RATIONALE

in a world without reason, turn to economics

1

Contents

1) **Heterogeneous Effects of Unemployment on Mental Health, Depending on Gender and Refugee Status: Evidence from Germany and the United States**
Anmol Kanotra, Kerry Neitzel, Maithili Modi, Xuyi Yang, Abi Thomas 1

2) **The Hedonic House Price Model in a Development Context – Evidence from the National Capital Region of the Philippines, 2000 - 2010**
Christopher Dann, Jakub Grohmann, Amy Ho, Sherman Khoo, Anna Kim, Swapnil Lal 23

3) **The Economic Consequences of Crime: Evidence from the *Unidades de Polícia Pacificadora* in Rio de Janeiro**
Savannah Kochinke, Edgar Akopyan, Somjeeta Chatterjee, Ardi Janjeva, Napat Matchaseth 59

4) **Anocracy in Ankara: Co-party Bias and the Turkish Housing Development Administration (TOKI), 2002 – 2016**
Ethan Caspi, Christopher Dann, Lutfi Sun, Andre Zeromski, Yihan Zhu 98

Preface

Although historically Rationale has been an 'Economist' zine-styled publication, the Economics Society's Research Division is incredibly proud to relaunch Rationale as the first-ever economics undergraduate/graduate working paper series at the London School of Economics. As such, the mission of the 'new' Rationale is simply to encourage independent undergraduate/graduate research in economics, its subfields, and social science to as great a degree as possible. Whilst the Society originally had intentions for an official journal, we have opted for a working paper series to overcome the peer-review process, which we felt could act as a 'barrier to entry' for other students.

Having assembled as four teams in October 2018, and working tirelessly this past academic year to pursue independent research, each team has been able to develop papers in the fields of health economics, urban economics, development economics, and political economy[1]. With research crossing four continents, and being related to highly topical policy and global issues, the entire process has been unequivocally stressful, albeit incredibly rewarding all the same. So on that note, a huge amount of gratitude is due.

Firstly, I personally would like to thank all of the respective team leaders – Anmol, Swapnil, Savannah and David (Yihan) – for their persistent efforts this year in organising their teams and for consistently liaising with myself. What has probably amounted to dozens of meetings in either the Garrick or Bean Counter in total, I am hugely grateful for their cooperation and hard work, despite all of the other pressing commitments undergraduate and graduate life at the LSE entails.

Secondly, I would also like to thank all of the research associates who would not have made this project possible in the first place. From manually scraping census data from the Philippine Statistical Authority and NASA satellites until 2am in the library, to running regressions and drafting everything in LaTeX till sunrise, there is no doubt a countless number of sleepless nights has made this working paper series achievable, and so for that I am incredibly grateful to the entire Research Division for their dedication.

Thirdly, on behalf of the team leaders and the entire Research Division, I would like to thank any and all professors and faculty who have kindly accommodated us in office hours for our projects, which includes, but is by no means limited to, Professors Sir Tim Besley, Simon Hix, Marcia Schafgans, Robin Burgess, Steve Pischke, Steve Gibbons, and Elisabetta De Cao, in addition to Matteo Sandi. I would also like to thank Judith Shapiro for her advice and support throughout the year.

And finally, I would like to extend my personal thanks to the 2018/19 Economics Society executive committee, who have been incredibly supportive and encouraging of the Research Division's efforts, and who have undoubtedly become dear friends in the process – Arsalan Kamal, Yootha

[1]For any feedback you may have on any of the papers, please find the emails of all coauthors at the back of the publication, and we are more than grateful for any suggestions and comments that can be provided.

Yong, Chris Shaw, Elle Chen, Carl Harper, Samir Patel, Aaron Luke, and Tom Glinnan.

With the re-inauguration of Rationale, we hope in future to extend our scope not just to the Society's Research Division, but to all students at the LSE, other UK institutions, and even other universities around the globe. Whilst the working papers in this edition are purely internal, I hope future Research Divisions will be able to keep this project alive, so as to maintain the LSE's mission of making the world a better place by virtue of its motto: *rerum cognoscere causas* – to understand the causes of things.

And with that, I would like to wish Kerry Neitzel, the incoming 2019/20 Head of Research, the best of luck for next year, and on behalf of the Research Division, and the Economics Society as a whole, we hope you enjoy this year's Rationale.

Christopher Dann
Head of Research 2018/19
LSE SU Economics Society

Heterogeneous Effects of Unemployment on Mental Health, Depending on Gender and Refugee Status: Evidence from Germany and the United States

Anmol Kanotra[1], Kerry Neitzel[1], Maithili Modi[1], Xuyi Yang[1], and Abi Thomas[1]

[1]The London School of Economics and Political Science, London, UK

Abstract

In this paper we study how the effect of unemployment on mental health differs depending on gender and refugee status using data from Germany and the United States. For the Germany case study, the effect of unemployment on life satisfaction is estimated using a panel data regression model. To capture possible heterogeneous effects depending on refugee status and gender, interaction terms between employment status and refugee status as well as between employment status and gender are included, yielding two separate regression models. For the United States, a similar model is used with an interaction term for gender and unemployment. Our main findings suggest that females tend to be more adversely affected by unemployment. Applying our model to the German data, the derived results suggest the opposite, i.e. that men are more impacted by unemployment in terms of mental health. Moreover, the analysis suggests that refugee's mental health is more severely impacted by unemployment, though this effect statistically insignificant. Consistent with theory and literature, our primary findings indicate that there are significant heterogeneities in the effect of unemployment depending on individual characteristics.

1 Introduction

Mental health has been a long-existing public health concern, but in recent years it started to gain increased attention. During the 2008-2013 recession, there were increasing suicide rates in the UK and other affected countries. It seems plausible that this phenomenon is related with the increases in unemployment, as suicide rates had been declining before the recession (NHS.UK, 2019). Therefore, understanding the dynamics of mental health and unemployment is increasingly becoming a policy concern.

While there is a wide literature documenting the effect of unemployment on mental health, not as many articles investigate how this effect differs depending on various social and demographic factors. Aiming to provide more insight into this dimension of research, this paper addresses two questions.

Firstly, we ask whether the effect of unemployment on mental health is different for refugees. In 2017 alone, there have been over 538,000 asylum seekers who were granted protection by EU countries, and over 600,000 illegal migrants into the EU (Europarl.europa.eu, 2019), making refugees a salient part of the labour force and population of these countries. Many refugees remain unemployed, due to lack of language skills, education or because they have no permission to work. According to the German Federal Ministry of Migration and Refugees, almost 40% of all registered refugees in Germany were out of work in December 2018. Given the pressing nature of the refugee crisis and the increasing significance of refugees in the working age population of some European countries, this group is especially interesting to analyse. The second objective of this paper is to examine whether there exist heterogeneities in the effect of unemployment on mental health depending on gender. Since women and men have historically had very different experiences in the labor market and are exposed to different social contexts than men, it seems intuitively plausible that heterogeneities exist. Previous efforts to study different effects of unemployment on mental health for males and females have been largely limited to localised studies within a relatively short period of time. Due of this, many existing findings are only representative of restricted geographical locations and points in time. In this research paper, we use nationwide data from the U.S. and Germany, which allows for a more representative analysis. Our research, thus, provides a fresh and better grounded perspective for understanding the mental health effects of unemployment.

2 Literature Review

Several articles have been published that examine the interaction between gender and unemployment. Reviewing previous research, Hammarström (1994) comprehensively looked at the relationship between youth unemployment and general health from a gender perspective utilising Finnish cross-sectional data. Her main finding was that unemployment is most likely causally linked with health deteriorating behaviour, such as increased alcohol consumption, tobacco consumption and illicit drugs use, the effects being stronger for males. Other detrimental consequences of unemployment that are relevant to mental health are a higher suicide rate, increased risk of alienation, lack of financial resources, criminality and future exclusion from the labour market.

Agquist and Starrin (2007) also studied the mental health effects of unemployment for young people, using a sample of young unemployed and trainees. Based on questionnaire data, capturing individual's retrospective accounts of the transition from employment to unemployment or vice versa as well as individual's mental health condition, they find an explicit relationship between unemployment and mental health issues. According to the study, one out of four men and one out of two women perceived that their mental health worsened upon becoming unemployed. This result further suggests that mental health effects of unemployment have a gender dimension. The direction of the effect, however, seems to go in a different direction than Hammarström's work suggests.

More recently, Strandh et al. (2012) investigated how the relationship between unemployment and mental health depends on gender. Their argument is that men and women experience unemployment differently since they aren't exposed to the same social context. Using representative longitudinal data from Sweden and Ireland, the authors look at both differences in mental health among unemployed men and women as well as changes in mental health when exiting unemployment. Based on their findings they conclude that, indeed, unemployment is experienced differently by women in Ireland, where women are less impacted by unemployment. In Sweden, however, there are no apparent differences. Therefore, the authors argue, the relationship between mental health and unemployment is largely context dependent.

A study that had a narrower scope than previous research discussed utilised a mail based survey to examine the effect of job loss for a sample of individuals in Brevard County, Florida, who were put out of work after the Space Shuttle Challenger disaster in 1986

(Leana 1991). Looking at the results from a gender perspective, no evidence was found that women or men respond differently to job loss in terms of psychological distress symptoms. The evidence did suggest, however, that men and women use different coping mechanisms to deal with unemployment. Men relied more on problem-focused activities such as job search, while women relied more on symptom-focused activities such as seeking social support. Moreover, Marital status was found to have a significant effect on several types of coping behaviours, but it did not significantly interact with gender. Although there are some conflicts in the results found so far, the literature does suggest, overall, that the effect of unemployment on mental health is heterogeneous depending on gender.

While there is a wide literature on interactions between unemployment and gender, there is limited evidence how unemployment affects the mental health of refugees. Pernice and Brook (1996) studied several demographic and post-immigration factors related to self-reported symptoms of mental health disorders of immigrants and refugees in New Zealand. Employing data from a survey administered to a sample of immigrants, including 129 refugees from Southeast Asia, they find that unemployment is one of the main predictors of depression and anxiety disorders among refugees.

By conducting a systematic literature review, Bogic Njoku and Priebe (2015) find that out of all articles considered, 84% suggest a positive association between unemployment and mental health disorders among refugees. In evaluation, the literature does suggest that refugees who are unemployed are more affected by a variety of mental health issues. It should be pointed out, however, that none of the studies we found used a comparison group of natives, to identify whether refugees are more affected than other residents in their host country. Providing this comparison between refugees and other residents is precisely the aim of this paper.

3 Theoretical Framework

To provide a theoretical foundation for the relationship between unemployment and mental health, utility theory is a useful framework. Utility can be seen in two theoretically distinct ways; decision utility is inferred from observed choices and material preferences, whereas experience utility more closely matches the notion of happiness or enjoying, and is a subjective state (Carter & McBride, 2011). Experience utility, and the notion of subjective wellbeing, are arguably more relevant than decision utility when looking at mental health outcomes. Van der Meer (2014) identifies different channels through which employment contributes to subjective wellbeing, which is dependent on physical and social well-being, not just on income. These include receiving stimulation and security from having a job, enjoying social status as an employed person, and receiving an income, which contributes to enjoyment of leisure time through consumption of goods and services. We can therefore see the ways in which unemployment can negatively impact subjective wellbeing. Not only is income lowered, but also one's social status is changed, and stimulation and security are decreased. These latter elements cannot be compensated for by unemployment benefits, which, in a decision utility framework, would be enough to restore utility by merely restoring income to close to previous levels. Thus, unemployment has a severe impact on subjective wellbeing (Cole et al. 2009; Creed and Macintyre, 2001), despite the fact that traditional economic theory considers 'work' to be an inferior good which provides disutility to workers.

This negative impact of unemployment on subjective wellbeing is not shared to the same extent by other individuals who are out of the labour market, such as homemakers and pensioners (Easterlin, 2005). This can play a part in explaining why differences in wellbeing can vary among unemployed men and women, with research finding that men are affected more severely by unemployment than women (Hultman et al. 2006). If people who become unemployed have a different way by which they can receive social approval, stimulation and security, and other things which augment wellbeing, then their happiness may not be negatively impacted so much. For many women, traditionally more so than for men, this may be through being a homemaker. This is reinforced by traditional, and often outdated, norms, which see it as more socially acceptable for women to stay at home while men are the main breadwinners within a household (Stam et al., 2015). Therefore, one might expect women to be less affected than men by unemployment in terms of mental health. Ultimately, however, the direction of the effect depends much on the specific social context which defines the channels through which men and women gain their wellbeing.

Another important determinant of experience utility is reference dependence. Reference-utility theory postulates that there are differences in the way that individuals value their subjective wellbeing depending on their individual starting points, or reference points, and whether outcomes are better or worse relative to this (Tversky & Kahneman (1992); Koszegi & Rabin, 2006). For example, it is found that relative income is more important to people's subjective wellbeing than absolute income, where relative income refers not only to income levels relative to those of others but also with respect to one's own past experiences and future expectations (Castilla, 2012). Carter and McBride (2011) find that people have an 'S-shaped' satisfaction, or utility function, in which the reference point depends on past expectations, social comparisons and subjective expectations. Aspirations also play an important role in this, with Easterlin (2001) finding that wellbeing is a function of the gap between what a person has and what they aspire to have, with those who achieve their aspirations considering themselves better off. Therefore, it seems plausible that unemployment has such a negative impact on mental health because of the sudden fall in subjective wellbeing relative to the reference point of being employed. Thus, the impact of unemployment on wellbeing may vary depending on how high this reference point was initially. It may be the case that for certain groups, such as refugees, whose reference point may be based on low levels of wellbeing experienced in the country from which they were forced to flee, that unemployment within their new country may not provide such a relative drop in wellbeing, such that they are still relatively content. Hence, it could be hypothesised that refugee's mental health is less negatively impacted by unemployment.

Not only does our wellbeing depend on our reference point, but it is found that reference points shift over time; our aspirations rise as previous ones are satisfied (Easterlin, 2005). This idea is known as the Hedonic Treadmill (or hedonic adaptation/habituation) (Brickman & Campbell, 1971) and explains how people adapt to improving circumstances to the point of affective neutrality. Evidence for this includes findings that lottery winners are generally no happier, and long-term paraplegics no less happy, than before these big changes occurred in their lives (Brickman et al., 1978). Kahneman and Kruger (2006) explain that many seemingly important life events, such as marriage and bereavement, have substantial short-run effects on life satisfaction, but that these effects are generally temporary. A similar impact is found for economic changes, such as large increases in income or standard of living. For example, Easterlin (1995) finds that the average self-reported happiness level did not increase in Japan between 1958 and 1987, although real income increased fivefold. Therefore, it may be interesting to see how the impacts of unemployment on different groups are felt as people's reference point shifts over time.

4 The German Case

4.1 Data and Variables

The empirical analysis is based on the German Socio-Economic Panel, which contains demographic data of roughly 20.000 residents each year, compiled during a cohort study running in Germany since 1984. In this paper, the data used ranges from 1984 to 2016. The main subpopulation of interest, for which data is not available in the U.S. data set, is residents who have been granted official refugee status. The total number of refugees who were interviewed is 2468, where all these responses were recorded in the year 2016. This temporal clustering of the refugee sample is unsurprising, as starting in 2015 large numbers of migrants from Syria and sub-Saharan Africa began to make their way across the Mediterranean Sea and the Balkans to seek asylum in Europe, resulting in what is commonly known as the European "refugee crisis".

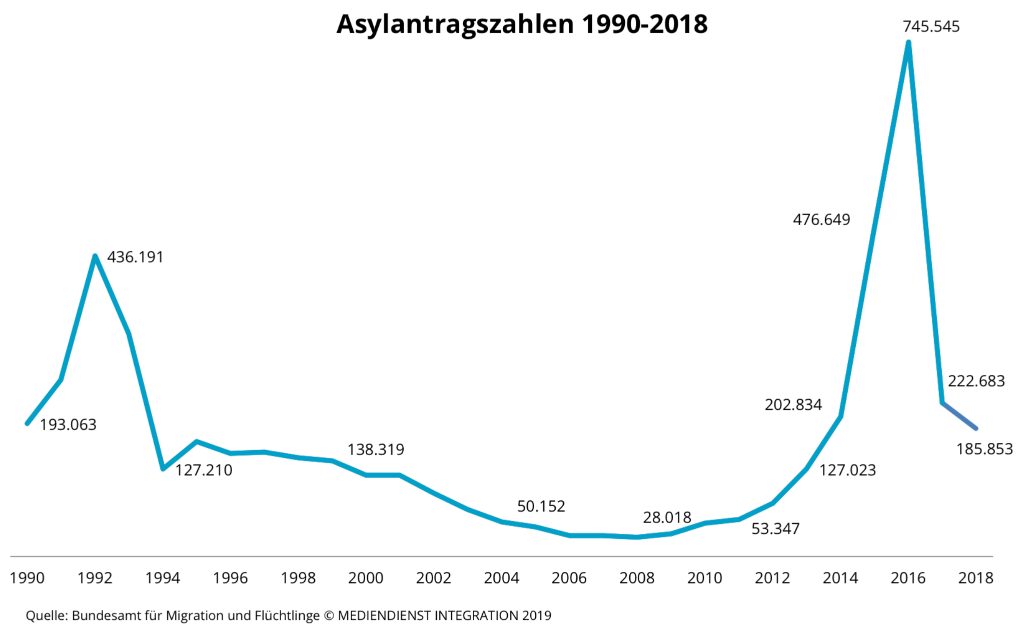


Figure 1: Asylantragszahlen 1990-2018: Number of Asylum Applications per Year

This phenomenon can be strongly seen from the graph above, showing the total number of applications for asylum in Germany since 1990. While the number is quite moderate for most of the period, with the refugee crisis in 2016 the number spikes significantly.

The data set doesn't contain explicit information about individuals' mental health. Instead, overall life satisfaction is used as a proxy. On the survey, individuals are asked the following question: "How satisfied are you at present with your life as a whole?". Respondents are then presented an ordinal scale ranging from 0 to 10, where 0 means "completely dissatisfied" and 10 means "completely satisfied". This measurement of psychological well-being is purely subjective, which leads to problems. Most importantly, interpersonal comparison may not be valid, because people "anchor" their scale at different points. Winkelmann, Liliana, and Rainer Winkelmann (1998) solve this problem by interpreting "anchoring" of the scale as idiosyncratic individual effects. These effects can then be considered using individual fixed effects in a regression model. In this case, however, it is not possible to use individual fixed effects, as two of the main treatment variables (gender and refugee status) are immutable characteristics and would, hence, be absorbed by the fixed effects. We will get back to this problem in the next section. The first main treatment variable used is employment status (Unemployed). For the purpose of this analysis, a binary variable is used, which takes the value 1 if a person is full-time or part-time employed and is coded 0 if a person is unemployed or has irregular or marginal employment. The other main treatment variable is refugee status (Refstat), which is a binary variable taking the value 1 if a person has been officially granted refugee status and is 0 otherwise. Gender is captured by a dummy variable that takes the value 1 for females (Female). Beyond these three variables, the other explanatory variables are age (Age), marital status (Married), years of education (Education) and the number of total months a person has been unemployed before (UnempExp).

4.2 Methodology

The effect of unemployment on life satisfaction is estimated using a panel data regression model. In order to capture possible heterogeneous effects depending on refugee status and gender, interaction terms between employment status and refugee status as well as between employment status and gender (Unempref and Femunemployed) are included, yielding two separate regression models. Moreover, time fixed effects capture possible temporal trends.

As mentioned above, there is an inherent problem with comparing life satisfaction scores across individuals, due to the subjective nature of the measurement. Since individual fixed effects are not viable in this case, the alternative strategy used here is to run two regressions. The first one estimates the effect of employment status on life satisfaction, with fixed effects. The second one does the same, but this time without individual fixed effects. If there is a bias due to anchoring of the scale, this should cause a difference in the effect of unemployment on life satisfaction. Hence, if the difference is negligible, this gives reason to believe that the estimates without fixed effects are trustworthy. Another important potential concern with the models is that there could be reverse causality, i.e. overall life satisfaction has a causal impact on unemployment status. We consider some evidence to assess whether this is the case. Our reasoning is that if the unemployed were inherently dissatisfied, then it should be the case that the unemployed are just as satisfied with their life as those who are employed but have been unemployed in the past. Therefore, we draw a comparison in mean life satisfaction between these different groups in the next section.

After having addressed these two issues of internal validity, we run the two main models:

Model with interaction term for gender:

$$\begin{aligned} LifeSatisfaction_{it} = {} & \beta_0 + \beta_1 \cdot Unemployed_{it} + \beta_2 \cdot Female_i + \beta_3 \cdot Femunemployed_{it} \\ & + \beta_4 \cdot Age_{it} + \beta_5 \cdot Education_{it} + \beta_6 \cdot Married_{it} + Year_t + \varepsilon_{it} \end{aligned} \tag{1}$$

Model with interaction term for refugee status:

$$\begin{aligned} LifeSatisfaction_{it} = {} & \beta_0 + \beta_1 \cdot Unemployed_{it} + \beta_2 \cdot Refstat_i + \beta_3 \cdot Unemprefstat_{it} \\ & + \beta_4 \cdot Age_{it} + \beta_5 \cdot Education_{it} + \beta_6 \cdot Married_{it} + Year_t + \varepsilon_{it} \end{aligned} \tag{2}$$

It should be noted that in the model with the interaction term for refugee status, it is not possible to control for gender, since all refugees in the sample are male. Furthermore, the refugee-model includes previous unemployment experience as a control, as this could be a factor determining individuals choice to leave their country and seek asylum abroad.

