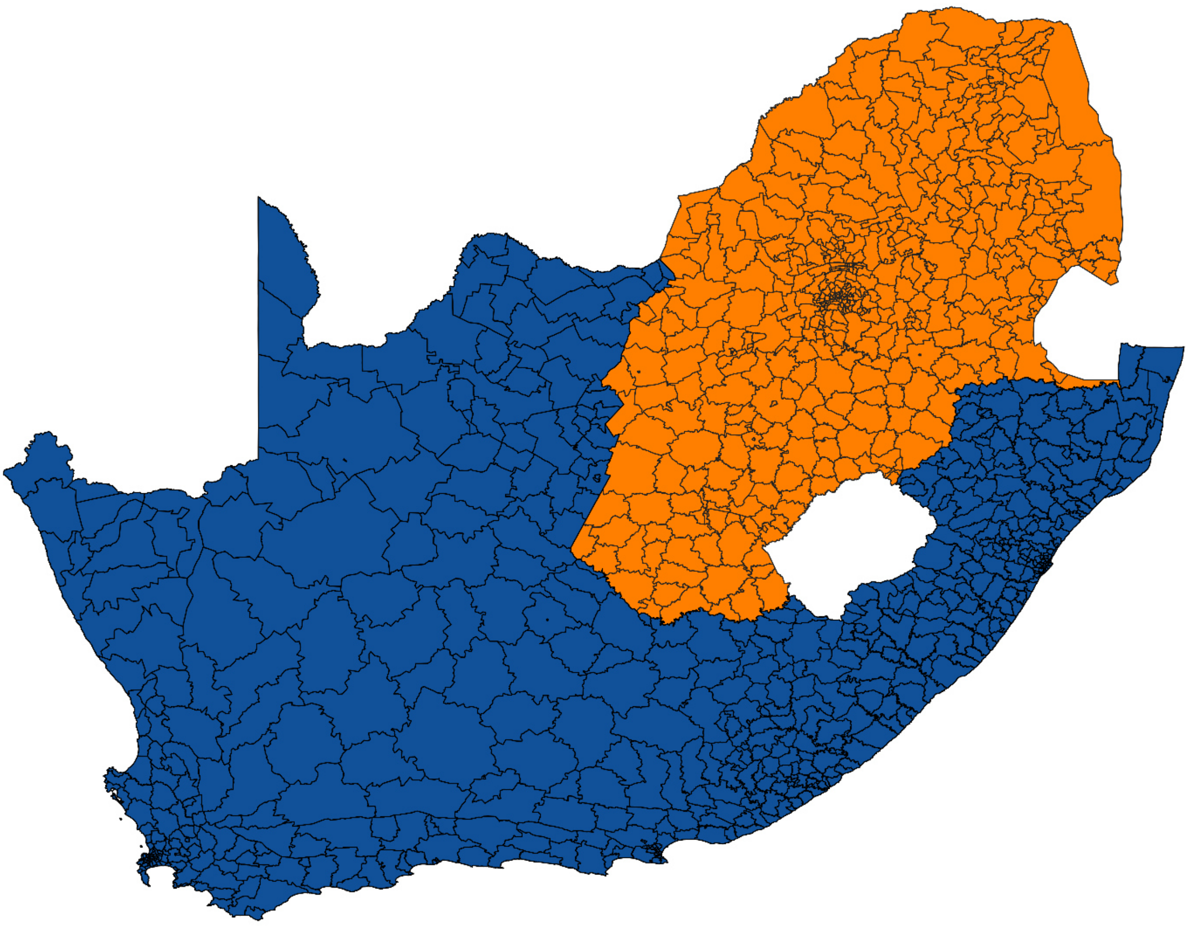
Notes: Each column refers to one regression. Dependent variables are IHS-transformed crime rates (crimes per 100,000 inhabitants). The spatial lag of the crime rate is instrumented with the spatial lags of first, second and third order of all independent variables. Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.
Source: Own calculations using data from ISS, SAPS, StatsSA, and WorldPop.

Appendix F: Heterogeneity analysis: geographic information on metropolitan municipalities and former colonies

This section provides information on the geographic pattern of British and Dutch colonization and the location of South Africa's metropolitan municipalities used in the heterogeneity analysis.

Figure A6 below classifies precincts according to the dominant former colonial power, i.e. whether the precincts were colonized the Dutch or the British Empire before the formation of the Union of South Africa in 1910. Blue precincts are formerly British dominated, i.e. they belonged to the Cape Colony or the Colony of Natal, and orange precincts are formerly Dutch dominated, i.e. they belonged to the Transvaal Republic or the Orange Free State.