5 The U.S. Case

5.1 Data and Variables

The data used has been sourced from "The Behavioural Risk Factor Surveillance System (BRFSS)" which is USA's premier system of health-related telephone surveys that collect state-level data about 18,000 U.S. residents regarding their health-related risk behaviours, chronic health conditions, and use of preventive services. While the dataset includes survey data from 1984 to 2018, we have only used data 1994 onwards as that is from when the mental health indicator was added to the survey.

The key mental health indicator is a self-reported status which is derived from asking the question, "Now thinking about your mental health, which includes stress, depression, and problems with emotions, for how many days during the past 30 days was your mental health not good?" Hence, the variable is discrete, ranging from 0 to 30. If people refused to answer, or answered by saying "I don't know", the observations were coded as missing values. While the variable is self-reported, it has fewer limitations in terms of comparability across individuals as the unit is number of days, as opposed to an arbitrary scale.

The main dependent variables is employment status ("Unemployed"), which is a dummy variable equalling 1 if a person is currently unemployed and zero otherwise. Beyond this, the additional regressors are relationship status ("NotSingle" taking the value one if a person is married or in a non-married couple), age ("Age") and education level ("Education"). Education level is a categorical variable dividing individuals up as follows: never attended school (1), attended elementary school (2), completed some high school (3), graduated from high school (4), completed some college or technical school (5), is a college graduate (6).

5.2 Methodology

To evaluate how mental health outcomes differ between men and women, with respect to unemployment, we use a cross-sectional dataset, across all 50 American states from 1994 to 2018. We then use a linear regression model using mental health outcome as the dependent variable and unemployment and gender as the independent variables/regressors. There are also other demographic factors such as age, education levels and relationship status that can cause a bias in our estimates, and hence have been included these as controls.

Since unemployment can be cyclical in nature, or be impacted by other time-variant factors, we want to avoid this bias in our estimates. Therefore, to get closer to causality, we control for those changes that occur economy-wide in a given time period, by running a time (year) fixed effects regression. Since it is a cross-sectional dataset, running entity fixed effects would not be viable.

We use the following Time Fixed Effects model:

$$\begin{aligned} MentalHealth_{it} = \beta_0 + \beta_1 \cdot Unemployed_{it} + \beta_2 \cdot Age_{it} + \beta_3 \cdot Educa_{it} + \beta 4 \\ \cdot NotSingle_{it} + \beta_5 \cdot Fem_i + \beta 6 \cdot Fem.unemployed_{it} + Year_t + \varepsilon_{it} \end{aligned} \tag{3}$$

6 Results

6.1 Germany

Before considering the main results, it is important to test whether the exclusion of individual fixed effects can be expected to have a biasing impact. As described above, the method used here is comparing two regressions for the effect of unemployment on life satisfaction, with and without individual fixed effects. The result from the first regression (Table 1), with time fixed effects only, is that unemployment has a significant negative effect on life satisfaction, both statistically ($p < 0.01$) and in terms of magnitude (unemployment is associated is associated with a 0.226-point drop in life satisfaction). Including individual fixed effects barely changes this result at all, causing the coefficient to decrease in absolute value by less than 1%. This clearly suggests that any bias caused by Individual differences in "anchoring" is negligible. It is still possible, however, that bias in the estimators for the coefficients of the interaction terms for gender and refugee status is introduced by the fact that members from either respective subpopulation systematically "anchor" their life satisfaction scale differently from the remaining population.

Next, we consider evidence to assess whether there is reverse causality. In Table 2, the mean life satisfaction scores are displayed for three categories of individuals: Employed, Unemployed and those who are currently employed but have been unemployed for some time in the past. What can be seen from the table is that, very clearly, unemployment

Table 1: Fixed Effects Regression Output

	Dependent Variable is Life Satisfaction	
	(1)	(2)
Unemployed	-0.226***	-0.225***
	(0.018)	(0.010)
Female	0.060**	
	(0.027)	
Age	-0.001	-0.024***
	(0.001)	(0.001)
Education	0.018***	-0.011***
	(0.005)	(0.002)
Married	0.320***	0.233***
	(0.024)	(0.014)
Constant	7.151***	8.250***
	(0.071)	(0.035)
Obs.	158868	536111
R-squared	0.0282	0.0064
Individual Fixed Effects	No	Yes
Time Fixed Effects	Yes	Yes

Robust standard errors in parentheses

***p<0.01, **p<0.05, *p<0.1

Table 2: Mean Life Satisfaction

	Mean	Std.Err.	[95% confidence interval]	
Employed (n=311,455)	7.1984	.00297	7.1925	7.2042
Employed with previous unemployment (n=111,276)	6.9901	.00518	6.9799	7.0003
Unemployed (n=264,831)	6.9027	.00385	6.8952	6.9103

is associated with a drop in overall life satisfaction compared with employment (change of ≈ -0.3). Moreover, those who are employed but have been unemployed in the past also have a lower mean score compared with the employed in total (change of ≈ -0.2). This difference, however, is smaller than the drop-in life between the employed and the unemployed. If the unemployed were inherently dissatisfied with their life, then these two differences in life satisfaction should be the same (as noted in the previous section), which suggests that any correlation between unemployment and life satisfaction is not entirely driven by a reverse causal relationship.

The main models both produce the result that unemployment negatively affects life satisfaction, in line with existing literature and theory, the effect being slightly stronger in the gender model. Considering the female subsample (see Table 3), women tend to have slightly lower life satisfaction than men, although this effect is not statistically significant at any standard level once controls are included. The interaction term with unemployment is positive and has a meaningful magnitude, in all models, where the interaction coefficient has a value around 0.2 in models 3 and 4. Moreover, the coefficient is highly statistically significant in all models ($p < 0.01$). This means that, using models 3 and 4 the estimated effect of unemployment for men is roughly -0.33 points on the life satisfaction scale while for women the effect is about -0.13. In other words, the effect of unemployment is about 60% weaker for women than for men. Clearly, this suggests is that for women, becoming unemployed has a less detrimental impact on overall mental health compared to men.

Considering the remaining regressors, age seems to have no significant impact once time fixed effects are included. In contrast, years of education and marital status both have statistically significant positive effects ($p < 0.01$) in all models. The coefficient for age squared is significant at the 10% level but is virtually zero in magnitude, suggesting that the effect of age doesn't change across a person's life span.

Table 3: Effect of Unemployment on Life Satisfaction dependent on Gender

	Dependent Variable is Life Satisfaction			
	(1)	(2)	(3)	(4)
Unemployed	-0.568***	0.339***	-0.333***	-0.322***
	(0.026)	(0.027)	(0.027)	(0.028)
Female	-0.119***	-0.031	-0.038	-0.038
	(0.029)	(0.030)	(0.029)	(0.029)
Femunemployed	0.358***	0.238***	0.205***	0.201***
	(0.034)	(0.035)	(0.035)	(0.035)
Age		-0.020***	-0.001	0.006
		(0.001)	(0.001)	(0.004)
Education		0.014***	0.017***	0.017***
		(0.005)	(0.005)	(0.005)
Married		0.318***	0.313***	0.299***
		(0.024)	(0.024)	(0.025)
Agesq				-0.000*
				(0.000)
Constant	7.258***	7.756***	7.194***	7.064***
	(0.019)	(0.068)	(0.071)	(0.104)
Obs.	164378	158868	158868	158868
R-squared	0.0091	0.0101	0.0292	0.0282
Time Fixed Effects	No	No	Yes	Yes

Robust standard errors in parentheses

**p<0.01, **p<0.05, *p<0.1

For the refugee subpopulation, the results look different. It can be seen from table 4 that refugees tend to have (statistically significantly) higher overall life satisfaction. With an effect of almost at least -0.2 points on the life satisfaction scale in all models, this effect is considerable. The interaction term, in contrast, is always negative with a value around -0.22 in the models 2 and 3. This implies that, for refugees the effect of unemployment is close to -0.36 while for other residents it is -0.14. As such, the results lead to the inference that refugees are much more impacted by unemployment in terms of mental health than other residents. It must be noted, however, that this effect is not statistically significant at any standard level. This is not surprising, given the small number of refugees in the sample and given the temporal clustering of the observations. What this means is that, ultimately, the data don't allow fully conclusive statements about the true interaction effect. Given these empirical results it might be speculated, however, that if there is a non-zero effect, it is more likely to be negative. Similar to before, age has a slight negative effect in the models 1-3. But when age squared is included the effect becomes positive and diminishing (at a very low rate). Years of education and marital status also have a positive effect as before, where this time both effects remain highly statistically significant in all models ($p < 0.01$).

Table 4: The Effect of Unemployment on Life Satisfaction dependent on Refugee Status

	The Dependent Variable is Life Satisfaction			
	(1)	(2)	(3)	(4)
Unemployed	-0.200***	-0.139***	-0.138***	-0.186***
	(0.012)	(0.012)	(0.012)	(0.013)
Refstat	0.340**	0.317**	0.244	0.231
	(0.160)	(0.161)	(0.161)	(0.161)
Unemprefstat	-0.128	-0.218	-0.219	-0.175
	(0.167)	(0.167)	(0.167)	(0.167)
Age		-0.008***	-0.008***	0.034***
		(0.001)	(0.001)	(0.003)
Education		0.002	0.001	0.005***
		(0.002)	(0.002)	(0.002)
Married		0.237***	0.242***	0.291***
		(0.016)	(0.016)	(0.016)
UnempExp		-0.320***	-0.321***	-0.290***
		(0.018)	(0.018)	(0.018)
Agesq				0.000***
				(0.000)
Constant	7.412***	7.665***	8.074***	8.557***
	(0.010)	(0.031)	(0.057)	(0.078)
Obs.	237046	226829	226829	226829
R-squared	0.0041	0.0344	0.0405	0.0480
Time Fixed Effects	No	No	Yes	Yes

Robust standard errors in parentheses

**p<0.01, **p<0.05, *p<0.1

6.2 U.S.

A coefficient equal to 3 persists for the unemployment dummy variable across all models, signifying that if a person is unemployed, the number of days that their mental health is bad are larger. Which implies that unemployment negatively affects mental health outcomes, as expected in theory and proven in the literature.

Women in general have more bad days of mental health than men – on average 1.15 more days (out of 30).

This regression model shows us that women who are unemployed have (on average) 3.08 number of bad mental health days as compared with women who are employed (i.e. $\beta_1 + \beta_6$). Moreover, we see that unemployment affects women (marginally) more than men, as unemployed men, on average, have 3 bad days of mental health. This result is significant at the 5% level, but not at the 1% level.

When we split categorical variables into their individual values, we observe that people below the age of 65 have a greater number of days of bad mental health, whereas those above 65 have fewer. This is interesting as 65 is the retirement age in USA. At this point it is important to note that while those who are unemployed have worse mental health than those who are not, during a lifespan of an individual, he/she is relatively better-off off during the retired years. However, this result is not of significance to our research question. Further research on this would be possible.

Finally, when comparing results of the time fixed effects model with those in the model without time fixed effects, we do not see any major differences. This implies that the model without fixed effects does not entail a lot of bias in its estimates. Moreover, both our models have significant estimates as all p-values are zero or close to zero.

Table 5: The Effect of Unemployment on Mental Health dependent on Gender

	Dependent Variable is Mental Health	
	(1)	(2)
Unemployed	3.041***	3.011***
	(0.029)	(0.029)
Age	-0.032***	-0.034***
	(0.000)	(0.000)
Education	-0.388***	-0.397***
	(0.004)	(0.004)
Not Single	-1.051***	-1.122***
	(0.008)	(0.008)
Fem	1.177***	1.153***
	(0.007)	(0.007)
FemUnemployed	0.059	0.077*
	(0.038)	(0.038)
Constant	6.228***	2.240***
	(0.023)	(0.033)
Obs.	3802369	3802369
R-squared	0.038	0.039
Time Fixed Effects	No	Yes

Robust standard errors in parentheses

**p<0.01, **p<0.05, *p<0.1

7 Discussion

To conclude, our first finding, based on the German data, is that the effect of unemployment on mental health is more negative for men than for women. This can be readily explained using our initial theoretical framework, which is mainly based on experience utility. Applying the theory, it could be argued that women may be able to derive utility from non-pecuniary channels to a greater extent than men. For example, unemployed women may still receive mental stimulation from looking after their family and receive social status as a homemaker (Stam et al., 2016).

Using the U.S. data, however, we find that women are marginally more affected by unemployment than men. There are several possible explanations for the difference in results from the two countries. One possibility is that the difference in our results between Germany and U.S. could be driven by the types of data used for these two countries. The German results are drawn from panel data, containing information of the same individuals over time. In contrast, the U.S. results are drawn from cross-sectional data, which means individuals are not tracked over time. Since our mental health measures are self-reported, looking at the same individuals over time is important to reduce bias, suggesting the German results are more robust. Another difference is that the interaction coefficient for gender has very low standard errors using the German data compared to the U.S. data, further suggesting that the former set of results is more reliable. Finally, it cannot be ruled out that the results reflect differing effects of unemployment on mental health for women in the U.S. and Germany. Most plausibly, differences in the structure of the labor market and social context cause women to experience unemployment differently across countries, which is also suggested by studies done by Strandh et al. (2012).

The second main finding is that, although refugees on average have a higher life satisfaction overall, there seems to be a stronger negative impact of unemployment on life satisfaction for refugees than for other residents though not significant. This may seem surprising, because theories of reference dependence seem to suggest that for groups like refugees, unemployment in their host country should result in a lower relative drop in well-being compared to other residents, because their reference point is lower. The reference dependence theory, however, is supported by the general finding that refugees are more satisfied with life on average than non-refugees. To account for the discovered interaction effect, it could be speculated that unemployment comes with a greater social stigma for refugees.

Moreover, unrealistic expectations about their opportunities in the host countries could lead to an inflated reference point. Finally, it is also plausible that refugees are more impacted by unemployment because tend to have weaker social and family networks who are available to support them in case of unemployment, compared to natives.

It must be kept in mind that our results on refugees, although meaningful in terms of the magnitude of the effect, are only indicative since they are not statistically significant, due to small sample sizes of refugees and temporal clustering of the observations. Given this limitation of our research, further work would benefit from larger sample sizes to see if significant results can be derived. Moreover, research could benefit from having refugee data on the U.S., to see whether the results vary and to do a proper cross-country analysis between the U.S. and Germany. This way, it can be explored how different systems for integrating refugees into the labor market and society influence the effect of unemployment on mental health for this group. There also is scope for future research to test a version of the Hedonic Treadmill effect regarding refugees, to see whether their self-reported mental health goes down over time in response to a shifting reference point of being in a new country with higher standards of living. Finally, something that our research has not addressed is what exactly is driving the effect of unemployment on different groups, although some potential explanations have been discussed above. Therefore, future work should break down exactly what is driving the stronger negative effect on the mental well-being of refugees.

References

Agquist, C., and B. Starrin. "Youth Unemployment and Mental Health – Gender Differences and Economic Stress." Scandinavian Journal of Social Welfare, John Wiley & Sons, Ltd (10.1111), 3 Apr. 2007,

Bogic, M., Njoku, A., & Priebe, S. (2015). Long-term mental health of war-refugees: a systematic literature review. BMC international health and human rights, 15(1), 29.

Brickman, P., & Campbell, D. (1971). Hedonic relativism and planning the good society. In M. H. Apley (Ed.), Adaptation-level theory: A symposium (pp. 287–302). New York: Academic Press.

Brickman, P. et al. (1978) "Lottery Winners and Accident Victims: Is Happiness Relative?" Journal of Personality and Social Psychology. 36(8), 917–927

Cole, K., Daly, A., & Mak, A. (2009). Good for the soul: The relationship between work, wellbeing and psychological capital. Journal of Socio-Economics, 38(3), 464–474.

Carter, S & Mcbride, M. (2013) Experienced Utility versus Decision Utility: Putting the 'S' in Satisfaction. Journal of Socio-Economics, 42: 13–23.

Carrie R. Leana. Gender Differences in Responses to Unemployment. Journal of Vocational Behaviour 38, 65-77 (1991) .

Castilla, C. (2012). Subjective Well-being and Reference Dependence: Insights from Mexico. Journal of Economic Inequality.10:219–238.

Creed, P.A. & Macintyre, S. (2001). The relative effects of deprivation of the latent and manifest benefits of employment on the well-being of unemployed people. Journal of Occupational Health Psychology 6(4): 324–31.

Easterlin, R, A.(1995). "Will Raising the Income of all Increase the Happiness of All?" Journal of Economic Behavior and Organization. 27:1, pp. 35– 47.

Easterlin, R, A. (2001) "Income and Happiness: Towards a Unified Theory," The Economic Journal, 111:473, 465-484.

Easterlin, R. A. (2005) —A Puzzle for Adaptive Theory. Journal of Economic Behavior and Organization, 56 (4), 513-521

Gunnell, D., Donovan, J., Barnes, M., Davies, R., Hawton, K., Kapur, N., Hollingworth, W. and Metcalfe, C. (2015)"The 2008 Global Financial Crisis: effects on mental health and suicide." Policy Report 3/2015

Hammarström, Anne. "Health Consequences of Youth Unemployment—Review from a Gender Perspective." Social Science & Medicine, vol. 38, no. 5, 1994, pp. 699–709.

Hultman, B., Hemlin, S., & Hornquist, J. O. (2006). Quality of life among unemployed and employed people in northern Sweden. Are there any differences? Work: A Journal of Prevention, Assessment and Rehabilitation. 26(1), 47-56.

Kahneman, D. & Krueger, A. (2006). "Developments in the Measurement of Subjective Well-Being." The Journal of Economic Perspectives. 20(1), 3–24.

Kőszegi, B. & Rabin, M. (2006). A Model of Reference-Dependent Preferences. The Quarterly Journal of Economics. 121(4), 1133–1165,

Pernice, R., & Brook, J. (1996). Refugees' and immigrants' mental health: Association of demographic and post-immigration factors. The Journal of social psychology, 136(4), 511-519.

Strandh, Mattias, et al. "Unemployment, Gender and Mental Health: the Role of the Gender Regime." Sociology of Health & Illness, John Wiley & Sons, Ltd (10.1111), 26 Sept. 2012,

Stam, K., et al (2016) Employment Status and Subjective Well-Being: the Role of the Social Norm to Work. Work, Employment & Society. 30(2): 309–333.

Tversky, A. & Kahneman, D. (1992) "Advances in Prospect Theory: Cumulative Representation of Uncertainty." Journal of Risk and Uncertainty. 5(4), 297–323.

Van der Meer, P. (2014). Gender, Unemployment and Subjective Well-Being: Why Being Unemployed Is Worse for Men than for Women. Social Indicators Research. 115 (1): 23–44

Winkelmann, L. & Winkelmann, R. (1998) "Why are the unemployed so unhappy? Evidence from panel data." Economica 65.257: 1-15.

The Hedonic House Price Model in a Development Context – Evidence from the National Capital Region of the Philippines, 2000 - 2010

Christopher Dann[1], Jakub Grohmann[1], Amy Ho[1], Sherman Khoo[1], Anna Kim[1], and Swapnil Lal[2]

[1]The London School of Economics and Political Science, London, UK
[2]University of Toronto, Toronto, Canada

Abstract

The paper analyses the housing market of the National Capital Region (NCR) of the Philippines, alongside contiguous provinces, using the hedonic house price model. Although we do not have precise house price data, we use rental values as a proxy for housing values, and look at a plethora of internal and external characteristics of houses, from the types of toilet facilities households use, to air quality data, as measured by nitrogen dioxide levels from satellites. Using the Philippine Statistical Authority's (PSA) 'Census for Population and Housing', we manually constructed a panel dataset for the years 2000 and 2010, and then use a range of spatial autoregressive models to empirically test the data, whilst simultaneously accounting for spatial autocorrelation. We then use panel data specifications, such as spatial autoregressive random effects regressions, of which we continue to obtain some robust and statistically significant results. Overall, we confirm that fuel for lighting, toilet facilities, and population density are significant determinants of monthly rental values in and around the cities of the NCR. Nonetheless, due to innumerable endogeneity issues, in addition to the Modifiable Areal Unit Problem (MAUP), we treat our findings as being purely suggestive, in hopes to introduce a future literature on housing inequality in the Philippines.

1 Introduction

Sen's (1999) approach to assessing the quality of life is underpinned by the notions of 'functionings', (states of 'being and doing', such as being well-nourished and having shelter) and 'capabilities' (the set of valuable functionings that a person has effective access to). In the capacity of shelter, it is apparent as to why housing is a fundamental human necessity, and thus a 'functioning' of paramount importance, in order for individuals to lead fulfilled lives. Albeit an emerging economy, the Philippines is still very much a developing country, with 21.6% of the country still living under the national poverty line in 2015 (Asian Development Bank, 2019). With regards to housing, the World Bank estimates that the number of informal settlers in the Philippines has increased gradually, from 4.1% of the total urban population in 2003 to 5.4% in 2012. In the National Capital Region (NCR) alone, an estimated 1.3 million people, or close to 11% of the population in the region, lived in informal settlements, which evidently highlights the nature of housing inequality in the Philippines. As such, we attempt to analyse the housing market of the Philippines' central business district (CBD) of the National Capital Region (NCR), in addition to the cities of neighbouring provinces.

Using census and satellite data, we apply the hedonic house price model to 108 cities, using city-level expected monthly rental values as a proxy for house prices, of which 16 are directly within the NCR province/region. Yet, despite its applications in many developed contexts, the hedonic house price model has not been sufficiently applied to many developing economies; we consequently seek to contribute to the literature in this manner. Firstly, we are able to use Brook's (n.d.) method to develop a rental Gini coefficient for every city within the NCR, and then use this to create two Lorenz curves for 2000 and 2010. Here, we show that while this measure of housing inequality is not stark, inequality has remained inertial over a decade, and chloropleth maps of individual city-level Gini coefficients visually illustrate which cities are most unequal in the NCR. Moreover, when running various spatial diagnostic tests, such as both global and local Moran's I tests and Anselin's (1996) Lagrangian multiplier tests, we are able to confirm the nature of clustering of rental values. These results evidence how 1) the NCR and surrounding provinces conform to Park and

Burgess' (1925) 'concentric zone model' of cities and Alonso's (1964) 'bid-rent curve', and 2) an urban-suburban divide between clusters of very expensive and very cheap housing persists. Finally, using a range of spatial econometric specifications, in addition to fixed and random effects models, we find that fuel for lighting, toilet facilities, and population density are statistically significant and robust determinants of monthly rental values for the cities within and surrounding the NCR.

2 Background

2.1 The Hedonic House Price Model

Property prices are dependent on various characteristics and are not easy to accurately determine. Pioneered by Court (1939), the hedonic house price model splits the price of a good into utilities derived from its individual attributes. This approach allows the relative contribution(s) made to the composite price to be estimated. Hedonic models have been applied is such various settings, such as art auctions, the US apple industry, and the UK car market (Ashenfelter and Graddy, 2003; Tronstad, Huthoefer and Monke, 1992; Requena-Silvente and Walker 2006). When applied to the property market, property prices can be considered as a function of its characteristics, usually split into structural, neighbourhood, and accessibility attributes (Xiao, 2016). The main disadvantage of hedonic models is usually the limited number of explanatory variables, which means the model does not fully represent the posited relationship between prices and valued characteristics. Moreover, as argued by Cotteleer, Gardebroek and Luijt (2008), standard hedonic models do not take into account the influence of market power, assuming perfect competition in explaining transaction prices. In markets with restricted entry, arguably the housing market, there remains an excess surplus divided between buyers and sellers, resulting in a deviation of the observed price from the hedonic price, determined by relative market power. Finally, as the appropriate functional form for a hedonic price equation cannot in general be specified on theoretical grounds, selection of the functional form presents another difficulty, with the choice usually determined by a question of convenience (Halvorsen and Pollakowski, 1981).

Although not prolific, there are some papers that apply the hedonic model in other developing/emerging countries. Ozus et al. (2007) examine the spatial distribution of housing prices at the metropolitan and district level of Istanbul. After conducting an analysis separately for five districts of Istanbul, they conclude that at the sub-market level, the most important variables change from district to district (Ozus et al., 2007). For example, 'seaview' becomes the most important variable in locations on the Bosphorus shores, whilst possession of a fireplace and kitchen are significant in sub-districts with higher socio-economic characteristics (Ozus et al., 2007). On the other hand, central heating and heat insulation are significant in sub-districts with lower socio-economic characteristics (Ozus et al., 2007). Moreover, Akbar and Altaf (1995) explore the hedonic price structure for housing in low income localities of Karachi, discovering a significant impact of utility services (water, gas and electricity) on rent; the presence of each increases the rent by more than 10%. Meanwhile, Chen and Jim (2010) examine the implicit amenity-disamenity effect of the urban landscape in Shenzhen by applying the hedonic pricing method. They conclude that urban parks, residential gardens and natural features of the landscape, like the Shenzhen Bay, are reflected in house prices (Chen and Jim, 2010). Visibility and proximity to each of the above characteristics not only increased the market price of an apartment, but also engendered an especially large premium of 17.2% for houses with a view of residential gardens (Chen and Jim, 2010). Given its success in other developing economies, we suspect that the hedonic valuation model will be able to give us significant results regarding price differentials in Philippines. Furthermore, this model also allows us to analyse, at least approximately, the marginal trade-offs that consumers make among characteristics that define a property.

2.2 Rental Policies of the Philippines

While the NCR is our main geographic area of interest, we increase the number of observations by also looking at all cities that comprise the contiguous provinces of Bulacan, Cavite, Laguna and Rizal. This gives us a total of 108 cities with which to study. As we can see from figure 1 below, in every province except Laguna, the proportion of total households renting has risen over a decade. Whilst this increase is by no means precipitous, it still evidences how the growth in the frequency of renting households subsequently means a lower home ownership rate. As we can see in the NCR, just less than one-third of the total populace are renting houses, and so, additional to the aforementioned lacklustre performance of government policy to assist home ownership rates, this further justifies studying the NCR's housing market.

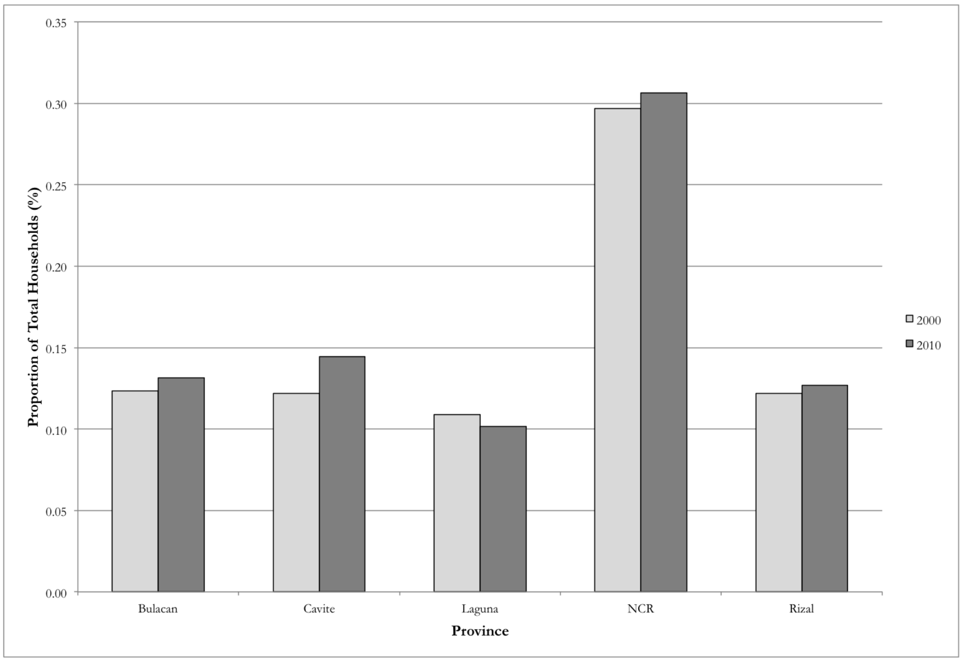


Figure 1: Proportion of Households Renting, 2000 – 2010

In terms of policy, the Urban Development and Housing Act (UDHA) was inaugurated in 1992 with objectives to uplift the conditions of the underprivileged and homeless citizens in urban areas. To achieve its objectives, several strategies were used for land acquisition, balanced housing, private sector participation, consultation and rural development. Strategies for land acquisition included various degrees of government intervention, such as community mortgage, land swapping, land consolidation, land banking, joint venture agreements and expropriation (Ramos, 2000). Moreover, the National Shelter Program (NSP) was one of the actions undertaken, and it is the Philippines' banner program for low-income housing provision. The NSP divides housing into 'socialised' (valued at less than \$6,000, targeted at households up to the 30th income percentile) and 'economic' housing units (valued at up to \$40,000 targeted at households up to the 50th income percentile). Within this framework, a total of 623,053 housing units were built by 2000 and they were mostly serviced with water, power, drainage and roadway systems. Policies related to land opened up possibilities for land ownership to people who would never have had access without government intervention. Incentives like tax deductions or exemptions, in addition to very low interest rates, has also encouraged private sector involvement (Ballesteros, 2010). However, current housing efforts appear inadequate since, on average, the NSP has only delivered 26% of its target.

2.3 Housing Inequality and Home Ownership

Fundamentally, individuals in extreme poverty cannot afford decent housing, which has caused high proportions of informal settlers in many Philippine cities. As mentioned previously, in the NCR alone, an estimated 1.3 million people, or close to 11% of the region's population, live in informal settlements. Various factors beget such housing inequality, of which a major reason is the dearth of housing supply in urban areas. Another factor is the inefficiency of the government's actions. Data obtained from the National Housing Authority showed that in 2011, the government built 4,000 housing units in the province of Bulacan, but only 800 units are currently being occupied (Saludes, 2016). Consequently, we thereby seek to quantify the magnitude of inequality by developing a 'Rental Gini' coefficient, outlining the cumulative proportion of the population that account for cumulative proportions of rent expenditure per city. To calculate this Rental Gini coefficient per city, we use Brooks' (n. d.) basic geometric method, whose fundamental formula is:

$$\text{Score} = \text{Fraction of Income} \times \Big[\text{Fraction of Population} + 2(1 - \text{Fraction of Population})\Big] \quad (1)$$

We use our Census of Population and Housing (CPH) data, for both 2000 and 2010, to calculate such “scores” for all 108 cities. According to Brooks (n. d.), a Gini coefficient for each city j can be calculated simply by subtracting each score by 1, such that:

$$\text{Rental Gini}_j = 1 - \text{Score}_j \tag{2}$$

Using the above method, we were thus able to calculate rental Gini coefficients for all cities within the NCR for both 2000 and 2010, which permits comparisons on housing inequality after a decade has passed. However, we treat the ‘fraction of income’ as being analogous to the ‘fraction of rent expenditure’ for our purposes. In terms of overall housing inequality in the NCR, we were thus able to construct two Lorenz curves.

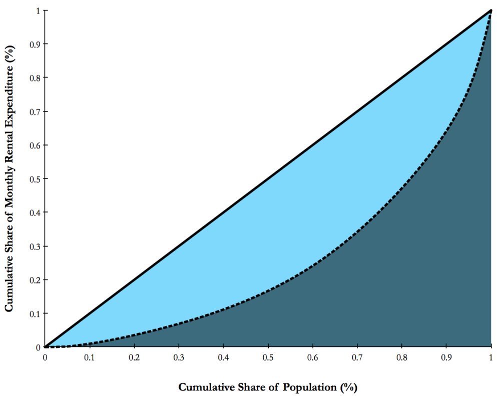


Figure 2: NCR Rent Lorenz Curve, 2000

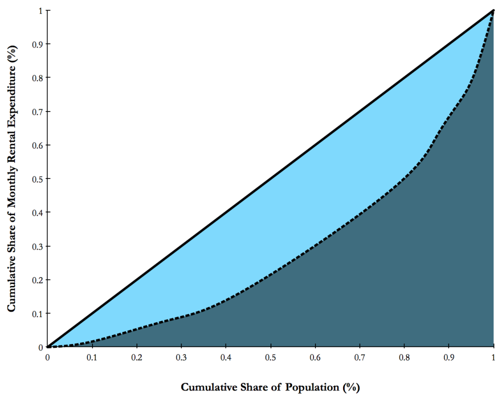


Figure 3: NCR Rent Lorenz Curve, 2010

As we can see between figures 2 and 3, whilst housing inequality has not been overly stark, with an overall NCR rental Gini coefficient of 0.4658 and 0.4209 for 2000 and 2010 respectively, the minor shift inwards towards the line of ‘unitary equality’ highlights the extent of inertia. In the context of an entire decade having passed, this extent is indeed large, and the difference between the 2000 and 2010 rental Gini coefficient is a meagre 0.0449. However, the above Lorenz curves are based on aggregations from each of the individual cities within the NCR. Thus, despite being a further descriptive measure, we try to also visually explore the geographical implications of the NCR Lorenz curves, and thus plot the calculated Gini coefficients for every city within the NCR via chloropleth maps.

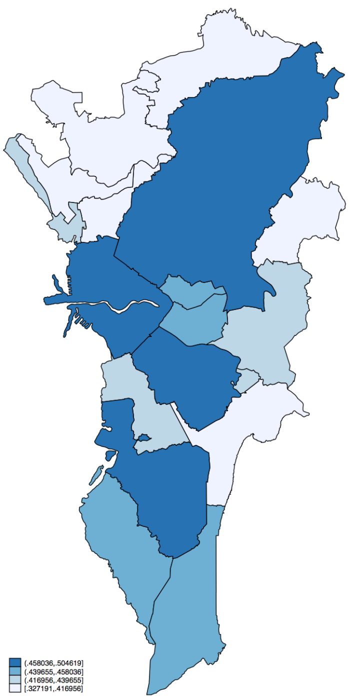


Figure 4: NCR Rent Gini Chloropleth Map, 2000

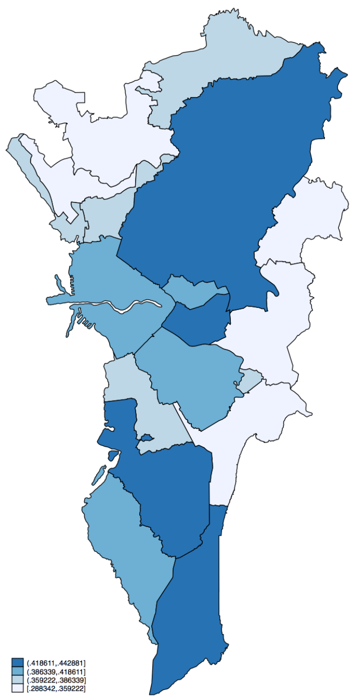


Figure 5: NCR Rent Gini Chloropleth Map, 2010

As we can see from figures 4 and 5, greater levels of housing inequality are positioned in the more central areas of the NCR, whereas the outskirts exhibit smaller Gini coefficients. However, despite a purely descriptive measure, chloropleth maps permit an easy visual analysis and check of the data. As we can see from the above maps, relative Gini coefficients have not changed to a large degree; those areas with higher Gini coefficients in 2000 are good predictors of higher Gini coefficients in 2010. Whilst some areas have indeed become more equal in terms of our Gini measure, other cities within the NCR have actually become more unequal over a decade, such as Mandaluyong (the most central city within the NCR) and Muntinlupa (the southern most city). Additional to the aggregate NCR Lorenz curves, these maps further paint the picture of a housing market that has maintained its levels of inequality, of which events over a decade appear to have not diminished or mitigated. Given the increase in renting as a form of tenancy, versus home ownership, as per figure

1, alongside the inertial levels of housing inequality as pictured above, understanding the determinants of rental values in the NCR region is thus of particular, and real world, relevance.

3 Data

Almost all data was collected from the PSA's 'Census for Population and Housing' (CPH) from 2000 and 2010. Our unit of analysis is each of the seventeen cities within the NCR, and we also look at ninety-one other cities comprising the four provinces that surround the NCR – Bulacan, Cavite, Laguna and Rizal. This allows us to not only see changes within NCR-specific cities, but allows to expand our sample size and research scope by looking at the more suburban areas outside of the NCR. This subsequently permits the potentiality of capturing any urban-rural differences and variations in our dependent variable and associated covariates.

Our dependent variable of interest is the 'expected value of monthly rent' per city, which we calculated via a weighted sum of averages; this was based on the number of census respondents who paid rental values within varying bracket bands. We thus took the average of the upper and lower values within each band, multiplied this by the proportion of individuals who reported paying rent within that specific bracket, and summated this to obtain on overall average for each city. Although our mean measure does not directly capture house prices from transactions, as is commonly performed in the hedonic house price model, we assume, not without caution, that rent values are a good approximation of house prices themselves. We then collected data on relevant variables for our internal and external characteristics. Our internal characteristics consist of:

- Roof/wall material
- Fuel for lighting
- Main source of drinking and/or cooking water
- Type of toilet facility

- Primary manner of garbage disposal
- Washing machine ownership

Our external characteristics consists of:

- Occupied housing units by condition
- Distance to the Central Business District (CBD)
- School location of students
- Population density
- Air quality

Table 1 describes our summary statistics for each dependent, independent, and confounding variable for our two time periods of 2000 and 2010 combined.

Table 1: Summary Statistics, 2000 – 2010

	N	Mean	Std. Dev	Min	Max
log(Average Rental Value)	216	3.1271	0.3106	–	3.7132
Concrete Wall and Iron Roof	216	0.5535	0.1447	0.2016	0.8676
Makeshift Wall and Makeshift Roof	216	0.0044	0.0042	0	0.0236
Electricity for Lighting	216	0.9222	0.0704	0.5609	1.2928
Kerosene for Lighting	216	0.0519	0.0619	0.0034	0.4259
Exclusive Sink as Water Source	216	0.4890	0.2237	0.0106	0.9290
Dug Well as Water Source	216	0.0167	0.0328	0	0.3563
Exclusive Toilet	216	0.6857	0.1510	0.0839	0.9344
Open Pit as Toilet	216	0.0183	0.0307	0	0.2981
Garbage Disposal by Truck	216	0.6610	0.6231	0	8.8504
Garbage Disposal by Dumping in Pit	216	0.0500	0.0489	0.0012	0.4166
Owned Washing Machine	216	0.4412	0.1421	0.0095	3.2178
Unowned Washing Machine	216	0.5321	0.2530	0.0134	0.9905
No Repair/Minor Damage of House	216	0.7328	0.1141	0.0714	0.9092
Dilapidated House	216	0.0091	0.0427	0	0.5703
School in Same City	216	0.8646	0.0673	0.5831	0.9848
School in Other Province	216	0.0525	0.0423	0.0020	0.2369
log(Population Density)	216	3.3064	0.6336	1.1386	4.6004
log(meannightlightpercapita)	216	-3.4202	1.2703	-5.8697	–
Distance to CBD	216	0.3220	0.1606	0	0.6237
NO2 Levels	216	0.2098	0.3089	0	1.3600
Dust and Sea-Salt Particulates	216	9.6779	5.3724	0	20.1900
Average Floor Space	216	47.1121	9.9593	17.7068	75.2014

For all covariates except 'distance to CBD', 'population density' and 'air quality', we simply take the proportion of households who responded as possessing those housing characteristics from the CPHs, and so nearly all variables are values between 0 and 1. Regarding population density, we simply take the natural logarithm of this, and for distance to the CBD, we take the centroid coordinates of each polygon from a shapefile of the Philippines' cities/municipalities (see figure 7 below), and calculate the Euclidean distance from the CBD (which we define as the City of Makati within the NCR). Finally, with regards to air quality, we use two broad measures based on available satellite data. Our measures are 1) nitrogen dioxide (NO2) levels and 2) concentrations (micrograms per cubic meter) of mineral dust and sea-salt filtered fine particulate matter of 2.5 micrometers or smaller. For NO2 levels, this data was sourced from various satellite instruments: the Global Ozone Monitoring Experiment (GOME), SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY (SCIAMACHY) and Global Ozone Monitoring Experiment-2 (GOME-2). For mineral dust and sea-salt, we acquired this data from further various satellite instruments: the NASA Moderate Resolution Imaging Spectroradiometer (MODIS), Multi-angle Imaging SpectroRadiometer (MISR), and the Sea-Viewing Wide Field-of-View Sensor (SeaWiFS). Both datasets were publicly provided by NASA's Socioeconomic Data and Applications Center (SEDAC). By importing both air quality datasets as TIFF files into ArcMap for 2000 and 2010, and by overlaying a shapefile of each city/municipality in the Philippines, we were then able to extract dataset observations for our 108 cities of interest.

We further make use of the census data provided to us by performing a 'quasi-robustness check'. Oftentimes the CPHs would provide different categories of their variables of interest, of which respondents would self-select into. For example, when asked about the usual manner of household garbage disposal, a respondent would be given a range of options: 'Picked up by garbage truck', 'Dumping in individual pit (not burned)', 'Burning', 'Composting', 'Burying', 'Feeding to animals', and 'Others'. Consequently, given this range of categories, we decided to take two versions of our internal characteristic variable – primary manner of garbage disposal – of which one we would expect a positive association with expected monthly rental values, and the other a negative association. With regards to this specific example, we thereby took 'Picked up by garbage truck' as the "good/best" possible version of garbage disposal, thus expecting a positive association with rental values, and took 'Dumping in individual pit (not burned)' as the "bad/worst" possible variant. Although this exercise is based purely on subjective interpretations of what comprises the best/worst versions of our internal and external housing characteristics, our selection of these variants is largely intuitive. In the case of garbage disposal, it is clear that the best version of the usual manner of garbage disposal, based on the categories provided, is being picked up by a dedicated truck, as this would mean 1) more frequent and organised garbage collection, and thus 2) cleaner houses, of which cleaner neighbourhoods, via positive externalities, may be exhibited as well.