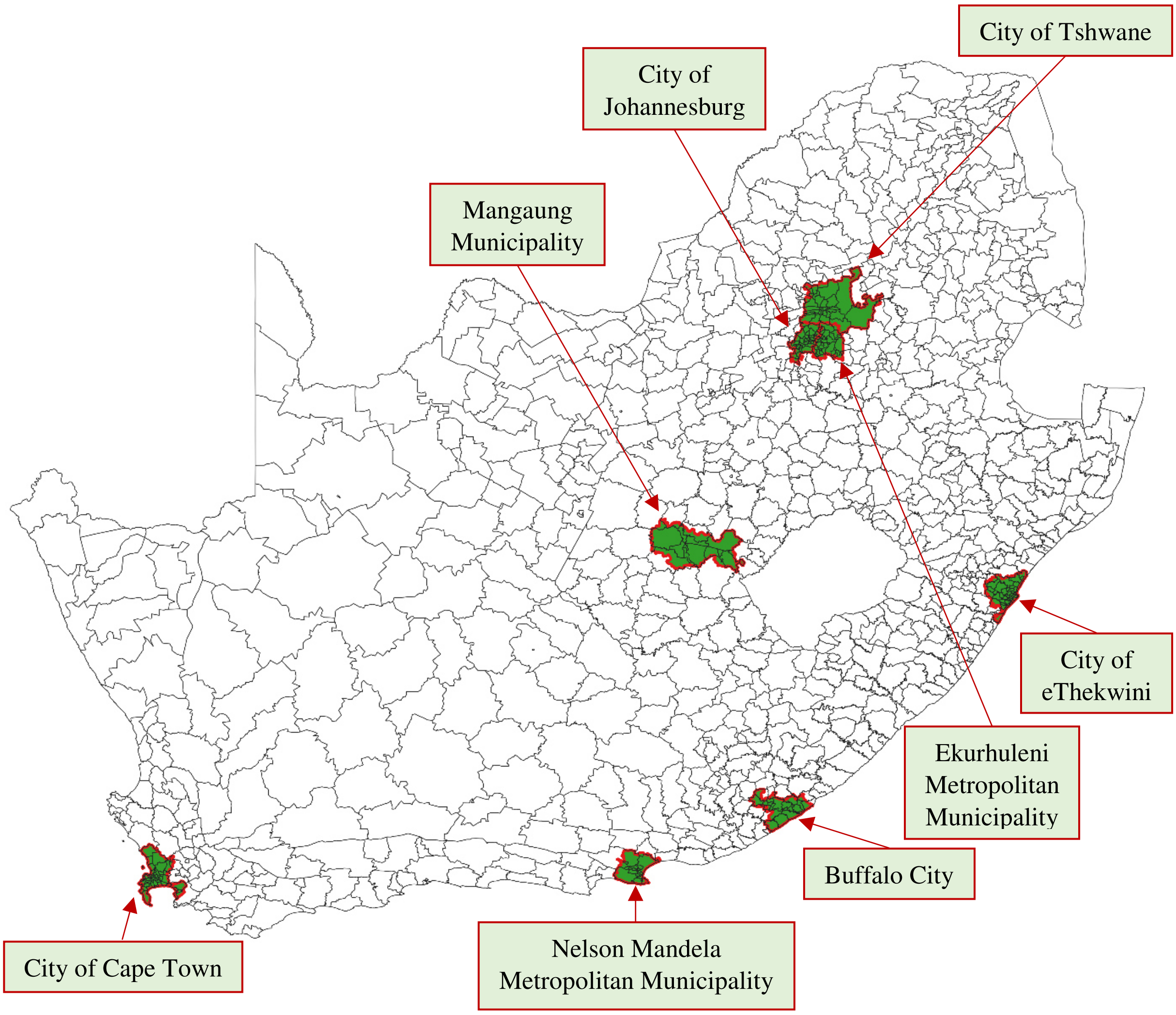
Figure A6: Division of police precincts by dominant former colonial power



Notes: Blue precincts are classified as formerly British dominated and orange precincts are classified as formerly Dutch dominated.
Source: Own calculations using data from SAPS.

Figure A7 below maps the locations of all eight metropolitan municipalities in relation to the country's 1,089 police precincts. Three are located in the province of Gauteng (City of Johannesburg, City of Tshwane, Ekurhuleni Metropolitan Municipality), two in the Eastern Cape (Buffalo City, Nelson Mandela Metropolitan Municipality), one in the Western Cape (City of Cape Town), one in the Free State (Mangaung Municipality), and one in KwaZulu-Natal (City of eThekwini). As visible in the map, the boundaries of the metropolitan municipalities frequently intersect with the boundaries of police precincts, i.e. some police precincts lie only partly inside a metropolitan municipality. For the heterogeneity analysis, I consider a police precinct as part of a metropolitan municipality if at least half of its area lies within the boundaries of a metropolitan municipality.

Figure A7: Locations of South Africa's eight metropolitan municipalities



Notes: Metropolitan municipalities are colored green and outlined red. Black lines denote police precinct boundaries.
Source: Own calculations using data from SAPS and StatsSA.

Appendix G: Robustness checks

In this section, I present the results from four types of robustness checks. On the next page, Table A11 shows bias-adjusted coefficients and degree of selection required for null effects, based on Oster (2019). Figure A8 shows the results from the randomization inference. Tables A12 and A13 show the results from regressions with the reduced and the expanded set of control variables, respectively. Lastly, Tables A14 to A16 show the results from regressions with log-transformed crimes, alternative inequality indices, and alternative grouping of education and housing categories, respectively.

A test of unobservable selection and coefficient stability

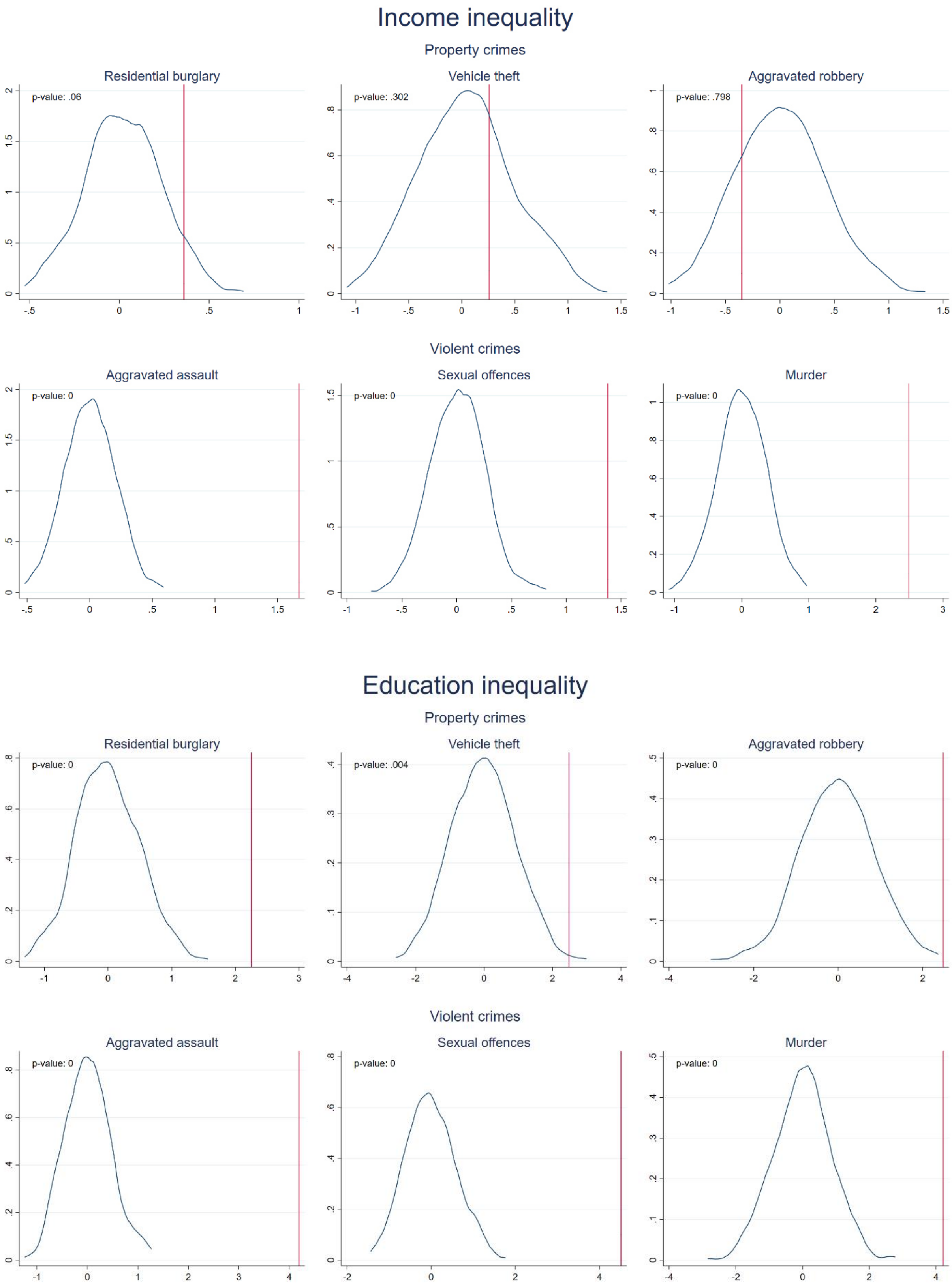
Table A11: Regressions of local crime rates on inequality within police precincts – bias-adjusted coefficients (β) and degree of selection required for null effects (δ)