In terms of the hedonic house price model, it is clear that these positive benefits surrounding a house should yield higher expected monthly rental values. On the other hand, dumping garbage into an individual pit, without being burned, could create further negative neighbourhood externalities, with the possibility of garbage being blown onto open streets and accumulating, thus bringing the value of houses down. Thus, it should be unanimously accepted that dumping garbage into a pit, relative to having garbage be picked up by a dedicated truck, yields a worst and best version of this internal characteristic of houses. In terms of these best/worst covariates acting as 'quasi-robustness checks', we thus run our specifications twice: once with the good characteristics, and once with the bad characteristics. What we subsequently expect, and as to why this acts as a type of robustness check on our regression specifications, is that coefficients should switch signs: all our regressors should have positive coefficients for the good characteristics, and these coefficients should switch when regressing with the bad characteristics. Table 2 outlines the good/bad variants for our main internal and external characteristics (of course, for the covariates such as air quality and nightlight, there is no "opposite version" of which to test against, and so these variables are used in all specifications for both the good and bad characteristic data panels).

Table 2: Good/Bad Housing Characteristics

Variables	Variants
Roof and Wall Materials	Concrete walls and iron roofs
	Makeshift walls and makeshift roofs
Fuel for Lighting	Electricity
	Kerosene
Primary Source of Water	Own use faucet community water system
for Cooking and Drinking	Dug well
Type of Toilet Facility	Water-sealed sewer septic tank used exclusively by household
	Open pit
Primary Manner of	Picked up by garbage truck
Garbage Disposal	Dumped in individual pit (not burned)
Washing Machine	Owned
Ownership	Not owned
Housing Unit Condition	No repair/minor damage
	Dilapidated
Location of School for	Same city
Students	Other province

Finally, although there are unequivocally a hive of possible confounding variables, data constraints limit the possibility of accessing, and thus incorporating, such variables. As such, the only 'blunt' confounder we employ into our specifications is 'City-level GDP per capita'. To acquire this data, we use the highly popular technique, as popularised by Vernon Henderson et. al. (2012), of using nightlights as a proxy for output. We do this primarily because city-level GDP data is non-existent, as it is too granular an areal unit. Similar to the air quality data, we used ArcMap to import a shapefile of the Philippines' cities/municipalities. We then imported TIFF files of nightlights from the National Oceanic and Atmospheric Administration's (NOAA) DMSP-OLS Nighttime Lights Time Series dataset. This measure is based on a 0-63 scale, whereby each pixel within the TIFF file, as mapped to the city-level shapefile, is given a value between 0 and 63, with the latter being the maximum amount of luminosity. For each city/municipality, we then take the mean level of luminosity, and divide by the city-level populations to acquire per capita values.

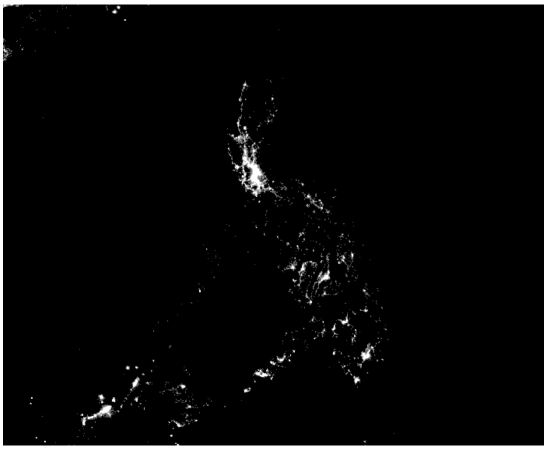


Figure 6: Raw Nightlight Map

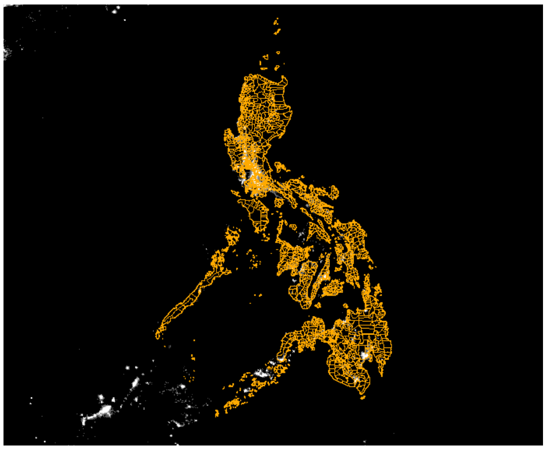


Figure 7: Raw Nightlight Map with City/Municipality Shapefile

4 Methodology

4.1 Spatial Diagnostic Tests

As per LeSage and Pace (2009), spatial dependence refers to the situation whereby observations in some location i concurrently depend on observations in nearby or neighbouring locations j. This is underpinned by Tobler's first law of geography, that "everything is related to everything else, but near things are more related than distant things". The basic empirical model we must first analyse, to assess this notion of spatial dependence with regards to rent values is thus a simple spatial autoregressive (SAR) model:

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\beta + \varepsilon \tag{3}$$

Here, $\mathbf{y}$ is an $n \times 1$ vector of our dependent variable, rent values, for our areal units of cities/municipalities. As such, given there are 108 cities/municipalities, $\mathbf{y}$ shall be 108×1. We first test the possibility of spatial dependence by creating a spatial weights matrix,

$\mathbf{W}$, and use a basic binary matrix, using queen contiguity, to specify all neighbouring constituencies with regards to the spatial proximity of units. The matrices for the two years of data, 2000 and 2010, were created using a shapefile from PhilGIS of all cities/municipalities in the Philippines, and this subsequently gave us a spatial lag on our dependent variable of rent values, $\mathbf{y}$. The spatial weights matrix, $\mathbf{W}$, is simply a square matrix with elements reflecting the degree of connectivity between areal units i and j. We use $\mathbf{W}$ to determine the value of the spatial lag, $\mathbf{Wy}$, which is simply the average values of our dependent variable of rental values, $\mathbf{y}$, for neighbouring cities; it is essentially the average rent values in neighbouring cities, relative to a specific city of interest. By using a spatial weights matrix, we thus create a first-order spatial lag, $\mathbf{Wy}$, on our dependent variable alongside a set of covariates, $\mathbf{X}$, the internal and external characteristics of houses.

In a simple OLS setting, we assume that our disturbance term, ε, is I.I.D, $\varepsilon \sim N(0, \sigma^2)$. However, typically when working with areal units, omitting a test for spatial dependence often leads to a disturbance term which cannot be assumed away as normally distributed I.I.D, thus leading to false inferences about relationships that exist within the data. Consequently, we first must examine whether any such spatial dependence exists within our two-time period panel dataset with spatial diagnostic tests. Referring back to the I.I.D assumption, we can test to see whether ε does exhibit spatial clustering using a Moran's I test, with our null hypothesis being that the disturbance term is I.I.D. For a further diagnostic test, we also use Anselin et.al's (1996) Lagrangian multiplier tests to not only identify whether spatial dependence does exist in the data, but to also outline whether the dependence stems from either clustering on unobservables (the error term drives spatial clustering) or diffusive processes (our dependent variable drives spatial clustering). However, there is a potential third-source of spatial clustering in the data, which is clustering on observables (our covariates drive the spatial clustering). To also test for this, we regress turnout against our covariates and use a Moran's I test again. If spatial clustering disappears when regressing alongside our covariates, then we can show that spatial association stems from clustering on observables; for this to occur, our second Moran's I test statistic would have to become statistically insignificant at any confidence level.

Table 3: Spatial Diagnostic Tests

		Good Housing Characteristics with Controls			Bad Housing Characteristics with Controls		
Census Year	Moran's I Chi-Square Test Statistic	LM_ρ	LM_λ	2nd Moran's I Chi-Square Test Statistic	LM_ρ	LM_λ	2nd Moran's I Chi-Square Test Statistic
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
2000	25.35***	1.7580	0.6676	0.0700	1.4599	2.1954	0.3900
2010	97.83***	4.9969**	0.0442	3.330	1.2371	3.1972	18.7800***

Table 1 gives us a range of results with which to heed attention towards[1]. For the year 2000, it appears none of our spatial diagnostic tests show that any spatial autoregressive model is necessary. Because both Lagrangian multiplier tests failed to reject the null hypotheses, $\rho = 0$ (diffusion) and $\lambda = 0$ (clustering on unobservables), this means we have clustering on observables, which is further confirmed by the fact that the 2nd Moran's I Chi-Square value was statistically insignificant. This subsequently means that the addition of all of our covariates, when regressing with either the good or bad characteristics of houses, accounts for most of the spatial autocorrelation in the data. Hence, we do not require any spatial autoregressive specification to account for this; a basic OLS would be sufficient.

On the other hand, we had much more interesting results for the year 2010. Here, when using the good housing characteristics as our covariates, we once again failed to reject the null hypotheses, $\rho = 0$ and $\lambda = 0$, which leaves us in the exact same position as the 2000 data: we have clustering on observables, so our covariates account for most of the spatial autocorrelation. However, with the bad housing characteristics, while we failed to reject the Lagrangian multiplier tests' null hypotheses, our 2nd Moran's I Chi-Square test statistic was very statistically significant, meaning even once our regressors were included, there was still spatial autocorrelation in the data. Thus, at least for this period of data, for the bad housing characteristics, we must account for spatial autocorrelation by including spatial lags.

[1] Because both the Moran's I and Lagrangian multiplier tests follow a chi-square distribution, our critical value is 3.841 (a chi-squared statistic with one degree of freedom where $\alpha = 0.10$).

4.2 Global Moran's I

Now we have confirmed that spatial association does exist in the data, we use a Moran's I test statistic for each of the census years to measure the intensity of the spatial association, and this acts as a global correlation between the values of an observation with that of its neighbours. The resulting statistics derived are recorded in table 2, and a graphical visualisation of each test is shown in the figures below.

Table 4: Moran's I Global Correlation

Census Year	Moran's I Statistic	Z-Score
2000	0.7310***	11.4600
2010	0.7350***	11.6120

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

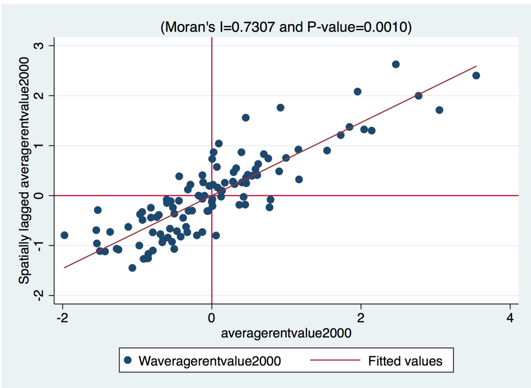


Figure 8: Moran's I Plot, 2000

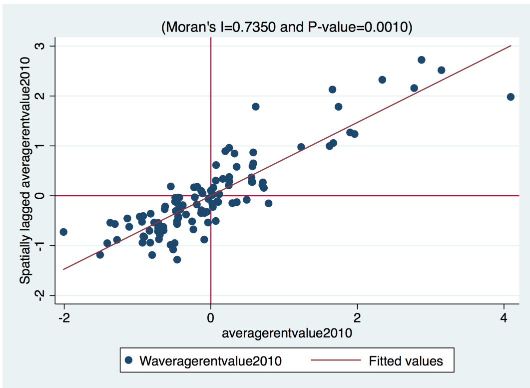


Figure 9: Moran's I Plot, 2010

The Moran's I scatterplot is standardised so that the data has a mean of 0 and standard deviation of 1, and the red trend line is simply an OLS fit through the data points. Fundamentally, the higher the value of the Moran's I test statistic, the greater the level of spatial association. As seen from the above graphs, it is clear that spatial association does exist in the data. Given the Moran's I statistics are 0.7310 and 0.7350 for 2000 and 2010 respectively, with both being statistically significant to the 99% level, we have very strong positive spatial autocorrelation, and thus strong evidence that random chance is not the principal driver of spatial association between cities/municipalities' rent values. To show this descriptively, we develop chloropleth maps of rent values across the five provinces and the cities/municipalities that comprise them.

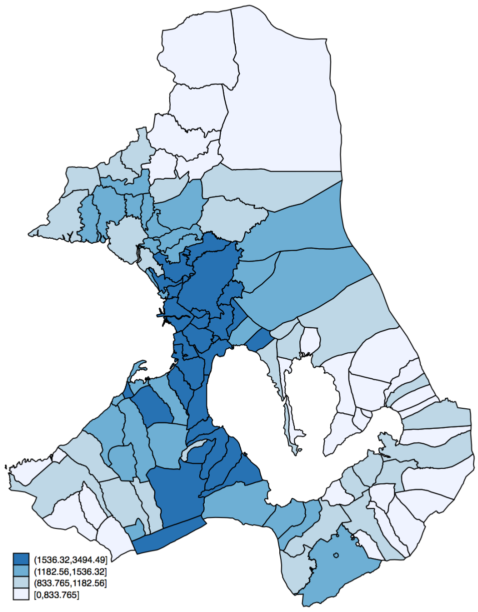


Figure 10: Rent Values Chloropleth Map for Cities/Municipalities of Bulacan, Cavite, Laguna, NCR and Rizal, 2000

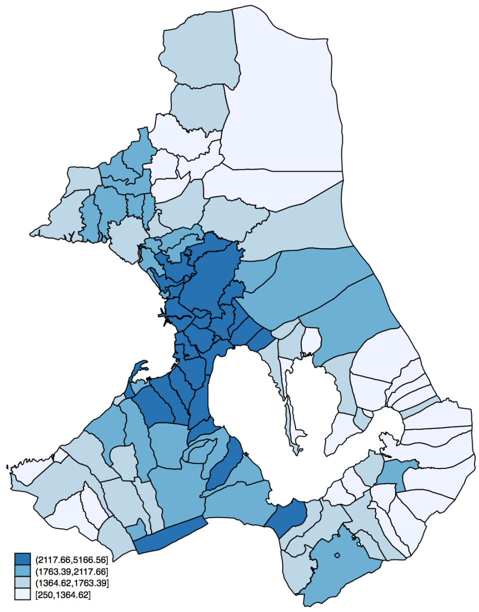


Figure 11: Rent Values Chloropleth Map for Cities/Municipalities of Bulacan, Cavite, Laguna, NCR and Rizal, 2010

As we can see visually from both chloropleth maps, there is definitely strong spatial clustering of rent values. More specifically, both maps evidence two potential theoretical models of interest – Park and Burgess' (1925) 'concentric zone model' and Alonso's (1964) 'bid-rent curve'. As we can see from the maps, the NCR province, being the central business district (CBD) of the Philippines as a whole, has more expensive average monthly rent values. Moving into the outside provinces, we see that monthly rent values become cheaper. With regards to Park and Burgess' (1925) model, while we do not plot, or explore, the impact of the industrial composition of each province, the chloropleth maps open up this possibility for future research; it is clear that proximity to the CBD has a positive impact on rental values.

4.3 Local Moran's I

Although we have successfully shown that spatial autocorrelation exists at the global level, we seek to further identify spatial autocorrelation at the local level. We thus use Anselin's (1996) 'local indicators of spatial association' (LISA), whereby we calculate a local Moran's I for each of the 108 cities. Doing so will enable us to identify either clusters (neighbouring units with positive spatial autocorrelation) or hotspots (spatial outliers).

In a similar fashion to figures 8 and 9, we construct two local Moran's I plots to identify potential clusters and outliers[2].

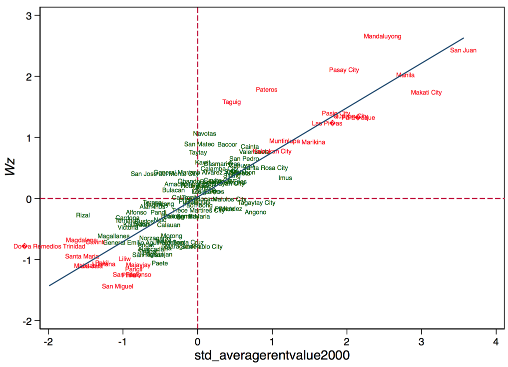


Figure 12: Local Moran's I Plot, 2000

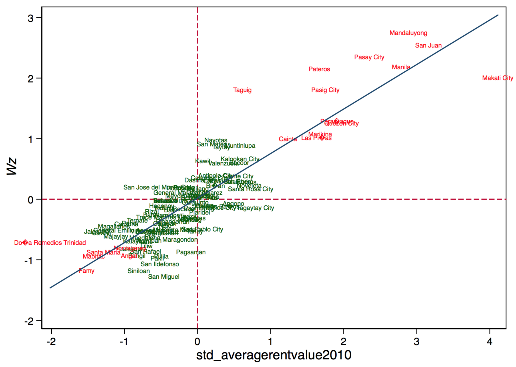


Figure 13: Local Moran's I Plot, 2010

As we can see from figures 12 and 13, there are undoubtedly some cities that cluster together. More specifically, there are three areas of clustering: Bulacan, Laguna, and the NCR. Those cities in the lower lefthand corner of both graphs indicate cities with the lowest monthly rent values that also directly neighbour cities with low rent values;

[2]Cities highlighted in red identify statistically significant z-scores to the 95%, and cities highlighted in green are statistically insignificant.

cities in the upper righthand corner indicate cities with the highest monthly rent values that also directly neighbour cities with high rent values. Based on the aforementioned provinces, from both 2000 and 2010, clusters of cities with low rental values are situated in the provinces of Bulacan and Laguna. On the other hand, again for both 2000 and 2010, the NCR exclusively encompasses a cluster of cities with high rental values. This further evidences the models of Burgess (1925) and Alonso (1964) – the NCR, being the CBD, is not only more expensive, but forms a significant 'high-high' cluster. More suburban cities, such as those in Bulacan and Laguna, comprise a few 'low-low' clusters. This finding could point to further ideas about the nature of housing inequality between the NCR and contiguous provinces, which further studies could explore.

4.4 Cross-Sectional Spatial Autoregressive Models

Now that we have confirmed there is definite spatial association in the data, at both the global and local levels, we can fully justify the use of spatial autoregressive specifications and their applications to the hedonic house price model. Consequently, we look at one-source and two-source models of spatial dependence. One-source models focus on whether spatial association stems from spatial lags (in either our dependent or independent variables) or spatial errors, identified with the spatial error model (SEM), spatial lag of X model (SLX), and the spatial autoregressive model (SAR)[3]. Two-source models focus on the possibility of there being two of either a spatial lag in $\mathbf{X}$ and/or $\mathbf{y}$, and/or spatial errors, identified with the spatial Durbin error model (SDEM), spatial Durbin model (SDM), and spatial autoregressive combined model (SAC) (again, we do not consider either of the spatial Durbin models, as they incorporate spatial lags of our regressors). Table 5 is a summary of all six specifications, of which we use three: the SAR, SEM, and SAC models.

[3]Due to a limited number of observations, and a plethora of independent variables, we decide not to use an SLX model; an SLX model would double the number of regressors by spatially lagging all of our independent variables, and would thus lead to overfitting.

Table 5: Spatial Econometric Model Typology

Name	Structural Model
Spatial Durbin Error Model	$\mathbf{y} = \mathbf{X}\beta + \mathbf{WX}\theta + \mathbf{u}$, $\mathbf{u} = \lambda\mathbf{Wu} + \epsilon$
Spatial Autoregressive Combined Model	$\mathbf{y} = \rho\mathbf{WY} + \mathbf{X}\beta + \mathbf{u}$, $\mathbf{u} = \lambda\mathbf{Wu} + \epsilon$
Spatial Durbin Model	$\mathbf{y} = \rho\mathbf{WY} + \mathbf{X}\beta + \mathbf{WX}\theta + \epsilon$
Spatial Autoregressive Model	$\mathbf{y} = \rho\mathbf{WY} + \mathbf{X}\beta + \epsilon$
Spatial Lag of X Model	$\mathbf{y} = \mathbf{X}\beta + \mathbf{WX}\theta + \epsilon$
Spatial Error Model	$\mathbf{y} = \mathbf{X}\beta + \mathbf{u}$, $\mathbf{u} = \lambda\mathbf{Wu} + \epsilon$

Although our spatial diagnostic tests highlight which specifications to use, we decide to run regressions using all three chosen specifications (in addition to a baseline OLS). If the diagnostic tests are correct, then use of, say, a SAR model for the 2000 good and bad characteristics should yield statistically insignificant coefficients for any spatial lags on our dependent variable. We thus have sixteen total regressions (four for each census year, and for the good and bad housing characteristics).

Table 6: Spatial Regression Results, 2000

	Good Characteristics				Bad Characteristics			
	OLS	SAR	SEM	SAC	OLS	SAR	SEM	SAC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Roof and Wall Materials	-0.0253	-0.0103	-0.0390	-0.0238	-2.4351	-2.5089	-1.9348	-2.1490
	(0.2129)	(0.2037)	(0.1981)	(0.2041)	(3.8386)	(3.550)	(3.5479)	(3.5413)
Fuel for Lighting	0.9047	0.8938	0.9213	0.9095	-0.5696	-0.5405	-0.5526	-0.5276
	(0.6428)	(0.5874)	(0.5938)	(0.5846)	(0.5190)	(0.4881)	(0.4868)	(0.4921)
Primary Water Source	-0.4743	-0.4656	-0.4627	-0.4553	-3.8755**	-3.9667***	-3.8706***	-3.9638***
	(0.3400)	(0.3090)	(0.3204)	(0.3137)	(1.5502)	(1.4651)	(1.4466)	(1.4765)
Toilet Facility	1.1967*	1.1912*	1.1976*	1.1923*	-2.8883**	-2.8738**	-2.9917**	-2.9596**
	(0.6761)	(0.6209)	(0.6388)	(0.6316)	(1.4047)	(1.3104)	(1.3192)	(1.3236)
Garbage Disposal Manner	0.0145	0.0137	0.0159	0.0151	0.3958	0.3867	0.3467	0.3472
	(0.0105)	(0.0095)	(0.0105)	(0.0102)	(0.4471)	(0.4075)	(0.4182)	(0.4102)
Washing Machine	1.1e-8*	1.03e-8**	1.2e-8**	1.1e-8***	2.2e-9	1.3e-9	3.01e-9	2.02e-9
	(5.7e-9)	(4.9e-9)	(5.4e-9)	(5.0e-9)	(3.6e-9)	(3.6e-9)	(3.21e-9)	(3.53e-9)
Building Condition	0.6287***	0.6374***	0.6314***	0.6393***	1.8254*	1.8408*	1.6847*	1.7234*
	(0.1677)	(0.1576)	(0.1553)	(0.1570)	(1.0637)	(0.9856)	(0.9841)	(0.9819)
CBD Distance	0.6775	0.6569	0.6585	0.6416	0.2342	0.2680	0.2731	0.3028
	(0.5524)	(0.4976)	(0.5088)	(0.4953)	(0.3129)	(0.2945)	(0.2982)	(0.3026)
School Location	-0.0937	-0.0854	-0.1015	-0.0935	0.8201*	0.8220*	0.8206*	0.8282*
	(0.2827)	(0.2637)	(0.2623)	(0.2636)	(0.4631)	(0.4229)	(0.4299)	(0.4240)
NO2 Levels	0.0690*	0.0736*	0.0693*	0.0735*	0.0668*	0.0574*	0.0620*	0.0535*
	(0.0402)	(0.0399)	(0.0371)	(0.0393)	(0.0365)	(0.0328)	(0.0330)	(0.0322)
Dust-Sea Salt Particulates	-0.0192	-0.0192	-0.0195	-0.0195	-0.0073	-0.0071	-0.0066	-0.0066
	(0.0130)	(0.0121)	(0.0121)	(0.0121)	(0.0057)	(0.0053)	(0.0053)	(0.0053)
log(Population Density)	0.4636**	0.4626***	0.4602***	0.4598***	0.2129***	0.2142***	0.2122***	0.2145***
	(0.1803)	(0.1660)	(0.1648)	(0.1640)	(0.0677)	(0.0631)	(0.0637)	(0.0640)
Average Floor Space	0.0030	0.0030	0.0026	0.0027	0.0067***	0.0067***	0.0067***	0.0067***
	(0.0025)	(0.0023)	(0.0023)	(0.0023)	(0.0024)	(0.0023)	(0.0023)	(0.0023)
log(meannightlightpercapita)	-0.0450*	-0.0459**	-0.0452**	-0.0459**	-0.0168	-0.0144	-0.0158	-0.0136
	(0.0229)	(0.0219)	(0.0214)	(0.0220)	(0.0252)	(0.0239)	(0.0233)	(0.0239)
ρ		-0.0071		-0.0065		0.0139		0.0141
		(0.0203)		(0.0203)		(0.0169)		(0.0171)
λ			-0.0731	-0.0695			0.1222	0.0935
			(0.1823)	(0.1807)			(0.2571)	(0.2676)
Observations	108	108	108	108	108	108	108	108
R^2	0.6785				0.7809			
Pseudo R^2		0.6789	0.6784	0.6788		0.7816	0.7807	0.7814

Robust standard errors in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

As we can see from table 6, we do have some statistically significant results, of which some also match our hypothetical expectations. For our good housing characteristics, increasing the proportion of houses who have exclusive toilet facilities, washing machines, and building conditions with minor repairs necessary subsequently increases monthly rental values. Although our washing machine coefficient can be considered a negligible effect, this is also intuitively concrete: the presence of a washing machine inside houses is unlikely to be the most significant determiner of monthly rental values. Nonetheless, there were still some surprising results. For example, the fact that our measure of city-level GDP per capita, log(meannightlightpercapita), was statistically significant but negatively associated with monthly rental values is odd, as one would expect the opposite. However, we can argue that this is largely driven by measurement error, in that *raw* values of nightlight are potentially not the most accurate measures of output levels. Moreover, it is surprising that distance to the CBD, albeit being positively correlated, is not statistically significant, given the above evidence from our global and local Moran's I tests. Yet, we attribute this again to measurement error, as we simply calculated the distance to the CBD by using the centroid of each city's polygon from the shapefile. Distance to the CBD may not necessarily be determined entirely through this perfect Euclidean geographical distance, as, in actuality, this will most likely be reflected by transport services, which we have not accounted for due to data limitations.

Across all our specifications, our results are very robust. The fact that our spatial lag coefficients (ρ and λ) for the SAR, SEM, and SAC models are statistically insignificant supports our results from our spatial diagnostic tests; the introduction of a spatially lagged dependent variable and/or error term, to account for any spatial autocorrelation, did not improve the results. Thus, as we expected, a simple OLS was sufficient for this cross-section of data. Perhaps our most robust finding is with regards to population density: regardless of the specification, and whether using the good or bad housing characteristics as our regressors, increasing the population density by 1% significantly increases monthly rental values. However, another robust finding is that of nitrogen dioxide (NO2) levels in the air: the greater the NO2 levels, the greater the monthly rental values. Whilst this is a less intuitive finding, there is the potential for omitted variable bias here. Cities, by design, are likely to be the most densely populated and richer areas, and so consumption of various nitrogen emitting amenities, such as cars, trains, airplanes, amongst others, are likely to contribute towards this result. Further specifications could interact this variable with a binary indicator for city or municipality status (in our dataset, some of our areal units are classified as ‘cities’ whilst others are classified as ‘municipalities’).

Table 7: Spatial Regression Results, 2010

	Good Characteristics				Bad Characteristics			
	OLS	SAR	SEM	SAC	OLS	SAR	SEM	SAC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Roof and Wall Materials	-0.5711***	-0.5752***	-0.5156***	-0.5157***	0.2344	0.4173	0.7849	1.0621
	(0.1111)	(0.1031)	(0.1079)	(0.1078)	(3.4678)	(3.2207)	(3.0519)	(3.0412)
Fuel for Lighting	0.1687	0.1864	0.1448	0.1649	0.1872	0.1747	0.1649	0.1311
	(0.4316)	(0.4055)	(0.3980)	(0.3999)	(0.5911)	(0.5511)	(0.5324)	(0.5263)
Primary Water Source	-0.0080	-0.0101	-0.0244	-0.0275	-0.4813	-0.4841	-0.4439	-0.4459
	(0.0329)	(0.0299)	(0.0342)	(0.0336)	(0.5500)	(0.5104)	(0.5115)	(0.5125)
Toilet Facility	0.4935***	0.4837***	0.4620***	0.4444***	-0.4185	-0.3929	-0.7852	-0.6928
	(0.1432)	(0.1340)	(0.1305)	(0.1319)	(0.6049)	(0.5716)	(0.5450)	(0.5424)
Garbage Disposal Manner	0.0621	0.0656*	0.0717*	0.0762**	0.0333	0.0286	-0.1073	-0.1208
	(0.0407)	(0.0358)	(0.0392)	(0.0372)	(0.2240)	(0.2077)	(0.2136)	(0.2080)
Washing Machine	0.6471***	0.6439***	0.6559***	0.6517***	-0.4637***	-0.4581***	-0.5523***	-0.5401***
	(0.1496)	(0.1374)	(0.1467)	(0.1450)	(0.1370)	(0.1264)	(0.1459)	(0.1439)
Building Condition	0.0711	0.0676	0.0738	0.0690	-0.0745	-0.0728	-0.0774	-0.0719
	(0.0816)	(0.0747)	(0.0746)	(0.0735)	(0.0618)	(0.0559)	(0.0566)	(0.0493)
CBD Distance	0.1534	0.1638	0.1389	0.1547	-0.0202	-0.0167	-0.0576	-0.0386
	(0.1258)	(0.1158)	(0.1259)	(0.1253)	(0.1205)	(0.1135)	(0.1350)	(0.1372)
School Location	-0.1165	-0.1097	-0.0807	-0.0723	0.2100	0.2011	0.1303	0.1282
	(0.1261)	(0.1185)	(0.1192)	(0.1207)	(0.2781)	(0.2585)	(0.2267)	(0.2322)
NO2 Levels	0.0690**	0.0638**	0.0538*	0.0458	0.0667*	0.0637*	0.0366	0.0274
	(0.0321)	(0.0312)	(0.0280)	(0.0293)	(0.0383)	(0.0357)	(0.0291)	(0.0295)
Dust-Sea Salt Particulates	-0.0046**	-0.0044**	-0.0031*	-0.0028	-0.0013	-0.0013	-0.0003	0.0000
	(0.0020)	(0.0019)	(0.0018)	(0.0019)	(0.0026)	(0.0025)	(0.0018)	(0.0018)
log(Population Density)	0.1904***	0.1893***	0.1784***	0.1778***	0.1878***	0.1870***	0.1520***	0.1537***
	(0.0369)	(0.0346)	(0.0345)	(0.0346)	(0.0391)	(0.0364)	(0.0351)	(0.0350)
Average Floor Space	0.0011	0.0011	0.0010	0.0010	0.0017**	0.0017**	0.0015	0.0014
	(0.0008)	(0.0007)	(0.0007)	(0.0007)	(0.0009)	(0.0008)	(0.0009)	(0.0009)
log(meannightlightpercapita)	-0.0154*	-0.0143**	-0.0119*	-0.0099	-0.0032	-0.0026	-0.0034	-0.0013
	(0.0075)	(0.0072)	(0.0064)	(0.0068)	(0.0085)	(0.0080)	(0.0065)	(0.0065)
ρ		0.0050		0.0073		0.0025		0.0104
		(0.0070)		(0.0069)		(0.0078)		(0.0074)
λ			0.4816**	0.5032**			0.6671***	0.6887***
			(0.2185)	(0.2168)			(0.1445)	(0.1472)
Observations	108	108	108	108	108	108	108	108
R^2	0.8479				0.7970			
Pseudo R^2		0.8483	0.8457	0.8456		0.7969	0.7898	0.7878

Robust standard errors in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

For our 2010 data, we again obtain some statistically significant results which match our hypothetical expectations. For our specifications with good housing characteristics, it is clear that increasing the proportion of households that have access to exclusive toilet facilities and washing machines increases monthly rental values. With regards to air quality, again NO2 levels were positively associated, which is again a surprising result which we attribute to omitted variable bias (using the same reasoning as above). Nonetheless, for most specifications, air quality as measured by dust and sea-salt particulates was negatively associated as expected, and was statistically significant for the majority of specifications involving good housing characteristics as our covariates. The most surprising result was regarding the roof and walls materials variable. The 'good' variant of this regressor was houses built with concrete walls and iron roofs, and our negatively associated and statistically significant coefficient suggests that increasing the proportion of households with houses constructed of concrete walls and iron roofs decreases monthly rental values. Again, our most robust finding is for population density, which was statistically significant to the 99% across all of our specifications. When taking our 2000 cross-sectional specifications into account, it is clear that population density unequivocally impacts monthly rental values, which supports the general literature on demographic control in the Philippines.

Unlike our 2000 cross-sectional regressions, the impact of spatial lags was more significant here, especially with regards to clustering of the error term. For both our SEM and SAC models, when using both good and bad housing characteristics as regressors, the spatial lag of the error term, λ, was positive and statistically significant to the 99% level. This subsequently means there are undoubtedly unobserved characteristics in neighbouring cities that positively impact monthly rental values, of which we have not accounted for. Yet, this result is surprising – our previous Lagrangian multiplier tests rejected the null hypothesis that $\rho = 0$ whilst concurrently failing to reject the null that $\lambda = 0$, meaning spatial autocorrelation should be driven by 'diffusion' in our dependent variable. Essentially, we should have expected the coefficient of the spatial lag of our dependent variable, ρ, to be statistically significant, whilst the coefficient of the spatial lag of the error term, λ, should have been statistically insignificant. Based on SEM and SAC specifications, this turned out to be the exact opposite of what the Lagrangian multiplier tests suggested. Whilst we are unsure of how to interpret this finding, we take note of this result for future exploration, and consequently treat the current results with a grain of salt.

4.5 Fixed Effects, Random Effects, and Spatial Autoregressive Models with Random Effects

Although we obtain some statistically significant results that match our hypothetical expectations, we attempt to make use of our two-period panel data, in order to account for any city-level unobserved heterogeneity that may augment omitted variable bias. For our good housing characteristic covariates, when running a Hausman test, we were unable to reject the fundamental null hypothesis that random effects, versus fixed effects, is the preferred specification, as we obtained a chi-squared test statistic of 22.41[4]. However, when further running a Breusch-Pagan Lagrangian multiplier test for random effects, we were also unable to reject the null hypothesis that $\text{Var}(\alpha_i) = 0$. Consequently, based on our inability to reject both the Hausman and Breusch-Pagan tests, the best specification for our data is a simple Pooled OLS. However, we still include the fixed effects and random effects specifications for illustrative and comparative purposes.

With regards to our bad housing characteristic covariates, when also running a Breusch-Pagan test, we were able to reject the null hypothesis that the variance of residuals across entities is zero, with a test statistic of 3.95, and subsequent p-value of 0.0235. Consequently, a random effects model, versus a Pooled OLS model, is justified for this panel of data. Our data also satisfies many of the associated necessary assumptions for random effects, although we heavily assume in this scenario that the unobserved effect, α_i, is uncorrelated with any of our explanatory variables, $\text{Cov}(x_{itj}, \alpha_i) = 0$ for $\forall t$, $\forall j$. Random effects is also advantageous here because it permits the incorporation of explanatory variables that are time-invariant, such as our covariate of 'distance to CBD' which fixed effects models would remove[5]. Additionally, because our census data relies upon random sampling of the

[4] We can only reject at the 90% level with a corresponding p-value of 0.0706.

[5] Distance to CBD is a time-invariant variable, as we measured this distance using the centroids of polygons within the city-level shapefile; this Euclidean geographic distance would thereby remain unchanged

population, and given we only study a sample of cities from the entire population, treating α_i as a random variable is not unreasonable. Also, we satisfy assumptions regarding no multicollinearity, as our variance inflation factor (VIF) values never reached greater than 10 for both good and bad housing characteristics. Finally, given the previous spatial diagnostic tests from table 3, we decide to also run our choice of spatial models – SAR, SEM, and SAC – with random effects.

between 2000 and 2010

Table 8: Fixed Effects and Spatial Random Effects Regression Results, 2010

	Good Characteristics			Bad Characteristics				
	Pooled OLS	Time FE	Time FE & RE	Pooled OLS	Time FE & RE	Time FE SAR RE	Time FE SEM RE	Time FE SAC RE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Roof and Wall Materials	-0.1105	-0.2330	-0.2079	0.5552	-0.1059	-0.1120	-0.1456	-0.1805
	(0.1438)	(0.1482)	(0.1520)	(3.3948)	(3.4486)	(3.3363)	(3.3456)	(3.3541)
Fuel for Lighting	1.0143***	1.1354***	1.0856***	-0.5074	-0.6642**	-0.6631**	-0.6322*	-0.6349*
	(0.2639)	(0.2633)	(0.2667)	(0.3326)	(0.3353)	(0.3288)	(0.3332)	(0.3337)
Primary Water Source	-0.0104	-0.0213	-0.0139	-0.8843**	0.0174	0.0054	0.1012	0.1101
	(0.0863)	(0.0850)	(0.0858)	(0.4494)	(0.4545)	(0.5379)	(0.5511)	(0.5534)
Toilet Facility	0.7790***	0.6530***	0.6244***	-3.7076***	-4.2721***	-4.2650***	-4.3446***	-4.3482***
	(0.1731)	(0.1763)	(0.1745)	(0.5037)	(0.4824)	(0.4912)	(0.5060)	(0.5063)
Garbage Disposal Manner	0.0171	0.0164	0.0177	0.3078	0.2866	0.2881	0.2744	0.2757
	(0.0217)	(0.0214)	(0.0214)	(0.2388)	(0.2413)	(0.2330)	(0.2336)	(0.2338)
Washing Machine	-0.0736	-0.0972	-0.1084	0.0057	0.0812	0.0813	0.0787	0.0792
	(0.1219)	(0.1249)	(0.1263)	(0.1048)	(0.1037)	(0.1000)	(0.0996)	(0.0996)
Building Condition	0.0277	0.1487	0.1299	-0.1921	-0.1122	-0.1142	-0.1014	-0.1022
	(0.1219)	(0.1276)	(0.1282)	(0.2596)	(0.2459)	(0.2390)	(0.2377)	(0.2377)
CBD Distance	0.2091	0.1463	0.1420	0.1202	0.1707	0.1678	0.1782	0.1736
	(0.1749)	(0.1735)	(0.1808)	(0.1473)	(0.1686)	(0.1656)	(0.1676)	(0.1706)
School Location	-0.1149	-0.3292	-0.3517	0.4099	0.2926	0.2960	0.2859	0.2885
	(0.2048)	(0.2158)	(0.2206)	(0.3477)	(0.3637)	(0.3526)	(0.3543)	(0.3548)
NO2 Levels	0.0603	0.0739*	0.0733	0.0635*	0.0595	0.0603	0.0586	0.0603
	(0.0434)	(0.0430)	(0.0451)	(0.0379)	(0.0433)	(0.0432)	(0.0419)	(0.0434)
Dust-Sea Salt Particulates	-0.0033	-0.0081*	-0.0080*	-0.0034	-0.0037	-0.0037	-0.0037	-0.0038
	(0.0041)	(0.0044)	(0.0046)	(0.0039)	(0.0044)	(0.0042)	(0.0043)	(0.0043)
log(Population Density)	0.2553***	0.2664***	0.2702***	0.2206***	0.2436***	0.2432***	0.2462***	0.2463***
	(0.0480)	(0.0474)	(0.0493)	(0.0447)	(0.0499)	(0.0485)	(0.0498)	(0.0498)
Average Floor Space	0.0028*	0.0021	0.0020	0.0026**	0.0013	0.0013	0.0015	0.0015
	(0.0015)	(0.0015)	(0.0015)	(0.0013)	(0.0013)	(0.0014)	(0.0014)	(0.0014)
log(meannightlightpercapita)	-0.0096	-0.0222	-0.0216	-0.0156	-0.0185	-0.0186	-0.0190	-0.0194
	(0.0141)	(0.0146)	(0.0153)	(0.0128)	(0.0147)	(0.0144)	(0.0143)	(0.0146)
ρ						-0.0010		-0.0021
						(0.0140)		(0.0143)
λ							0.1119	0.1151
							(0.5300)	(0.1793)
Observations	216	216	216	216	216	216	216	216
R^2	0.6658	0.6782	0.6769	0.7555	0.7477			
Pseudo R^2						0.7480	0.7467	0.7466

Standard errors in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Based on the results from table 8, for our good characteristics, it is clear that the inclusion of fixed effects and random effects specifications does nothing substantial to the data relative to our Pooled OLS estimates. This subsequently supports our results from the Hausman and Breusch-Pagan tests. However, for our bad characteristics, random effects does alter the results relative to the Pooled OLS estimates. For example, our random effects specifications remove any statistical significance for the 'primary water source' covariate, and even switches the coefficient sign. Nevertheless, our spatial models with random effects also did nothing substantial to alter the results, with both spatial lags on the dependent variable and error term being statistically insignificant for all specifications – SAR, SEM, and SAC.