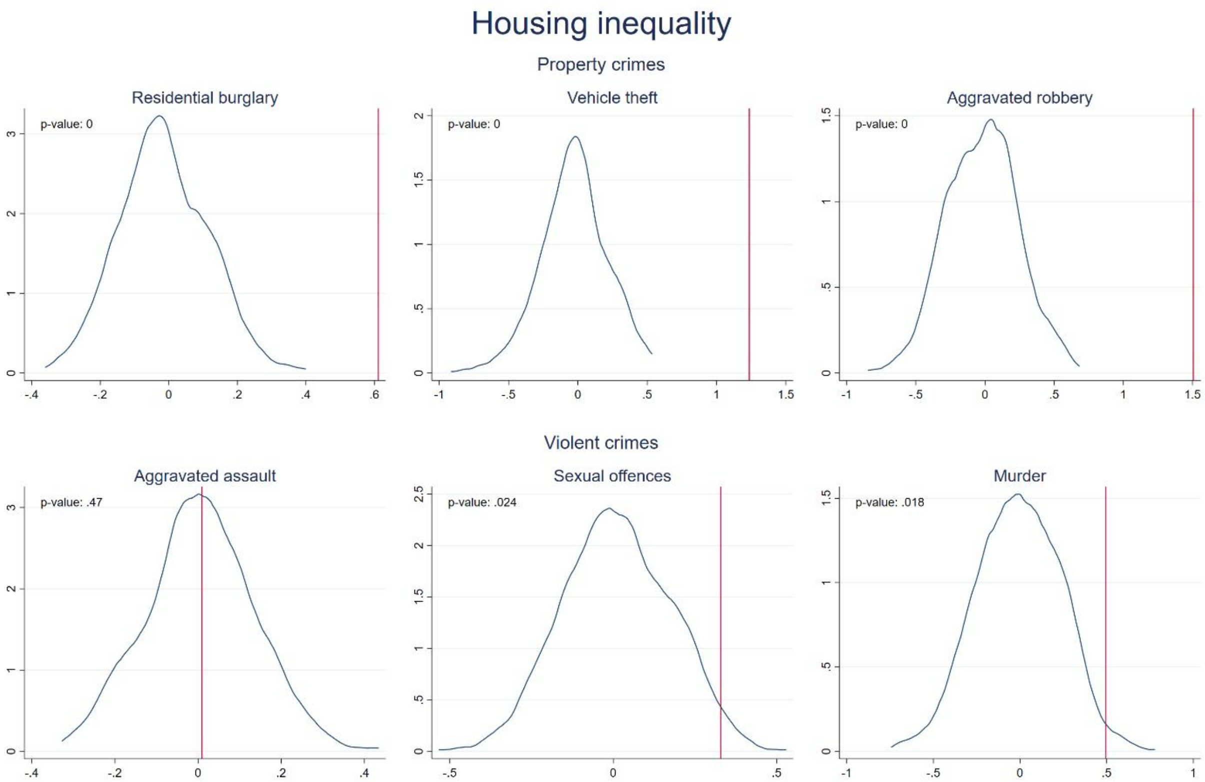
	Residential burglary		Vehicle theft		Aggravated robbery		Aggravated assault		Sexual offences		Murder	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Income inequality	-0.409	-0.315	0.948	0.681	12.497	2.972	2.663†	1.995†	3.530†	2.490†	10.920†	4.919†
(Gini index)	(0.610)	(0.660)	(-2.225)	(-3.134)	(0.105)	(0.225)	(2.716)	(4.610)	(-2.981)	(-4.408)	(-0.797)	(-1.732)
Education inequality	1.265†	1.353†	-0.745	0.305†	-5.159	-0.502	-0.540	2.187†	1.041†	2.967†	-2.748	2.134†
(Jenkins index)	(2.095)	(2.289)	(0.791)	(1.126)	(0.401)	(0.865)	(0.904)	(1.806)	(1.230)	(2.234)	(0.658)	(1.767)
Housing inequality	-0.256	-0.147	-1.306	-0.424	-0.724	0.559†	-0.643	-0.360	-0.184	0.022	-0.522	0.124
(Jenkins index)	(0.712)	(0.811)	(0.497)	(0.750)	(0.701)	(1.535)	(-0.012)	(-0.022)	(0.665)	(1.067)	(0.523)	(1.308)
Cultural heterogeneity	-5.116	-1.971	-16.146	-3.535	120.276†	27.109†	9.906†	4.152†	8.747†	4.011†	11.788†	3.704†
(Gen. var., norm.)	(0.092)	(0.195)	(0.144)	(0.396)	(0.428)	(0.906)	(-2.954)	(-4.871)	(2.221)	(3.334)	(-0.343)	(-0.785)
Assumption on R_{max}	$1.3\tilde{R}$	$\tilde{R} + (\tilde{R} - \dot{R})$	$1.3\tilde{R}$	$\tilde{R} + (\tilde{R} - \dot{R})$	$1.3\tilde{R}$	$\tilde{R} + (\tilde{R} - \dot{R})$	$1.3\tilde{R}$	$\tilde{R} + (\tilde{R} - \dot{R})$	$1.3\tilde{R}$	$\tilde{R} + (\tilde{R} - \dot{R})$	$1.3\tilde{R}$	$\tilde{R} + (\tilde{R} - \dot{R})$
Observations	2,170	2,170	2,170	2,170	2,170	2,170	2,170	2,170	2,170	2,170	2,170	2,170