Nonetheless, it is worth noting that for all of our statistically significant results, for both the good and bad housing characteristics, the direction of the coefficients matches our hypothetical expectations. For example, for our 'toilet facility' covariate, the proportion of households that have 'water-sealed sewer septic tank used exclusively be household' is statistically significant and positive, whereas the proportion of households that use 'open pits' is statistically significant and negative. Another very robust result is with regards to 'fuel for lighting': the proportion of households that have access to 'electricity' is statistically significant and positive, whereas the proportion of households that use 'kerosene' is statistically significant and negative. Albeit a less robust result, we do see that air quality is consistently negatively correlated with monthly rental values, and for columns 2 and 3 is statistically significant to the 90%. This result is to be expected – as poorer air quality increases, as measured by the level of dust and sea-salt particulates in the air, monthly rental values should decrease. Moreover, although being statistically insignificant across all specifications, 'building condition' switched coefficient signs as expected when using the good and bad panels of data. Finally, and as has been shown in our cross-sectional specifications, 'population density' is positively associated with expected monthly rental values across all specifications.

Given the above results, our 'quasi-robustness check' of using good and bad variants for each of our internal and external housing covariates appears to be a valid method for empirically assessing our data. Although some coefficients have unexpected and opposite associations to monthly rental values, such as the good variant of 'roof and wall materials', these coefficients are statistically insignificant, and so we have confidence in those results that are statically significant. Overall, based on the results in table 8, we can firmly say that fuel for lighting, toilet facilities, and population density are significant determinants of monthly rental values in and around the cities of the NCR.

5 Discussion

This paper has successfully applied the hedonic house price model to the NCR and its contiguous provinces. Using a range of cross-sectional spatial econometric specifications and fixed and random effects models, we show that fuel for lighting, toilet facilities, and population density are key determinants of monthly rental values in and around the cities of the NCR. In terms of policy implications, our results evidence the importance of fuel for lighting and toilet facilities as internal characteristics of houses. Moreover, the robust significance of population density supports the findings in the literature that population growth is a huge factor in increasing rent values, and so demographic controls should be a key priority of the Philippine government, as it precipitously increases monthly rental values. Nonetheless, in spite of our findings, there are still a hive of limitations which plague the paper. Firstly, there are clear endogeneity issues in terms of omitted variable bias and reverse causality. We try to use panel data methods to control for as much omitted variable bias as possible, but using only one blunt confounder – GDP per capita – is clearly not an ideal strategy. Due to the lack of a valid instrument, we are also unable to solve reverse causality either: we still do not know whether rent values increase the quality and/or quantity of internal and external housing characteristics, or whether the relationship works in reverse.

The largest constraint this paper faced was with regards to data. Firstly, whilst the PSA has an abundance of interesting data with which to utilise, its disorganised arrangement made it incredibly difficult to obtain for regression purposes; the fact that most census data was provided in PDF format engendered the painful and painstaking process of manually scraping data into spreadsheets. Secondly, whilst we are fortunate to have the covariates we use, more granular data for the external housing characteristics, such as crime rates and natural disasters, would be incredibly beneficial to expand upon the scope of factors that impact monthly rental values. Finally, because our unit of analysis is the city-level, given we do not have individual housing transaction data, it is apparent that our paper suffers from the Modifiable Areal Unit Problem (MAUP), as we are significantly aggregating our covariates upwards. As such, if better, more accessible data could be provided by an authority such as the PSA, that would hugely increase the research capacity on the Philippine housing market.

Thus, given all of the aforementioned limitations, we treat our results as being purely correlational and suggestive. However, whilst we do not seek to overemphasise the paper's methodological drawbacks, our efforts are by no means futile. It is clear that fuel for lighting, toilet facilities, and population density are the biggest determinants of monthly rental values in the Philippines' major cities, with housing inequality remaining inertial over a decade; expensive and cheaper monthly rental values also cluster in urban and suburban provinces respectively. Consequently, we hope to have developed a brief literature on studying the housing market of the NCR within the Philippines, of which future studies should unequivocally explore further.

References

Akbar, M., & Altaf, A. (1995). A Rent Hedonic Model of the Low Income Housing Market in Karachi, Pakistan. Pakistan Economic and Social Review, 79-90.

Alonso, W. (1964). Location and Land Use: Toward a General Theory of Land Rent. Boston: Harvard University Press.

Anselin, L., Bera, A., Florax, R., & Yoon, M. (1996). Simple diagnostic tests for spatial dependence. Regional Science and Urban Economics, 77-104 .

Ashenfelter, O., & Graddy, K. (2003). Auctions and the Price of Art . American Economic Review, 763-787.

Asian Development Bank. (2019). Poverty in the Philippines. Retrieved January 10, 2019, from Asian Development Bank: https://www.adb.org/countries/philippines/poverty

Ballesteros, M. M. (2002). The Dynamics of Housing Demand in the Philippines: Income and Lifecycle Effects. Philippine Institute for Development Studies,2002(01). Retrieved February 11, 2019, from https://pdfs.semanticscholar.org/c6d9/f7582f77bef0a7b225d8af37c 3fd4b6d3fd1.pdf.

Brooks, W. (n.d.). Notes on how to compute Gini Coefficient. Retrieved January 25, 2019, from Wyatt Brooks: https://www3.nd.edu/ wbrooks/GiniNotes.pdf

Chen, W., & Jim, C. (2010). Amenities and disamenities: a hedonic analysis of the heterogeneous urban landscape in Shenzhen (China). The Geographical Journal.

Cotteleer, G., Gardebroek, C., & Luijt, J. (2008). Market Power in a GIS-Based Hedonic Price Model of Local Farmland Markets. Land Economics, 573-592.

Court, A. (1939). Hedonic Price Indexes with Automotive Examples. Dynamics of Automobile Demand, 99-117.

Halvorsen, R., & Pollakowski, H. (1981). Choice of functional form for hedonic price equations. Journal of Urban Economics, 37-49.

LeSage, J., & Pace, R. (2009). Introduction to Spatial Econometrics. Florida: Chapman and Hall/CRC .

Ozus, E., Dokmeci, V., Kiroglu, G., & Egdemir, G. (2007). Spatial Analysis of Residential Prices in Istanbul. European Planning Studies, 707-721.

Park, R., & Burgess, E. (1925). The City. Chicago: University of Chicago Press.

Requenta-Silvente, F., & Walker, J. (2006). Calculating Hedonic Price Indices with Unobserved Product Attributes: An Application to the UK Car Market. Economica, 509-532.

Sen, A. (1999). Development as Freedom. Oxford: Oxford University Press.

Socioeconomic Data and Applications Center. (n.d.). Global 3-Year Running Mean Ground-Level NO2 Grids from GOME, SCIAMACHY and GOME-2, v1 (1996?–?2012). Retrieved from Socioeconomic Data and Applications Center: http://sedac.ciesin.columbia.edu/data/set/sdei-global-3-year-running-mean-no2-gome-sciamachy-gome2

Socioeconomic Data and Applications Center. (n.d.). Global Annual PM2.5 Grids from MODIS, MISR and SeaWiFS Aerosol Optical Depth (AOD) with GWR, v1 (1998?–2016). Retrieved from Socioeconomic Data and Applications Center: http://sedac.ciesin.columbia.edu/data/set/sdei-global-annual-gwr-pm2-5-modis-misr-seawifs-aod

Tronstad, R., Huthoefer, L. S., & Monke, E. (1992). Market Windows and Hedonic Price Analyses: An Application to the Apple Industry. Journal of Agricultural and Resource Economics.

Xiao, Y. (2016). Hedonic Housing Price Theory Review. Springer Geography Urban Morphology and Housing Market,11-40. doi:10.1007/978-981-10-2762-8_2

The Economic Consequences of Crime: Evidence from the *Unidades de Polícia Pacificadora* in Rio de Janeiro

Savannah Kochinke[1], Edgar Akopyan[1], Somjeeta Chatterjee[1], Ardi Janjeva[1], and Napat Matchaseth[1]

[1]The London School of Economics and Political Science, London, UK

Abstract

Brazil has been one of the most prominent country case studies in academic research charting the relationship between crime and economic development. The issue of violent crime, in particular, has taken centre stage in the past year, as record high murder rates have coincided with the election of President Jair Bolsonaro. His campaign promised of tough anti-crime policies, which resonated with Brazil's track record of a *mano dura* - heavy-handed - approach to public security. The city of Rio de Janeiro, specifically, has undertaken a process of militarisation and implementation of police operations in its favelas to combat violence and improve security before the 2014 FIFA World Cup and the 2016 Olympic Games. Using an instrumental variable regression, we control for endogeneity in crime rates across the Integrated Circumscriptions of Public Security (CISP) in the municipality of Rio de Janeiro. This is done by utilising the implementation of the 38 Unidades de Policia Pacificadora (Police Pacification Units, UPPs) to study the effects of crime on nightlight data, a proxy for income at the local level. Our regression results suggest that, when controlling for homicide rates and income levels in neighbouring CISPs in the prior month, UPPs had a statistically significant effect on reducing crime rates. Contrary to the academic literature, however, our primary findings indicate that decreases in crime rates actually corresponded with statistically significant decreases in future nightlights. We explore possible reasonings for this, potential for future studies, and the effects of other types of violent crime, such as rape and theft, robberies, and deaths by police officers on future income levels.

1 Introduction

Brazil has been one of the most prominent country case studies in academic research analysing how violent crime affects economic development. This attention has only intensified given its hosting of two major sporting events in 2012 and 2014. The Olympic Games and the FIFA World Cup presented Brazil's politicians with a fresh incentive to address violent crime rates which have plagued the country for so many years. It also provided them with the chance to recognise the nuances showing how the effects of crime are heterogeneous across different segments of society. In Brazil, the victims of the most violent crimes are disproportionately young, male and poor – 'Brazilian males aged 15-24 are 15 times more likely to be killed than women', and Afro-Brazilians are two-thirds more likely to be killed than whites' (Kingstone, 2012). From 2008, the Brazilian government rolled out its *Unidade de Polícia Pacificadora* (UPP) programme in Rio De Janeiro, a state where murder rates have reached highs of 57 per 100,000 this century (Kingstone, 2012). Since then, there has been a great deal of controversy as to whether this new approach has adequately accounted for those complexities, or merely continued a long trend of heavy-handed, 'mano-dura' crime policy. Understanding the extent to which the UPP's in Rio De Janeiro have affected economic development through crime reduction is the fundamental undertaking of this paper. Our regression results did establish that when controlling for homicide rates and income levels in neighbouring CISPs in the prior month, UPPs had a statistically significant effect on reducing crime rates. We then establish how decreasing crime rates may affect income levels. Using nightlight data as a proxy for income at the local level, we analyse the impact that crime has on these datasets. Our primary findings here are that decreases in crime rates actually correspond with statistically significant decreases in future nightlights. We elaborate on this somewhat surprising result in our discussion section. First, we set out the appropriate contextual framework for our economic analysis. We look at the broad literature that ties crime and development outcomes together, before focusing on crime trends and crime policy in Brazil in more detail. A brief history of the UPP is outlined based on 'Werling's Model on Stages of Pacification'. Our methodology section explains how ISP data and nightlight data formed the basis of our economic analysis, and ends by considering some limitations to these methods. The aforementioned regression results are explained in more detail, and we conclude with a summary of our findings and suggestions for areas of further research.

2 Motivation and Context

2.1 Crime and Development

Crime is a major issue in many countries, particularly in the global South (Carrington et al., 2018). Due to this reason, a lot of effort has been put into trying to understand the dynamics of trends in crime and economic indicators as well as their causal relationship. High crime rates have multiple destructive socio-economic effects. Macro research conducted by Enamorado et al. (2014) shows that drug-related crimes substantially affected economic growth in areas of Mexico. Similarly, analysis of crime in Italy (Detotto, Otranto, 2010) showed that a 1% increase in crime causes a reduction in GDP growth of 0.00022%. In practice, long run crime is 5% more expensive during recessions.

Those studies primarily look at the overall effect of crime on economic development, but there is a considerable literature discussing the potential mechanisms through which crime deters economic growth. To some extent, this could be explained by reductions in economic diversification and economic complexity and increases in sector concentration (Rios, 2016). Although some studies do not find that crime affects educational decisions (Marquez-Padilla, Fernanda, Francisco Perez-Arce, and Carlos Rodriguez-Castelan, 2015), human capital and labour generally can be affected in other ways. Ben, Yishay and Pearlman (2013) found a negative relation between homicide and labour supply: an increase of homicide rate by 10 per 100 000 reduces weekly hours worked by 0.3 hours and this effect is greater for self-employed people. Studies also point to an exodus of labour due to high drug crime rates (Ríos, 2014). Crime and violence levels also have an impact on firm behaviour: an increase in the amount of violence leads to decreases in competitiveness and reduces sales growth, (Gaviria, 2002) while firm-specific kidnappings create disincentives for investment (Pshisva, Suarez, 2010). Empirical evidence also indicates a heterogeneity in the way crime affects different firms: bigger firms tend to be less affected than smaller and medium sized businesses (Islam, 2014). This could be due to the ability of larger firms to spend money on protecting against theft. It may also be linked to their larger revenues enabling them to corrupt local officials and stay in business. Given all these negative effects of crime on economic development, crafting effective policies in this area is crucial for sustainable long-run growth.

The research into the relationship between crime and development is further complicated by the problem of causality. While it could be the case that crime has a negative effect on economic activity and growth, it could also be the case that poverty and the other symptoms stemming from low economic and social welfare increase the rate of crime in communities or that economic growth can make crimes such as theft much more lucrative. There are also many other aspects of the economy that can significantly affect the amount of crime in a region. In their analysis of cities in the United States, Fleisher (1966) and Ehrlich (1973) looked at the effect of unemployment rates and inequality on crime and concluded that both have a crime-inducing impact. More recently, Fajnzylber et al. (2002) analysed 39 countries from 1965–95 and 37 countries from 1970–1994 to show that an increase in income inequality raises crime rates. Moreover, as the author highlights, "both economic growth and income inequality are robust determinants of violent crime rates". Interestingly, however, violent crime decreases with income growth.

2.2 Crime in Brazil

Brazil is recognised as one of the most violent countries in the world. As depicted by Figure 1, in 2019, the Global Crime Index Map gave Brazil a score of 70.24 out of 75 on the Crime Index, making it the 7th most crime-ridden country in the world, behind Venezuela, Papua New Guinea, Honduras, South Africa, and Trinidad and Tobago. However, the majority of deaths in Brazil are by homicide. Figure 2 shows that cities such as Porto Alegre, Manaus, Caruaru and Maraba are just some of the cities in Brazil, which are among the 50 cities with the highest homicide rates per 100,000 population in 2016. (The Economist, 2017). Historically, Rio de Janeiro, which is the primary target of the UPP, witnessed 56.5 homicides per 100,000 inhabitants in 2002; although this number had decreased to 20.2 by 2012 (Waiselfisz, 2014).

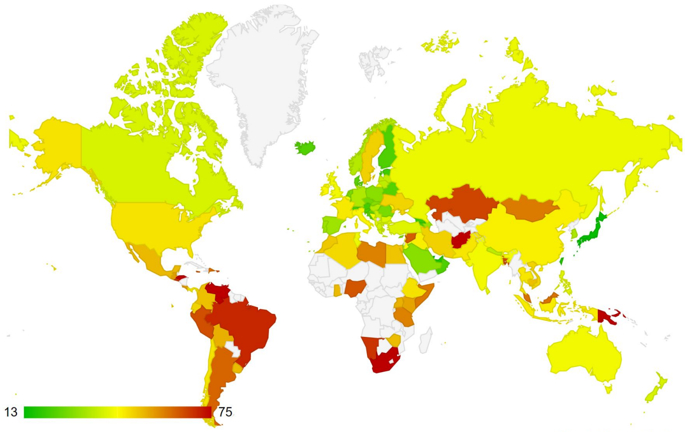


Figure 1: Global Crime Index Map (2019)

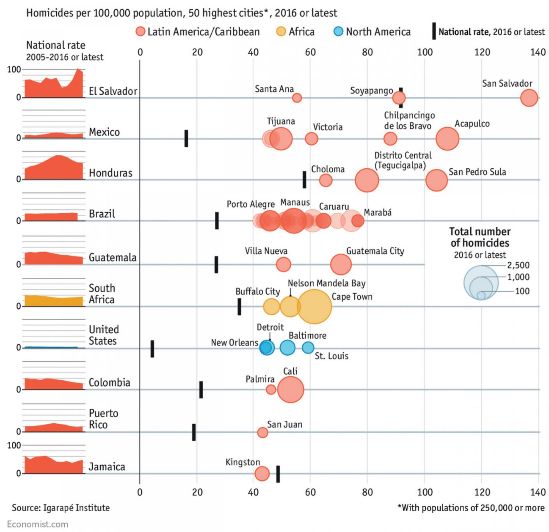


Figure 2: Globally top 50 cities with the highest homicide rates in 2016

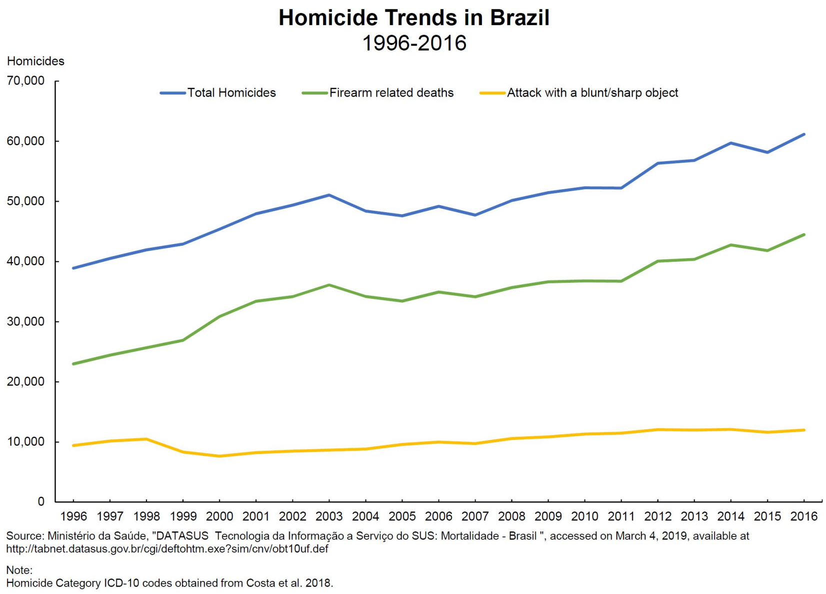


Figure 3: Homicide Trends in Brazil (1996-2016)

Figure 3 further highlights the fact that homicides have increased in Brazil between 1996–2016. Furthermore, it is clear that firearm related deaths, as well as attacks with a sharp or a blunt object, tend to be the most common types of homicides that are carried out in Brazil. An important societal development over the last decade has been rapid urbanisation, a trend that Brazil has shared with much of Latin America. This has resulted in the creation of areas devoid of basic infrastructure and public goods. These vulnerable areas are hotspots for violence and crime, and are dominated by gangs and drug traffickers. Thus, they are undeniably a persistent problem for densely populated cities like Rio de Janeiro (Miraglia, 2016). The haphazard urbanisation, coupled with extremely weak state presence, has made Brazilian cities a lucrative hub of violence - one that enables both organised gangs and petty criminals to flourish.

It is unsurprising that much of the criminal activity in Brazil is based upon drug-trafficking. Crimes such as robberies, homicides, abductions and youth violence as well as many non-lethal crimes are often provoked as a side-effect of drug-peddling. Illicit drugs such as crack cocaine began gaining prominence in Brazil from 1987, (Domanico, 2010) and gained popularity amongst the poverty-stricken residents situated in Salvador and São Paulo.

By sharing borders with three key producers of cocaine, namely 'Colombia, Peru and Bolivia' (Miraglia, 2016), Brazil's emergence as the second largest consumer of cocaine in the world (Central Intelligence Agency, 2019). Hence, this makes it a prime target for crimes related to drug-trafficking and abuse. *Primeiro Comando da Capital* (PCC) is undisputedly the largest criminal organisation in Brazil (OSAC, 2018), which operates large-scale international arms and drug-trafficking. Based in São Paulo, this violent prison gang poses as a serious hindrance to the Brazilian government's efforts at significantly reducing violent crime throughout country.

To the extent that these factors combine to exacerbate the deep-rooted levels of inequality in Brazil, much of the Brazilian population remain trapped in a cycle of violence, or at least the spectre of it. In response to this, they are met with violent police strategies, further reinforcing a climate of fear and persistent violence. In such an environment, it is no surprise that investment climate assessments in Brazil regularly identify crime and violence as a key constraint to business growth (World Bank, 2006). World Bank Enterprise Surveys show crime to be one of the major hurdles for productivity and micro and macroeconomic growth throughout Latin America (Ascher, 2012). This is a reflection of how a drive for elite capital accumulation strategies have come at the expense of widening urban inequalities; when these widen far enough, the implications for crime rates can be significant. The fact that there was a bias towards implementing UPP's closer to richer areas, or stadiums hosting big sporting events, indicates a drive to control informal settlements instead of integrating them and promoting their rights to social citizenship. Rather than conceiving the Brazilian state as an absent actor in these areas, it should be held to account for its complicity in perpetuating urban inequalities over generations, and in turn entrenching criminal practices across communities.

2.3 Crime and Violence Policy in Brazil

Having an understanding of why *mano dura* policies prior to the UPP period have failed is crucial to understanding the limitations of the UPP programmes themselves. The failure of successive governments to curb violent crime rates is what paved the way for the possibility of state police entering a *favela* with such a powerful mandate. The so-called '*mano dura*' approach to violent crime has been prevalent across Latin America in recent decades. These policies may include increased police discretion in arresting suspects on subjective evidence, as well as sentencing for minor offences. Such practices tend to be followed by pre-trial detention, unauthorised searches, and extrajudicial confessions, fundamentally threatening basic human rights and the rule of law. The presence of military police on a more permanent basis than a state of 'emergency crisis' is what usually gives the practice a perceived legitimacy (Igarape Institute, 2018). In Brazil, these elements have traditionally reinforced an environment where repressive policing has crowded out the building of trust between police and communities. Some initiatives which exemplify this are the implementation of *Posto de Policiamento Comunitário* (Community Police Posts) and the *Grupamento de Policiamento em Áreas Especiais* (GPAE) (Police Grouping in Special Areas) (Janssen, 2015). In response to the former, some civil society groups were so disillusioned by acts of police abuse that they instituted their own human rights training projects in areas like Pernambuco and Sergipe (Leeds, 2007).

There has been a trend of responding to crises with grandiose gestures rather than committing to a long-term approach that actors across the political spectrum can rally around. In 2000, President Cardoso launched the 124-step National Plan for Public Security aimed at encouraging policy reform at the state level. One of these steps was the Plan for Integration and Monitoring of Social Programs for the Prevention of Urban Violence (PIAPS) which aimed at coordinating different levels of political power and different subsections of social policy to improve the life prospects of the youngest and most vulnerable. By 2003, however, this plan had been replaced by President Lula's Single Public Security System (SUSP), which itself failed to have any lasting impact (Cano, 2006). The various failures at government level in coordinating a comprehensive crime policy package have only fed into the proliferation of repressive approaches at the local level. The most recent developments in Brazil seem to confirm the new government's intention to give *mano dura* policies a new lease of life. In February 2019, Justice and Public Security Minister Sergio Moro unveiled the government's new anti-crime package, prompting concerns over human rights protections and leniency towards the use of police force. (Reuters, 2019). While of course still very early to judge, the policy direction the country is headed in now brings to mind the old adage that '*mano dura* works, it just has not been done properly yet'. The fact that this approach to crime has remained resilient to a number of historical failures makes understanding the implications of such policies extremely important for the future of Brazil.

2.4 History of the UPP

Solving the problem of violent crime became a priority amongst Brazilian policy makers once again when it was announced that Rio would host both the 2014 FIFA World Cup, as well as the 2016 Olympics. Historically, government intervention in Rio de Janeiro's *favelas* was limited to short term initiatives and military invasions (Werling 2014, Skogan 2013, Cano 2012, WB 2012, Da Cunha and Mello 2011). Without the required political commitment and support, these programs were often discontinued, leaving behind negative perceptions of the police in communities (Werling 2014; CESeC 2011, Freeman 2012, WB 2012, Da Cunha and Mello 2011). Pioneered in 2008 by José Mariano Beltrame, the State's Public Security Secretary at that time, the UPP program attempted to break from this history, and aimed to integrate military interventions in *favelas* with the implementation of a permanent police presence and long-term social development and monitoring programs.

Pacification had two vital aims:

1. Resumption of territories that are dominated by militias, drug trafficking or other criminal organisations (Article 1, Government of the State of Rio de Janeiro, 2011).
2. To permanently guarantee security and respect for human rights of the local population, and to allow social occupation of the *favelas* (Article 1, Government of the State of Rio de Janeiro, 2011).

A legal decree also specifies that the local government should coordinate all related public and private entities and provide public services to the residents of the *favela* within 120 days of occupation by the UPP (Rekow, 2016; Article 6, Government of the State of Rio de Janeiro, 2011). The core idea was to use these methods to facilitate communication between the residents of the *favelas* and the local government in order to strengthen local leadership and increase the safety of the community.

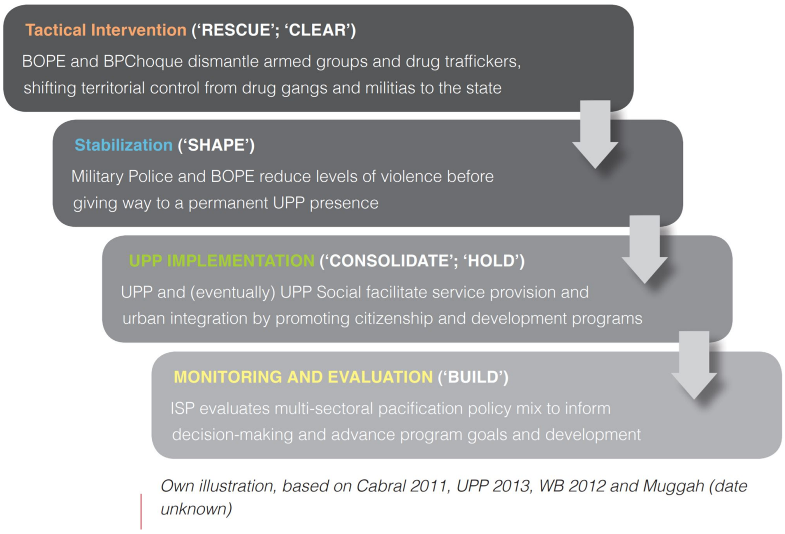


Figure 4: Werling's Model on Stages of Pacification

As observed in Figure 4, Werling's model of the UPP pacification process (Werling, 2014), there are four overlapping stages of pacification. The *Batalhão de Operações Policiais Especiais* (Special Police Operations Battalion, BOPE) and *Batalhão de Polícia de Choque* (Police Shock Battalion, BPChoque) first arrest criminals and seize drugs, arms, and ammunitions and during the Tactical Intervention stage, thereby taking control of the *favela* from the local gang (Werling, 2014; Rekow, 2016). Simultaneously, citizens are encouraged to report criminal activity during the overlapping Stabilisation period, further reducing crime in the area. The Implementation stage is carefully coordinated with the exit of the BOPE, who are then replaced by a new permanent police force, who are supposed to go on to conducting the Monitoring and Evaluation Stage. Overall, as shown in Figure 5 below (Introduction to International and Global Studies, 2015), the UPP programme had been established in 39 out of more than 700 *favelas* in Rio de Janeiro, many of which continue to be inflicted with violence. As mentioned above in Section 2.3, the *favelas* nearest to Olympic and World Cup zones were prioritised for pacification.

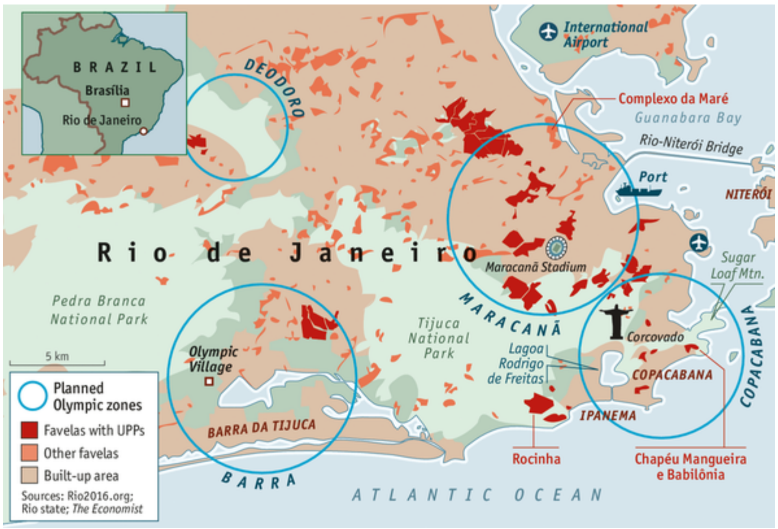


Figure 5: The location of UPPs and Olympic zones in Rio de Janeiro

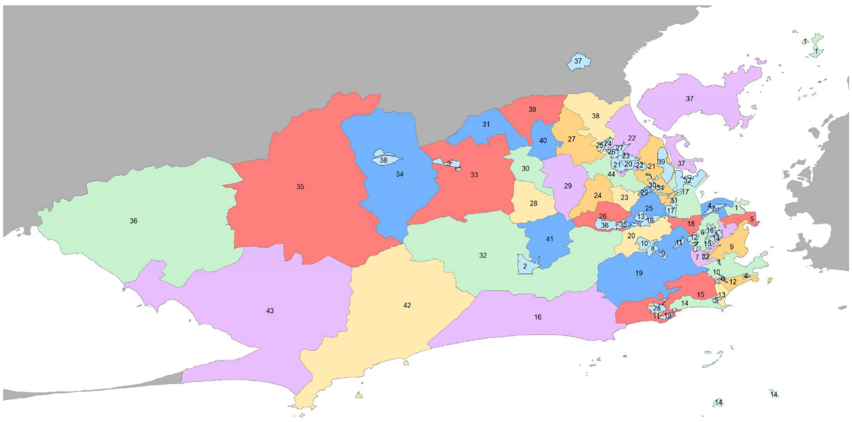
3 Methodology and Data

3.1 ISP Data

Data on violence statistics in the Rio de Janeiro municipality were obtained from the Instituto de Segurança Pública (ISP), the Public Security Institute, of Rio de Janeiro. The ISP was created in 1999 through Law 3,329 with the purpose of assuring, executing, managing and administering, through the Civil and Military Police, the public security policy of the State of Rio de Janeiro (Government of the State of Rio de Janeiro, 1999). The primary mission of the ISP is to produce information and disseminate research and analysis to assist with the implementation of public security policies.

The ISP has a detailed database of the criminal records and police activity in the state of Rio de Janeiro. Security statistics are based on Occurrence Records from the Civil Police Stations within the state, and additional information from the Military Police. The ISP consolidates the data after carrying out a quality control review of the data through the *Corregedoria Interna da Polícia Civil* (COINPOL), the Internal Control of the Civil Police. The ISP also provides additional information on the territorial divisions within the state, population distributions at this level, and other institutional information. Territorial division, from broadest to narrowest, are the Integrated Regions of Public Security (RISP), Integrated Public Security Areas (AISP), and the Integrated Circumscriptions of Public Security (CISP), which is our primary area of interest. Figure 6 shows the CISPs in the municipality of Rio de Janeiro.

Municipality of Rio De Janeiro division by Integrated Circumscriptions of Public Security (CISP)[1]



Source: Instituto de Segurança Pública, (ISP) ISPDatos, accessed March 9, 2019.

Note:
[1] Integrated Circumscriptions of Public Security (CISP) constitute the sphere of territorial integration, at an operational level, of the Integrated Military Police Companies of the State of Rio de Janeiro (PMERJ). The basic principle of CISP is the concept that the responsibility for policing a subarea of the Integrated Military Police Company should, whenever possible, coincide with the circumscription of a police station. Thus, the CISP correspond to the territorial areas of action and joint responsibility of the Integrated Companies and Police Stations.
[2] ISP geographic data is not updated to include CISP 45, which is within the southern boundaries of CISP 22.

Figure 6: CISP Map

Overlaid on the map are the 39 UPPS that were implemented between November, 19, 2008 and March 3, 2014. Figure 14 in the Appendix provides the occupation and installation dates of the 39 UPPs. The implementation of a UPP of Complexo da Maré, close to Rio's international airport, which was the most recent UPP, was so difficult that it was indefinitely postponed (Tardáguila, 2016). The ISP and UPP websites no longer list Complexo da Maré as a UPP, but we include the occupation date in our analysis. Using the geospatial mapping software ArcMap and the territorial shapefiles from the ISP, we mapped each UPP to the CISP or CISPs in which the policing unit was installed. Our analysis relies on the ISP's monthly evolution of statistics within CISPs. We analyse the violent crime statistics and population data from the 38 CISPs in the municipality of Rio de Janeiro for 190 months of data, starting in January of 2003, giving us a total of 7,220 observations. In addition to monthly population at the CISP level, we specifically utilise the

Table 1: ISP Crime Variables

Original Variable	Translated Variable
Homicídio doloso	Willful homicide
Lesão corporal seguida de morte	Bodily injury followed by death
Latrocínio (roubo seguido de morte)	Robbery (theft followed by death)
Morte por intervenção de agente do Estado	Death by intervention of State agent
Tentativa de homicídio	Attempted homicide
Lesão corporal dolosa	Willful bodily injury
Estupro	Rape
Total de roubos	Total robbery
Total de furtos	Total thefts

3.2 Nightlight Data

As the ISP does not collect income or other economic growth-related statistics, we utilised satellite data on night lights as a proxy for income growth measures. We follow the statistical framework developed by Henderson, Storeygard, and Weil (2012) to estimate growth for countries with poor income and GDP data to estimate the monthly income changes at the CISP level. Nightlights data are collected by the United States Air Force Defence Meteorological Satellite Program's (DMSP) polar orbiting satellites, which record radiance at night. The National Oceanic and Atmospheric Administration's (NOAA) National Geophysical Data Center (NGDC) then processes these raw data to isolate man-made light, from sources such as cities, industrial sites, or gas flares. However, before proceeding further we need to establish that changes nightlights can be used as a proxy for changes in income. To establish the appropriateness of light data we study the correlation between the annual nightlight composites for all of Brazil to annual GDP data for the country from 1993 to 2013 (Henderson, 2012). The data for GDP purchasing power parity (PPP) is obtained from the World Bank (2019) and is computed at constant 2011 international dollar prices. This choice is motivated by the fact that we are interested in real changes in the GDP measure and not just in nominal changes. Figure 7 plots this data and we have found a correlation of approximately 0.70, which we believe establishes it as a viable proxy for income.

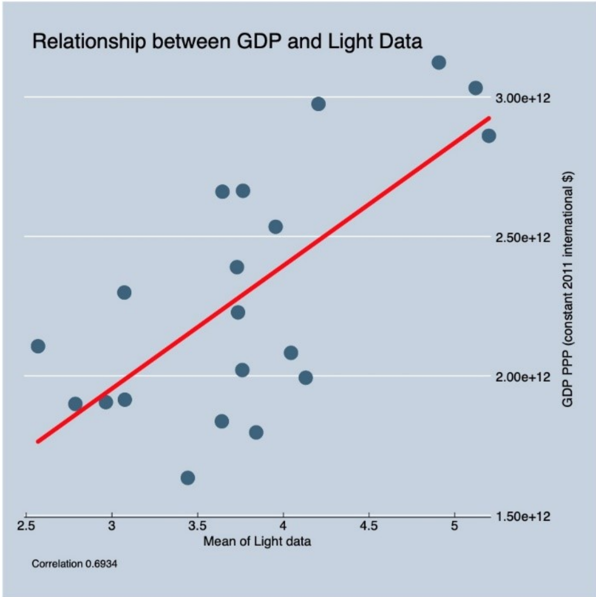


Figure 7: Correlation between GDP and Nightlight Data

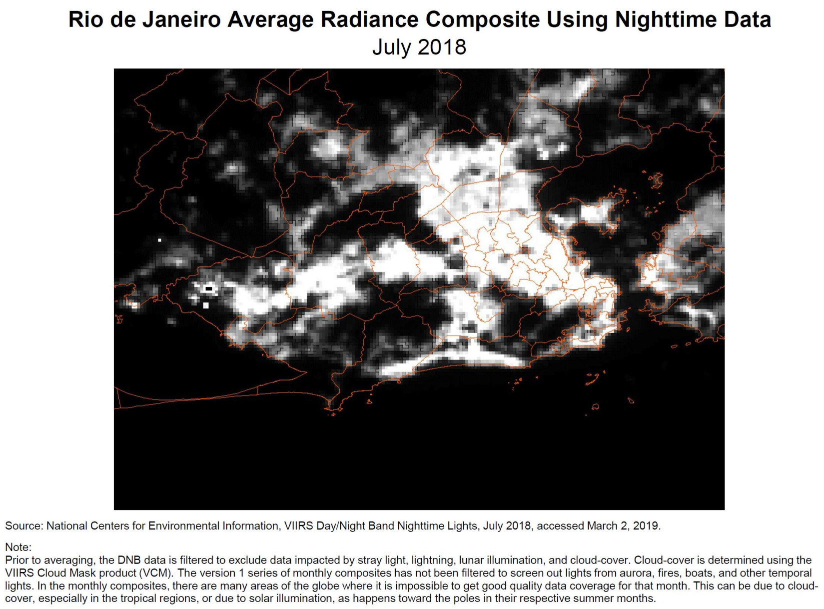
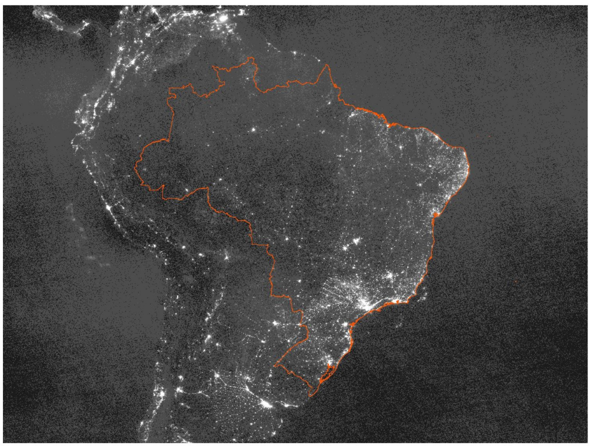


Figure 8: Average Radiance Composite Light Composite for Rio de Janeiro

We depict the most recently available nightlights data for Brazil in Figure 8, which shows that there is a clustering of lights around major cities such as Brasilia, Rio de Janeiro and São Paulo. Having shown that changes in nightlights are a useful proxy for economic activity, we then utilise the NOAA's Earth Observations Group's (EOG) Version 1 VIIRS Day/Night Band Nighttime Lights dataset to calculate the mean night lights for CISP. The Day/Night Band data excludes natural sources of light just as lightning, lunar illumination, and cloud-cover, and is averaged on a monthly basis from 2012 through 2018 (noting that June 2018 appears to be missing from the EOG's database). Figure 9 shows an example of the average radiance composite overlaid with the CISP boundaries. Using the Spatial Analyst Tools in ArcMap, we created zonal statistics tables that averaged the radiance within each CISP area for each month of available data and merged this with our existing dataset of crime statistics.

Brazil Average Visible Nighttime Light Composite Data
2013



Source: National Centers for Environmental Information, DMSP-OLS Nighttime Lights Time Series, 2013, available at https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html, accessed March 2, 2019.

Note:
The NOAA makes their cloudfree nighttime light composites using all the available archived DMSP-OLS smooth resolution data for calendar years 1992–2013.

Figure 9: Average Visible Nighttime Light Composite for Brazil

To control for the income of neighbouring CISPs, we additionally created a spatially weighted matrix of average radiance neighbouring CISPs. This includes data of some of the CISPs that border the CISPs that outside of the municipality of Rio de Janeiro.

3.3 Methodology and Empirical Strategy

We first perform a preliminary analysis of the effects implementing a UPP on the crime rate using pooled OLS and CISP and Time Fixed effects regressions. Given CISP i at month-year t, we calculate the change in the *CrimeRate* after the implementation of a UPP, given by the dummy variable *PostUPP*. We perform this model using the homicide, thefts and robberies, deaths caused by police officers, and rape rates. As the tactical intervention stage was a rapid process that resulted in the immediate arrests of major criminals in CISPs, we look at the contemporaneous effect of crime rates after a UPP occupation as our primary regression, but also consider regressions with a one month, six month, and one year lag of UPP occupation to study the effect over time. We also include control variables for a one month lag of the *CrimeRate*, to control for violence in the previous period, a log of the mean night lights, which we will call *LogIncome*, and the logged value of the mean income of neighbouring CISPs *LogIncomeNeighbor*, to control for the wealth of the surrounding communities. As we only have data on nightlights from May 2012 through the present, this reduces our observations by 4,332 to 2,888. Our autoregressive distributed lag model is:

$$\begin{aligned} CrimeRate_{it} = {} & PostUPP_{it} + CrimeRate_{it-1-s} \\ & + LogIncome_{it-1-s} + LogIncomeNeighbor_{it-1-s} \end{aligned} \tag{1}$$

where $s = 1, 6$ or 12.

We additionally run these regressions with entity and time fixed effects to account for idiosyncratic differences across different times of year (as we will demonstrate below, we found some seasonality in homicide rates that corresponds with the existing literature on seasonality of crime) and the characteristics of the CISPs. Next, we ran an OLS regression for the change in nightlights on the crime rate. We run three iterations of this model, lagging the crime rate by 1, 3, and 12 months to examine the effects that crime rate had on changes in nightlights over time. Similarly repeated this with controls for LogIncome and *LogIncomeNeighbor* that are lagged an additional month and with and without entity and time fixed effects. In this regression, we omit the lags *CrimeRate*, because of concerns regarding multicollinearity.

$$LogIncome_{it} = CrimeRate_{it-s} + CrimeRate_{it-1-s} + LogIncome_{it-1-s} + LogIncomeNeighbor_{it-1-s} \tag{2}$$

where $s = 1, 6$ or 12.