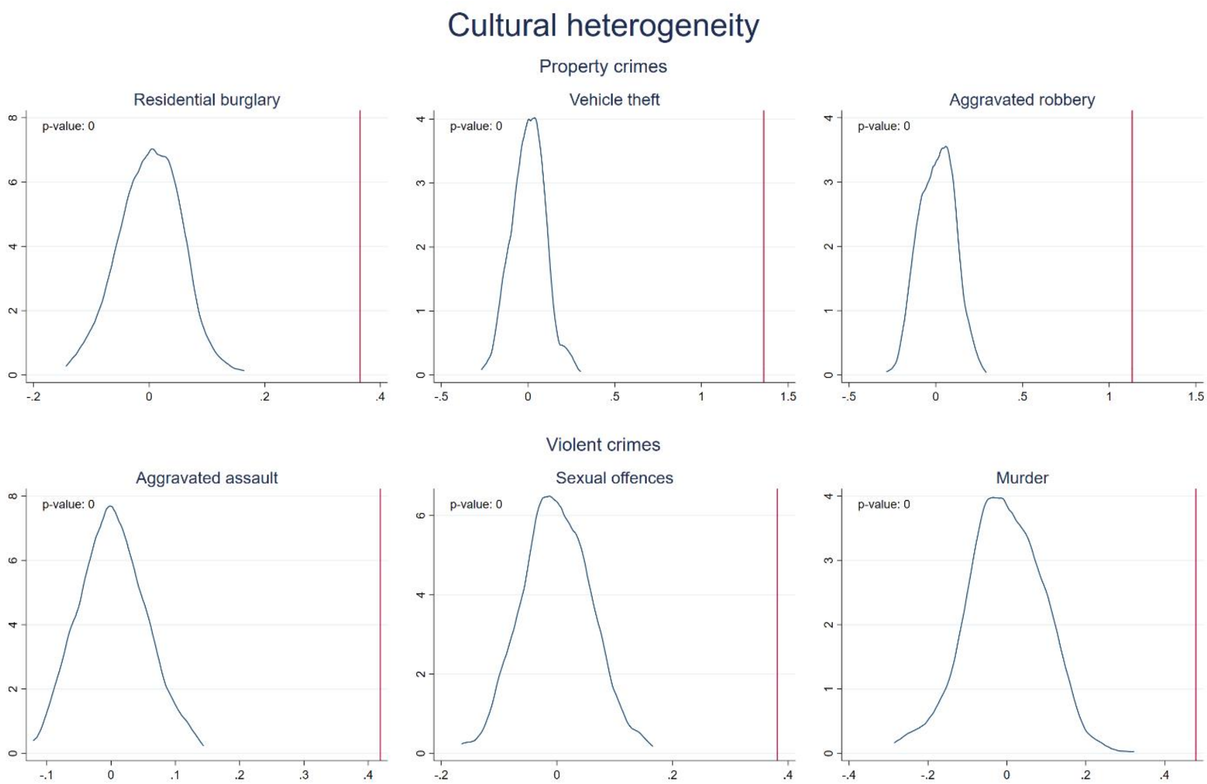
Notes: For each combination of crime type and inequality dimension, i.e. for each regression, I calculated a value for the bias-adjusted coefficient (β) and for the degree of selection required for null effects (δ). β is presented in the first row and δ in the row below in parentheses. The values are calculated based on the assumption that $R_{max} = 1.3\tilde{R}$ and alternatively $R_{max} = \tilde{R} + (\tilde{R} - \dot{R})$. † indicates robustness to unobservable selection. The reduced set of controls only comprises the spatial lag, cluster-fixed effects and province-specific time-effects. The full set of controls additionally includes ln mean p.c. household income, ln mean p.c. household income, ln population density, unemployment rate, share of adolescents, share of recently moved, share of female-headed hh, share of urban hh, share of coloureds. Dependent variables are IHS-transformed crime rates (crimes per 100,000 inhabitants). The spatial lag of the crime rate was not instrumented as the Stata command used for the calculation of β and δ, psacalc, does not support instrumental variable regressions. However, this should not influence the results in a meaningful way, as the results from the main regressions barely differ between instrumenting or not instrumenting the spatial lag of the crime rate.
Source: Own calculations using data from ISS, SAPS, StatsSA, and WorldPop.

Randomization inference

Figure A8: Randomization inference







Notes: The red line indicates the coefficient obtained with the true sample and the blue line the distribution of the coefficients from the permuted samples.
Source: Own calculations using data from ISS, SAPS, StatsSA, and WorldPop.

Robustness to reduction and expansion of the set of controls

Excluding potentially endogenous control variables

Table A12: Regressions of local crime rates on inequality within police precincts, excluding potentially endogenous controls

Independent variables	Property crimes			Violent crimes		
	Residential burglary	Vehicle theft	Aggravated robbery	Aggravated assault	Sexual offences	Murder
	(1)	(2)	(3)	(4)	(5)	(6)
Income inequality						
Gini index	0.781* (0.4059)	0.652 (0.7447)	0.595 (0.7688)	1.689*** (0.3916)	1.605*** (0.4727)	2.560*** (0.7117)
Education inequality						
Jenkins index	2.587*** (0.7271)	2.609** (1.3016)	2.668** (1.1813)	4.885*** (0.8726)	4.852*** (0.9627)	4.659*** (1.3152)
Housing inequality						
Jenkins index	0.604*** (0.2134)	1.220*** (0.3733)	1.423*** (0.3815)	-0.033 (0.1783)	0.305 (0.2609)	0.445 (0.3466)
Cultural heterogeneity						
Generalized variance (norm.)	0.283*** (0.1061)	1.204*** (0.2036)	0.816*** (0.2075)	0.476*** (0.1053)	0.365*** (0.1230)	0.458** (0.1911)
Province-specific time-effects	yes	yes	yes	yes	yes	yes
Spatial lag	yes	yes	yes	yes	yes	yes
Cluster-fixed effects	yes	yes	yes	yes	yes	yes
Reduced set of controls	yes	yes	yes	yes	yes	yes
Observations	2,170	2,170	2,170	2,170	2,170	2,170

Notes: Each column contains four regressions; one for each inequality dimension. The reduced set of controls includes ln mean p.c. household income, ln population density, share of adolescents, share of recently moved, share of urban hh. Dependent variables are IHS-transformed crime rates (crimes per 100,000 inhabitants). The spatial lag of the crime rate is instrumented with the spatial lags of first, second and third order of all independent variables. Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.
Source: Own calculations using data from ISS, SAPS, StatsSA, and WorldPop.

Adding further control variables

Table A13: Regressions of local crime rates on inequality within police precincts, additional controls

Independent variables	Property crimes			Violent crimes		
	Residential burglary	Vehicle theft	Aggravated robbery	Aggravated assault	Sexual offences	Murder
	(1)	(2)	(3)	(4)	(5)	(6)
Income inequality						
Gini index	0.319 (0.4095)	-0.679 (0.7764)	-1.609** (0.7801)	1.728*** (0.4008)	1.210** (0.4865)	1.669** (0.7655)
Education inequality						
Jenkins index	2.380*** (0.7153)	1.645 (1.3324)	1.756 (1.1868)	4.222*** (0.8627)	4.313*** (0.9531)	3.153** (1.3219)
Housing inequality						
Jenkins index	0.603*** (0.2129)	0.991*** (0.3633)	1.304*** (0.3598)	0.047 (0.1817)	0.265 (0.2606)	0.364 (0.3470)
Cultural heterogeneity						
Generalized variance (norm.)	0.340*** (0.1162)	1.163*** (0.2135)	0.905*** (0.2077)	0.430*** (0.1059)	0.345*** (0.1338)	0.375* (0.1971)
Province-specific time-effects	yes	yes	yes	yes	yes	yes
Spatial lag	yes	yes	yes	yes	yes	yes
Cluster-fixed effects	yes	yes	yes	yes	yes	yes
Expanded set of controls	yes	yes	yes	yes	yes	yes
Observations	2,170	2,170	2,170	2,170	2,170	2,170

Notes: Each column contains four regressions; one for each inequality dimension. The full set of controls includes ln mean p.c. household income, ln population density, unemployment rate, share of adolescents, share of recently moved, share of female-headed hh, share of urban hh, share of coloureds, dummy for coastal precincts, dummy for border precincts, ln distance to nearest metropolitan municipality, ln precinct area, dummy for split precincts. Dependent variables are IHS-transformed crime rates (crimes per 100,000 inhabitants). The spatial lag of the crime rate is instrumented with the spatial lags of first, second and third order of all independent variables. Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.
Source: Own calculations using data from ISS, SAPS, StatsSA, and WorldPop.

Further robustness checks

Log-transformed crime rates

Table A14: Regressions of local crime rates on inequality within police precincts, log crime rates

Independent variables	Property crimes			Violent crimes		
	Residential burglary	Vehicle theft	Aggravated robbery	Aggravated assault	Sexual offences	Murder
	(1)	(2)	(3)	(4)	(5)	(6)
Income inequality						
Gini index	0.154	0.023	-0.386	1.317***	0.810**	1.748***
	(0.3621)	(0.7116)	(0.5953)	(0.3091)	(0.3312)	(0.5637)
Education inequality						
Jenkins index	1.992***	2.144*	1.543*	3.687***	3.257***	3.382***
	(0.5756)	(1.1944)	(0.8642)	(0.6129)	(0.6035)	(0.9632)
Housing inequality						
Jenkins index	0.604***	1.132***	1.064***	0.140	0.242	0.398
	(0.1728)	(0.3329)	(0.2683)	(0.1326)	(0.1598)	(0.2506)
Cultural heterogeneity						
Generalized variance (norm.)	0.386***	1.294***	0.890***	0.332***	0.327***	0.372***
	(0.0958)	(0.1911)	(0.1555)	(0.0870)	(0.0907)	(0.1443)
Province-specific time-effects	yes	yes	yes	yes	yes	yes
Spatial lag	yes	yes	yes	yes	yes	yes
Cluster-fixed effects	yes	yes	yes	yes	yes	yes
Full set of controls	yes	yes	yes	yes	yes	yes
Observations	2,170	2,170	2,170	2,170	2,170	2,170

Notes: Each column contains four regressions; one for each inequality dimension. The full set of controls includes ln mean p.c. household income, ln population density, unemployment rate, share of adolescents, share of recently moved, share of female-headed hh, share of urban hh, share of coloureds. Dependent variables are log-transformed crime rates (crimes per 100,000 inhabitants). The spatial lag of the crime rate is instrumented with the spatial lags of first, second and third order of all independent variables. Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.
Source: Own calculations using data from ISS, SAPS, StatsSA, and WorldPop.

Alternative inequality indices

Table A15: Regressions of local crime rates on inequality within police precincts, alternative indices

Independent variables	Property crimes			Violent crimes		
	Residential burglary	Vehicle theft	Aggravated robbery	Aggravated assault	Sexual offences	Murder
	(1)	(2)	(3)	(4)	(5)	(6)
Income inequality						
Mean log deviation	-0.032 (0.1179)	-0.089 (0.2446)	-0.392 (0.2431)	0.333** (0.1577)	0.365*** (0.1410)	0.659*** (0.2369)
Theil index	-0.068 (0.1176)	-0.139 (0.2170)	-0.304 (0.2274)	0.218* (0.1187)	0.207 (0.1337)	0.453** (0.2252)
Education inequality						
Cowell-Flachaire-index (α=0)	1.303*** (0.4187)	1.444* (0.7481)	1.889*** (0.6935)	2.036*** (0.4858)	2.302*** (0.5160)	1.872** (0.7676)
Cowell-Flachaire-index (α=0.25)	0.712*** (0.2272)	0.714* (0.4295)	0.642* (0.3807)	1.343*** (0.2675)	1.459*** (0.3058)	1.423*** (0.4285)
Cowell-Flachaire-index (α=0.5)	1.154*** (0.3594)	1.202* (0.6638)	1.245** (0.5975)	2.086*** (0.4218)	2.278*** (0.4701)	2.127*** (0.6691)
Cowell-Flachaire-index (α=0.75)	0.712*** (0.2272)	0.714* (0.4295)	0.642* (0.3807)	1.343*** (0.2675)	1.459*** (0.3058)	1.423*** (0.4285)
Cowell-Flachaire-index (α=0.9)	0.316*** (0.1034)	0.309 (0.1986)	0.250 (0.1744)	0.610*** (0.1218)	0.662*** (0.1415)	0.661*** (0.1968)
Housing inequality						
Cowell-Flachaire-index (α=0)	0.728*** (0.2202)	1.657*** (0.3780)	1.676*** (0.3802)	0.019 (0.1658)	0.407 (0.2678)	0.267 (0.3501)
Cowell-Flachaire-index (α=0.25)	0.187*** (0.0717)	0.376*** (0.1256)	0.507*** (0.1270)	0.012 (0.0586)	0.077 (0.0861)	0.187 (0.1191)
Cowell-Flachaire-index (α=0.5)	0.378*** (0.1326)	0.788*** (0.2310)	0.972*** (0.2337)	0.022 (0.1060)	0.177 (0.1604)	0.309 (0.2174)
Cowell-Flachaire-index (α=0.75)	0.187*** (0.0717)	0.376*** (0.1256)	0.507*** (0.1270)	0.012 (0.0586)	0.077 (0.0861)	0.187 (0.1191)
Cowell-Flachaire-index (α=0.9)	0.074** (0.0300)	0.146*** (0.0527)	0.208*** (0.0532)	0.005 (0.0248)	0.028 (0.0359)	0.082 (0.0502)
Province-specific time-effects	yes	yes	yes	yes	yes	yes
Spatial lag	yes	yes	yes	yes	yes	yes
Cluster-fixed effects	yes	yes	yes	yes	yes	yes
Full set of controls	yes	yes	yes	yes	yes	yes
Observations	2,170	2,170	2,170	2,170	2,170	2,170

Notes: Each column contains twelve regressions; two for income inequality, and five for education and housing inequality, respectively. The full set of controls includes ln mean p.c. household income, ln population density, unemployment rate, share of adolescents, share of recently moved, share of female-headed hh, share of urban hh, share of coloureds. Dependent variables are IHS-transformed crime rates (crimes per 100,000 inhabitants). The spatial lag of the crime rate is instrumented with the spatial lags of first, second and third order of all independent variables. Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.
Source: Own calculations using data from ISS, SAPS, StatsSA, and WorldPop.

Alternative grouping/ordering of education and housing categories

Table A16: Regressions of local crime rates on inequality within police precincts, alternative grouping

Independent variables	Property crimes			Violent crimes		
	Residential burglary	Vehicle theft	Aggravated robbery	Aggravated assault	Sexual offences	Murder
	(1)	(2)	(3)	(4)	(5)	(6)
Education inequality						
Jenkins index (8 cat.)	2.685***	4.287**	4.056***	5.457***	5.188***	5.083***
	(0.8140)	(1.7088)	(1.5254)	(1.1562)	(1.2153)	(1.6897)
Jenkins index (5 cat.)	1.778**	4.474**	0.916	4.340***	3.254***	2.506
	(0.7365)	(1.8777)	(1.6529)	(1.1564)	(1.1694)	(1.7315)
Housing inequality						
Jenkins index (7 cat.,	0.617***	1.227***	1.472***	0.007	0.334	0.500
alt. order)	(0.2040)	(0.3697)	(0.3772)	(0.1731)	(0.2570)	(0.3497)
Jenkins index (5 cat.)	0.680***	1.263***	1.471***	-0.025	0.423	0.585
	(0.2049)	(0.3854)	(0.3958)	(0.1889)	(0.2632)	(0.3627)
Province-specific time-effects	yes	yes	yes	yes	yes	yes
Spatial lag	yes	yes	yes	yes	yes	yes
Cluster-fixed effects	yes	yes	yes	yes	yes	yes
Full set of controls	yes	yes	yes	yes	yes	yes
Observations	2,170	2,170	2,170	2,170	2,170	2,170

Notes: Each column contains four regressions; two for education and housing inequality, respectively. The full set of controls includes ln mean p.c. household income, ln population density, unemployment rate, share of adolescents, share of recently moved, share of female-headed hh, share of urban hh, share of coloureds. Dependent variables are IHS-transformed crime rates (crimes per 100,000 inhabitants). The spatial lag of the crime rate is instrumented with the spatial lags of first, second and third order of all independent variables. Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.
Source: Own calculations using data from ISS, SAPS, StatsSA, and WorldPop.

References

Demombynes, G. and B. Özler (2005), Crime and local inequality in South Africa, *Journal of Development Economics*, 76(2): 265-292.

Elbers, C., J. O. Lanjouw and P. Lanjouw (2003), Micro-level estimation of poverty and inequality, *Econometrica,*, 71(1): 355-364.

Li, X., Y. Zhou, M. Zhao and X. Zhao (2020), A harmonized global nighttime light dataset 1992–2018, *Scientific Data*, 7(168): 1-9.

Oster, E. (2019), Unobservable selection and coefficient stability: Theory and evidence, *Journal of Business & Economic Statistics*, 37(2): 187-204.

Zhao, Q. and P. Lanjouw (n.d.), *Using PovMap2: A user's guide*, World Bank, Washington, DC.

Volkswirtschaftliche Reihe der Passauer Diskussionspapiere

Bisher sind erschienen:

V-1-98	Gerhard Rübel, Can adjustments to working hours help reduce unemployment?
V-2-98	Martin Werding, Pay-as-you-go Public Pension Schemes and Endogenous Fertility: The Reconstruction of Intergenerational Exchange
V-3-98	Carsten Eckel, International Trade, Direct Investment, and the Skill Differential in General Equilibrium
V-4-98	Reinar Lüdeke, Das Staatsbudget und intergenerationelle Umverteilung, Das Staatsvermögen als Instrument intergenerativer Verteilungspolitik und der "generational accounting"-Ansatz: Alter Wein in neuen (höherwertigen) Schläuchen?
V-5-98	Anja Klüver und Gerhard Rübel, Räumliche Industriekonzentration und die komparativen Vorteile von Ländern - eine empirische Studie der Europäischen Union
V-6-98	Klaus Beckmann und Elisabeth Lackner, Vom Leviathan und von optimalen Steuern
V-7-98	Martin Werding, The Pay-as-you-go Mechanism as Human Capital Funding: The "Mackenroth hypothesis" Revisited
V-8-98	Reinar Lüdeke und Klaus Beckmann, Social Costs of Higher Education: Production and Financing. The Case of Germany (1994)
V-9-98	Gerhard Rübel, "Faire" Löhne und die Flexibilität von Arbeitsmärkten in einem Zwei-Sektoren-Modell
V-10-98	Klaus Beckmann, Notizen zum Steueranteil von Rentenversicherungsbeiträgen im Umlageverfahren
V-11-98	Christian Jasperneite und Hans Joachim Allinger, Trendwende am westdeutschen Arbeitsmarkt? - Eine ökonometrische Analyse
V-12-98	Christian Jasperneite und Hans Joachim Allinger, Langfristige Perspektiven für den westdeutschen Arbeitsmarkt: Was sagen die Gesetze von Okun und Verdoorn?
V-13-98	Hans Joachim Allinger und Christian Jasperneite, Saisonbereinigung von Arbeitsmarktdaten bei aktiver Arbeitsmarktpolitik
V-14-99	Reinar Lüdeke und Klaus Beckmann, Hochschulbildung, Humankapital und Beruf: Auswertung einer Längsschnittsbefragung Passauer Absolventen 1988 - 1998

V-15-99 Gerhard Rübel, Volkseinkommenssteigerung durch ausgabenfinanzierte Steuersenkung - Eine Umkehrung des Haavelmo-Theorems für offene Volkswirtschaften

V-16-99 Silke Klüver, Konzentrationsursachen in der europäichen Versicherungsbranche - eine empirische Untersuchung

V-17-99 Reinar Lüdeke, Familienlastenausgleich, Elternleistungsausgleich und die Neufundierung der umlagefinanzierten Altersversorgung

V-18-99 Anja Klüver und Gerhard Rübel, Industrielle Konzentration als Kriterium für die Geeignetheit eines einheitlichen Währungsraums – Eine empirische Untersuchung der Europäischen Union von 1972 bis 1996

V-19-00 Carsten, Eckel, Fragmentation, Efficiency-seeking FDI, and Employment

V-20-00 Christian Jasperneite, Understanding Hysteresis in Unemployment: The German Case

V-21-00 Jörg Althammer, Reforming Family Taxation

V-22-00 Carsten Eckel, Labor Market Adjustments to Globalization: Unemployment versus Relative Wages

V-23-00 Klaus Beckmann, Tax Competition through Tax Evasion

V-24-01 Klaus Beckmann, Steuerhinterziehung, begrenzte Rationalität und Referenzabhängigkeit: Theorie und experimentelle Evidenz

V-25-01 Klaus Beckmann, Solidarity, Democracy, and Tax Evasion: an Experimental Study

V-26-04 Michael Fritsch, Udo Brixy und Oliver Falck, The Effect of Industry, Region and Time on New Business Survival - A Multi-Dimensional Analysis

V-27-04 Gerhard D. Kleinhenz, Bevölkerung und Wachstum - Die Bevölkerungsentwicklung in Deutschland als Herausforderung für Wirtschafts- und Sozialpolitik

V-28-04 Johann Graf Lambsdorf, The Puzzle with Increasing Money Demand - Evidence from a Cross-Section of Countries

V-29-04 Frauke David, Oliver Falck, Stephan Heblich und Christoph Kneiding, Generationsgerechtigkeit und Unternehmen

V-30-04 Roland Engels†, Zur mikroökonomischen Fundierung der Geldnachfrage in allgemeinen Gleichgewichtsmodellen

V-31-05 Johann Graf Lambsdorff, Between Two Evils – Investors Prefer Grand Corruption!

V-32-05 Oliver Falck, Das Scheitern junger Betriebe – Ein Überlebensdauermodell auf Basis des IAB-Betriebspanels

V-33-05 Raphaela Seubert - On the Nature of the Corrupt Firm: Where to Situate Liability?

V-34-05 Johann Graf Lambsdorff – Consequences and Causes of Corruption – What do We Know from a Cross-Section of Countries?

V-35-05 Stephan Heblich - Arbeitszeitflexibilisierung Revisited

V-36-05 Oliver Falck und Stephan Heblich - Das Konzept der eigenverantwortlichen Generation zur Bewältigung des demographischen Wandels

V-37-05 Florian Birkenfeld, Daniel Gastl, Stephan Heblich, Ferry Lienert, Mascha Maergoyz, Oksana Mont und Andrius Plepys - Product ban versus risk management by setting emission and technology requirements – the effect of different regulatory schemes taking the use of trichloroethylene in Sweden and Germany as an example

V-38-05 Johann Graf Lambsdorff - Determining Trends for Perceived Levels of Corruption

V-39-05 Oliver Falck - Mayflies and Long-Distance Runners: The Effects of New Business Formation on Industry Growth

V-40-05 Johann Graf Lambsdorff und Christian Engelen - Hares and Stags in Argentinean Debt Restructuring

V-41-05 Johann Graf Lambsdorff und Mathias Nell – Let Them Take Gifts, and Cheat Those Who Seek Influence

V-42-06 Hans Joachim Allinger – Bürgerversicherung und Kopfpauschale haben vieles gemeinsam – Anmerkungen zur Diskussion einer Reform der gesetzlichen Krankenversicherung

V-43-06 Michael Schinke und Johann Graf Lambsdorff - Insider Trading among Central Bankers – a Treatise on Temptation and Policy Choice

V-44-06 Johann Graf Lambsdorff und Hady Fink - Combating Corruption in Colombia: Perceptions and Achievements

V-45-06 Oliver Falck und Stephan Heblich - Corporate Social Responsibility: Einbettung des Unternehmens in das Wirtschaftssystem

V-46-06 Johann Graf Lambsdorff und Luka Bajec - There Is No Bank Lending Channel!

V-47-06 Christian Engelen und Johann Graf Lambsdorff - Das Keynesianische Konsensmodell

V-48-07 Stephan Heblich - Eigenverantwortliche Individuen und Pro-Aktive Unternehmen

V-49-07 Christian Engelen und Johann Graf Lambsdorff - Das Keynesianische Konsensmodell einer offenen Volkswirtschaft

V-50-07 Christian Engelen und Johann Graf Lambsdorff - Fairness in Sovereign Debt Restructuring

V-51-07 Johann Graf Lambsdorff und Björn Frank - Corrupt Reciprocity - an Experiment

V-52-07 Mathias Nell - Strategic Aspects of Voluntary Disclosure Programs for Corruption Offences - Towards a Design of Good Practice -

V-53-07 Mathias Nell - Contracts Induced by Means of Bribery - Should they be Void or Valid?

V-54-08 Michael Pflüger – Die Neue Ökonomische Geographie: Ein Überblick

V-55-08 Florian Birkenfeld und Shima'a Hanafy - Wie zentral sind die Abschlussprüfungen an deutschen Schulen wirklich?"

V-56-08 Florian Birkenfeld - Kleine Klassen und gute Luft - Warum sind die Grundschulen auf dem Land besser?

V-57-08 Johann Graf Lambsdorff – The Organization of Anticorruption – Getting Incentives Right!

V-58-09 Oliver Farhauer und Alexandra Kröll - Verfahren zur Messung räumlicher Konzentration und regionaler Spezialisierung in der Regionalökonomik

V-59-09 Oliver Farhauer und Alexandra Kröll - Die Shift-Share-Analyse als Instrument der Regional- und Clusterforschung

V-60-10 Johann Graf Lambsdorff – Deterrence and Constrained Enforcement – Alternative Regimes to Deal with Bribery

V-61-10 Johann Graf Lambsdorff – Who Accepts Bribery? Evidence from a Global Household Survey

V-62-10 Oliver Farhauer und Alexandra Kröll – What We Can and What We Can't Say About Employment Growth in Specialised Cities

V-63-11 Johann Graf Lambsdorff, Manuel Schubert und Marcus Giamattei – On the Role of Heuristics - Experimental Evidence on Inflation Dynamics

V- 64-12 Manuel Schubert und Johann Graf Lambsdorff – On the Costs of Kindness - An Experimental Investigation of Guilty Minds and Negative Reciprocity

V-65-12 Manuel Schubert – Deeds rather than Omissions - How Intended Consequences Provoke Negative Reciprocity

V-66-12 Agnes Binagwaho, Renate Hartwig, Denyse Ingeri und Andrew Makaka – Mutual Health Insurance and its Contribution to Improving Child Health in Rwanda

V-67-15 Marcus Giamattei – Cold Turkey vs. Gradualism – Evidence on Disinflation; Strategies from Laboratory Experiment

V-68-15 Marcus Giamattei, Johann Graf Lambsdorff – classEx – an Online Software for Classroom Experiments

V-69-15 Johann Graf Lambsdorff – Preventing Corruption by Promoting Trust – Insights from Behaviorial Science

V-70-16 Lukas Schötz – Power and Rents – political influence on rent extraction. Quasi-experimental evidence for Bavaria

V-71-16 Lukas Schötz – Mayor Games in Bavaria: Self Selection of Local Politicians is not influenced by constitutionally defined remuneration increases

V-72-16 Johann Graf Lambsdorff, Marcus Giamattei, Katharina Werner, Manuel Schubert-Emotion vs. Cognition – Experimental Evidence on Cooperation from the 2014 Soccer World Cup

V-73-16 Katharina Werner – Whom do People Trust after a Violent Conflict? - Experimental Evidence from Maluku, Indonesia

V-74-16 Katharina Werner, Johann Graf Lambsdorff – Emotional Numbing and Lessons Learned after a Violent Conflict – Experimental Evidence from Ambon, Indonesia

V-75-17 Ramona Molitor – Publicly Announced Speed Limit Enforcement and Its Impact on Road Safety – Evidence From the German Blitzmarathons

V-76-17 Johannes Lorenz – Population Dynamics of Tax Avoidance with Crowding Effects

V-77-19 Katrin Huber – The role of the career costs of children for the effect of public child care on fertility and maternal employment

V-78-19 Markus Fritsch, Andrew Adrian Yu Pua, Joachim Schnurbus – Revisiting habits and heterogeneity in demands

V-79-20 Susanna Grundmann – Do just deserts and competition shape patterns of cheating?

V-80-20 Susanna Grundmann, Marcus Giamattei, Johann Graf Lambsdorff – On the downward rigidity of wages – evidence from an experimental labour market with monetary neutrality

V-81-20 Michael Grimm, Renate Hartwig, Ann-Kristin Reitmann, Fadima Yaya Bocoum – Can informal redistribution withstand formal safety nets? Insights from urban-rural transfers in Burkina Faso

V-82-20 Ann-Kristin Reitmann – Changing environmental conservation attitudes: Evidence from a framed field experiment among small-scale coffee farmers in Colombia

V-83-20 Eleonora Nillesen, Michael Grimm, Micheline Goedhuys, Ann-Kristin Reitmann, Aline Meysonnat – On the malleability of gender attitudes: Evidence from implicit and explicit measures in Tunisia

V-84-21 Manuela Fritz – Temperature and non-communicable diseases: Evidence from Indonesia’s primary health care system.

V-85-21 Katharina Grabrucker – Effects of internal rural-urban migration on rural non-farm enterprises: Evidence from Thailand and Vietnam

V-86-21 Pamina Koenig, Sebastian Krautheim, Claudius Löhnert, Thierry Verdier – Local Global Watchdogs: Trade, Sourcing and the Internationalization of Social Activism

V-87-22 Claudius Löhnert – Firms under Pressure: International Trade and Social Activism

V-88-22 Claudius Löhnert – Plausible Deniability: Integration vs. Outsourcing with Heterogeneous Firms and Unethical Suppliers