We utilise this approach to show some of the initial relations between UPP occupation and crime rates, and crime rates and income. There are various reasons why these OLS estimates could potentially be biased or inconsistent. These include the possibility of reverse causation between crime rates and changes in income, omitted variable bias, and possible measurement error with respect to certain types of crime (some types of violent crime, such as rape or theft may be under reported while some types of violent crime, such as homicide are more difficult to hide). In light of this, we employ a more robust model below.

As described in Section 2.1, the issue of reverse causality arises when analysing the effect of crime on poverty, or poverty on crime. A potential strategy to deal with these endogeneity concerns is to apply instrumental variable techniques. The conditions for an instrumental variable are relevance of the variable, that the IV is related to the potentially endogenous variable, and the exogeneity of the IV with respect to the dependent variable. We posit that the occupation of UPPs, which we hypothesise have an immediate subsequent reduction in crime rates, can function as an instrument for changes in income on crime rates as UPP occupation can only change income through reduction of the crime rate. Here, we utilise the Two Stage Least Squares estimation method. In our first stage, we regress the potentially endogenous variable, *CrimeRate* on the exogenous variable *PostUPP*. Again, we run this regression contemporaneously, and run versions of the model with and without fixed effects and with and without controls for prior *CrimeRate*, *LogIncome*, and *LogIncomeNeighbor*. To prove the relevance of *PostUPP* and *CrimeRates*, we expect to see a strong and significant relationship between the two variables.

$$CrimeRate_{it-12} = \gamma_0 + \gamma_1 PostUPP_{it-12} + \gamma_2 LogIncomeNeighbor_{it-13} + \gamma_3 LogIncome_{it-13} + \gamma_4 CrimeRate_{i,t-13} + v_{it} \tag{3}$$

We then use the fitted values of *CrimeRate* derived in the first stage and run our Second Stage instrumental variable regression for the change in mean income levels, again, running with and without CISP and time fixed effects idiosyncratic differences across CISPs and time trends related to crime.

$$\begin{aligned} LogIncome_{it} = \beta_0 + \beta_1 Crime\widehat{Rate}_{it-s} + \beta_2 LogIncomeNeighbor_{it-1-s} \\ + \beta_3 LogIncome_{it-1-s} + \beta_4 CrimeRate_{it-1-s} + u_{it} \end{aligned} \quad (4)$$

3.4 Limitations

We first want to recognise that, given the inherent limitations of our data, we focus on showing the descriptive relationship between the implementation of the UPPs and their effect on crime, and the economic consequences of the change in crime rates. In their current database, the ISP has only collected information on the crime statistics and population of the CISPs, and does not include other CISP level data, such as law enforcement spending, unemployment, schooling, and other demographic data. The literature on crime suggests that these factors are importantly correlated with both income and crime rates. As the CISP regions were constructed and are monitored by the ISP, it is difficult to find CISP level data from other sources. We attempt to estimate these descriptive relationships by utilising fixed effects for the CISPs as well as time fixed effects. We also construct controls from the data available, but understand that there may still be bias in our estimators. We utilise different regression models below to demonstrate correlation that of the implementation of UPPs has had on different types of violent crime rates in Rio de Janeiro, and how changes in the violent crime rate correspond with changes in night lights, and thus income. We hope that the results of our analysis set the foundation for future causal analysis on the effects of crime on economic development within Rio de Janeiro

4 Estimated Results and Discussion

4.1 Homicide Statistics

In this section we focus on the descriptive statistics for homicides and show regression results utilising the homicide rate in the following section (Additional regression tables showing the results for regression income change on thefts and robberies resulting in deaths,

deaths caused by police officers, or rape rates can be found in the Appendix). We highlight our results decided to focus on homicide as the variable of interest for crime and violence as homicide numbers are thought to be the least concealable due to the nature of the crime involving corpses (Fajnzylber et al., 2002), hence less likely to be prone to underreporting in comparison with other crime types such as thefts and robberies.

Figure 10 illustrates that the homicide counts exhibit a negative trend overall in Rio de Janeiro from 2003 through 2015. The steepest decline is attributed to 2009-2012 which lies within the period where most UPPs are implemented, stipulating that UPPs have an effect on the decreasing homicide levels. However, there is important to note that before the pacification, there already exists a decline in overall homicide levels since 2003. Additionally, the declining trend had mostly subsided by 2014, after the last UPP had been installed, as well as a rebound in homicide rates after 2015. This establishes our doubt in the sustainability of the effect of UPP programmes upon crime rates.

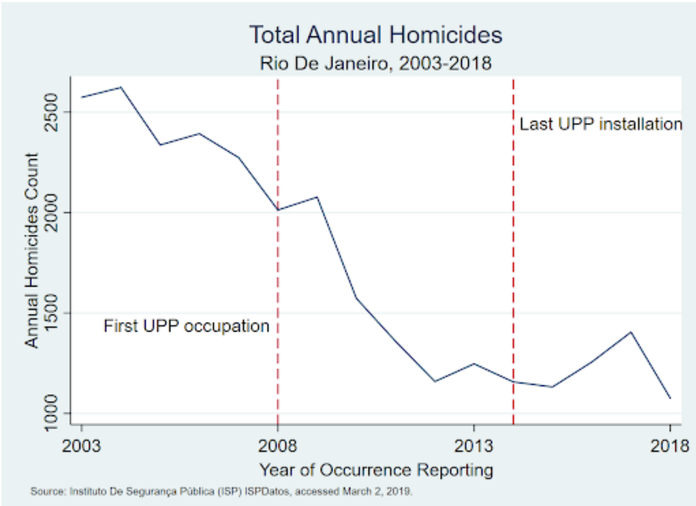


Figure 10: Total Annual Homicides in Rio de Janeiro, from 2003-2018

We then separated Rio de Janeiro into treated and untreated CISPs (with treated containing at least 1 UPP-implemented area) in order to look for any apparent differences between areas that contained UPPs and areas that doesn't. From Figure 11 and 12, the untreated CISPs are generally more violent and volatile compared to treated areas. Looking at homicide shares over time in Figure 11, untreated CISPs have around 1.25-2 times more homicides count relative to treated CISPs, consistent over the whole period. However, homicides count might not be the best measure for comparison as it can be misrepresentative. Figure 12 therefore shows the average homicide rates for every 100,000 population in specific months, again categorised by treated and untreated CISPs. Figure 12 not only suggests higher homicide rates in untreated CISPs, consistent with Figure 11; but also indicate the existence of seasonality through a clear variation of homicide rates across specific months: in December, mean homicides rate per 100,000 people is 5.61 compared to 2.51 in August.

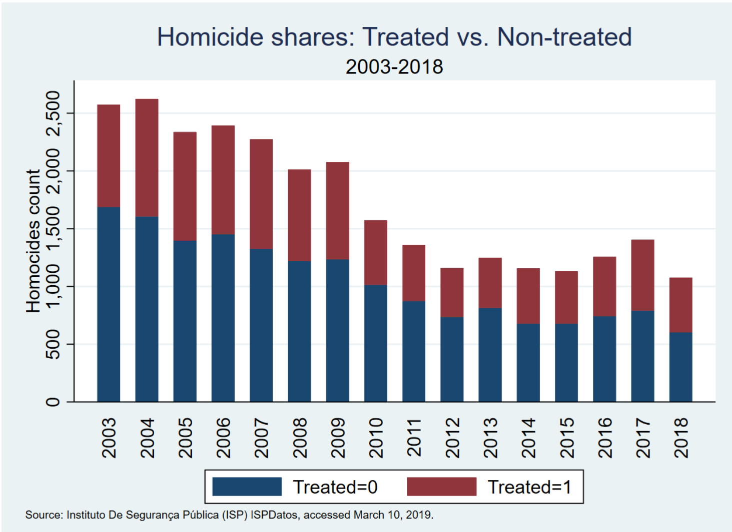


Figure 11: Homicide shares in Rio de Janeiro, in CISPs with UPP versus CISPs without UPP

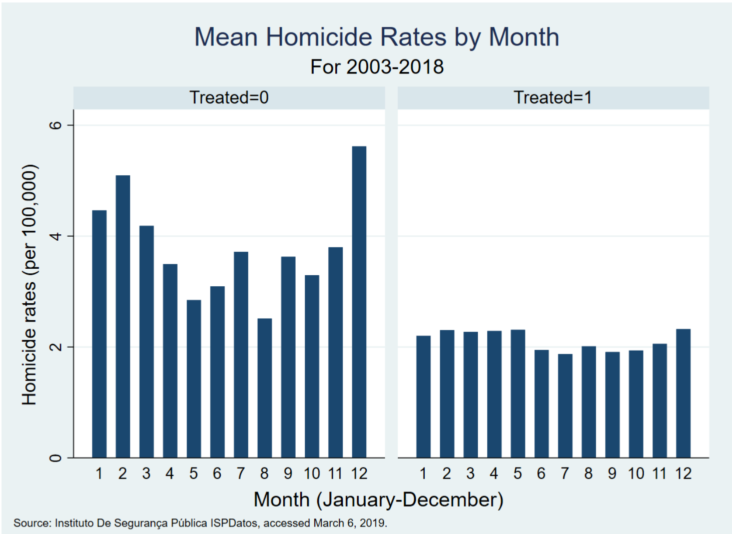


Figure 12: Mean Homicide Rates by Month

To further examine the nature of homicides in Rio de Janeiro on a more granular level, Figure 13 shows a heat map/table for the average annual homicide rate for each CISP from 2003-2018. This allows us to portray how homicide rates are distributed across districts and over time. Looking at Figure 13, we find that the distribution of homicides rates is not uniform: it is higher in certain areas or point of time. It is mostly concentrated in CISP 1, 4, 6, 29 and 39 with these areas having homicide rates over 90 per 100,000 people. The concentration is also mostly skewed towards pre-2008, signifying an overall decrease in homicides rate across all CISPs in question.

Rio De Janiero Homicide Rate by Integrated Circumscriptions of Public Security

2003–2018

CISP[2][3]	Homicide Rate[1] 2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
1	129.0	512.3	381.3	189.2	688.5	372.6	739.6	306.0	243.2	302.4	120.3	119.8	59.7	59.4	118.1	117.3
4	93.7	**99.6**	**161.4**	**127.5**	**84.4**	**80.5**	**108.7**	**63.5**	**41.0**	**40.8**	**37.4**	**49.7**	**52.6**	**52.4**	**52.0**	**57.8**
5	50.3	47.2	38.5	38.2	48.8	29.6	16.0	39.8	13.2	13.1	39.1	38.9	18.1	28.3	23.0	12.7
6	109.8	**125.8**	**87.5**	**80.5**	**73.5**	**66.7**	**64.9**	**42.1**	**22.2**	**13.5**	**14.6**	**21.8**	**41.1**	**36.1**	**20.3**	**32.1**
7	46.4	**53.7**	**38.1**	**20.2**	**40.0**	**17.4**	**34.5**	**19.6**	**21.9**	**12.1**	**4.8**	**9.6**	**26.2**	**23.8**	**9.4**	**28.1**
9	6.2	**16.9**	**13.7**	**15.1**	**9.0**	**7.4**	**15.5**	**5.1**	**8.7**	**4.3**	**8.6**	**2.9**	**4.3**	**3.6**	**5.0**	**2.8**
10	11.2	**14.2**	**2.0**	**13.0**	**10.9**	**7.9**	**7.8**	**2.9**	**3.9**	**1.0**	**3.8**	**4.7**	**2.8**	**3.8**	**16.8**	**13.9**
11\15	2.1	**21.2**	**25.2**	**17.4**	**13.1**	**16.4**	**7.5**	**10.8**	**6.7**	**12.7**	**6.5**	**14.5**	**17.1**	**11.8**	**31.2**	**32.3**
12	5.8	**12.6**	**6.8**	**3.4**	**6.7**	**1.1**	**6.6**	**4.4**	**3.3**	**3.2**	**8.6**	**3.2**	**5.3**	**14.8**	**8.4**	**5.2**
13	7.6	**7.5**	**19.4**	**11.9**	**8.8**	**2.9**	**5.8**	**2.9**	**0.0**	**2.8**	**1.4**	**4.2**	**0.0**	**0.0**	**8.3**	**6.9**
14	9.5	**8.3**	**5.9**	**9.3**	**10.4**	**5.7**	**6.8**	**3.4**	**3.4**	**1.1**	**1.1**	**3.3**	**3.3**	**6.6**	**4.3**	**2.2**
16/42	23.2	17.2	15.7	17.3	12.8	14.7	9.3	23.0	18.9	14.9	26.5	28.3	20.5	30.0	35.5	31.5
17	32.8	**58.7**	**42.7**	**55.2**	**28.0**	**25.3**	**21.3**	**27.4**	**24.8**	**11.1**	**26.9**	**24.4**	**30.4**	**25.4**	**42.1**	**10.7**
18	18.7	24.7	26.1	19.8	21.2	19.5	14.9	10.3	2.9	4.4	5.8	5.8	10.1	2.9	14.3	9.9
19	13.6	**13.5**	**17.2**	**20.0**	**13.2**	**14.6**	**10.9**	**5.7**	**7.8**	**2.8**	**5.6**	**6.3**	**2.8**	**8.4**	**11.1**	**6.2**
20	16.1	**12.8**	**17.1**	**24.5**	**34.9**	**24.1**	**16.0**	**6.1**	**6.1**	**4.8**	**6.6**	**5.4**	**5.9**	**11.3**	**11.2**	**5.8**
21	73.3	**82.2**	**53.5**	**42.6**	**44.2**	**34.3**	**44.7**	**30.4**	**28.4**	**21.9**	**17.7**	**24.6**	**18.0**	**27.0**	**41.7**	**22.3**
22/45	23.3	**23.6**	**30.8**	**27.5**	**42.7**	**31.8**	**29.8**	**23.4**	**13.7**	**12.8**	**15.1**	**35.1**	**27.7**	**43.8**	**54.8**	**28.8**
23	38.9	18.2	25.9	14.5	12.2	24.2	28.4	18.4	15.1	12.9	8.5	6.4	4.2	11.6	18.8	12.0
24	71.8	50.7	33.2	40.4	35.9	38.7	41.6	23.7	34.8	27.5	30.4	20.2	19.1	20.0	19.0	19.9
25	44.7	**33.8**	**28.8**	**44.4**	**40.2**	**35.2**	**36.5**	**25.4**	**16.1**	**23.6**	**12.9**	**18.8**	**18.8**	**17.2**	**30.5**	**22.9**
26	0.0	**0.0**	**0.9**	**23.9**	**31.6**	**21.8**	**16.4**	**15.5**	**15.4**	**12.7**	**11.8**	**6.7**	**13.4**	**8.4**	**7.9**	**10.5**
27	45.5	36.9	60.2	47.2	47.4	40.1	52.5	32.7	26.2	28.1	36.2	20.6	20.0	22.5	23.9	15.1
28	49.6	40.7	41.0	32.3	20.2	13.6	21.7	20.3	20.8	23.5	28.6	22.2	24.9	28.8	14.5	25.1
29	95.7	**91.9**	**86.3**	**83.7**	**94.7**	**89.2**	**87.5**	**86.0**	**55.0**	**34.9**	**44.1**	**70.1**	**40.1**	**32.5**	**38.1**	**27.9**
30	62.1	54.9	55.3	36.9	26.8	38.7	36.1	30.2	37.9	19.7	18.8	28.0	20.2	13.9	24.4	21.0
31	6.0	0.0	26.3	39.7	31.0	23.1	4.5	1.9	38.9	36.2	29.2	20.4	24.0	31.3	38.4	23.6
32	22.3	**24.4**	**27.0**	**27.7**	**24.1**	**23.4**	**21.3**	**12.2**	**11.3**	**11.8**	**15.7**	**22.0**	**23.0**	**21.9**	**17.3**	**16.9**
33	50.4	**59.9**	**59.9**	**50.5**	**56.0**	**48.1**	**54.3**	**23.5**	**37.6**	**29.7**	**25.5**	**23.4**	**22.9**	**26.0**	**28.2**	**23.3**
34	44.4	57.0	45.6	56.4	37.8	37.3	27.3	22.4	29.7	19.9	24.4	19.9	21.7	23.2	23.2	20.8
35/43	70.0	88.9	94.3	98.8	93.2	78.6	89.0	70.8	51.2	38.0	27.7	27.5	26.9	33.8	32.8	19.5
36	75.1	79.9	60.6	76.7	75.8	71.6	73.3	60.6	50.7	37.7	50.1	41.1	35.7	34.9	30.3	23.7
37	33.3	39.4	24.0	28.6	18.3	16.2	17.1	14.1	10.3	10.7	13.9	11.5	11.5	10.1	20.4	12.6
38	33.4	**37.6**	**38.8**	**42.5**	**34.3**	**25.8**	**35.2**	**20.1**	**15.2**	**25.1**	**22.1**	**13.6**	**15.9**	**17.7**	**22.2**	**19.3**
39	177.0	160.3	85.4	74.2	84.1	94.4	110.9	91.9	38.9	48.8	55.2	52.8	48.7	55.1	59.7	41.9
40	72.8	51.6	51.2	75.8	52.3	40.0	44.3	43.1	37.4	35.5	38.8	27.2	28.9	27.9	37.2	31.8
41	25.8	38.1	39.2	24.5	25.7	14.2	12.0	14.0	8.3	7.6	10.3	11.0	8.2	10.2	11.2	5.1
44	65.8	43.3	38.5	27.8	57.8	36.8	39.9	21.1	26.0	20.0	29.0	15.7	32.9	30.3	24.4	21.0

Source: Instituto de Segurança Pública, (ISP) ISPDatos, accessed January, 30, 2019.

Note:
[1] Homicides per 100,000
[2] Integrated Circumscriptions of Public Security (CISP) constitute the sphere of territorial integration, at an operational level, of the Integrated Military Police Companies of the State of Rio de Janeiro (PMERJ). The basic principle of CISP is the concept that the responsibility for policing a subarea of the Integrated Military Police Company should, whenever possible, coincide with the circumscription of a police station. Thus, the CISP correspond to the territorial areas of action and joint responsibility of the Integrated Companies and Police Stations.
[3] Includes all Municipalities in the Municipality of Rio De Janiero. Some original CISP boundaries have been split up over time. In these cases, homicides and populations are aggregated for the original CISP. Bold indicates that a UPP was installed at some point since 2008. Bold indicates a UPP was established in this CISP.

Figure 13: Homicide Rate by CISP, 2003 – 2018

4.2 Regression Results

Table 2 presents the results of the regression between *HomicideRate* on *PostUPP* lagged by 1, 6, and 12 months (contemporaneous effects can be seen in our primary 2SLS regression in Table 4) with the control variables. Though we lose statistical significance when applying our time and entity fixed effects, we can observe that within the first 6 months of occupation, homicide rates fall, but after a year, we see that UPP implementation is corresponds with an increase in homicide rates. Across all iterations of our regressions, we observe a small R-squared value, which is indicative of the limitations that we outlined in Section 3.3.1, that, given the ISP has created these specific CISP boundaries and focuses primarily on collecting crime activity, no other Rio de Janeiro government agencies provide the demographic and financial data that are correlated with crime and UPP implementation. So, we use this table primarily to observe trends, but hesitate to make strong causal inferences. Table 3 shows our OLS regressions of the change in income on homicide rates and our control variables, this time omitting the lagged homicide rate to avoid multicollinearity concerns. Focusing on the fixed effect models in columns 4 through 6, we see some interesting relations between the homicide rate and changes in income. We observe that, initially, decrease the homicide rate corresponds with increase in income. However, in the long run, we now see a positive statistically significant relationship between the homicide rate and change in income. However, we have not yet controlled for possible reverse causality issues between the homicide rate and mean log income, so, we turn to the two stage least squares instrumental variable regression.

Table 2: OLS Regressions of Homicide Rates on UPP Occupation

Panel A	Dependent Variable is $HomicideRate_t$					
	Pooled OLS	CISP & Time FE	Pooled OLS	CISP & Time FE	Pooled OLS	CISP & Time FE
	(1)	(2)	(3)	(4)	(5)	(6)
$PostUPP_{t-1}$	-0.589^{**}	-0.23				
	(0.21)	(0.1)				
$PostUPP_{t-6}$			-0.557^{*}	-0.113		
			(0.23)	(-0.12)		
$PostUPP_{t-12}$					-0.407	0.118
					(0.22)	(0.12)
$HomcideRate_{t-2}$	0.078	-0.031^{**}				
	(0.07)	(0.01)				
$HomcideRate_{t-7}$			0.159	0.061^{***}		
			(0.11)	(0.01)		
$HomcideRate_{t-13}$					0.082	-0.13^{*}
					(0.08)	(0.00)
$LogIncome_{t-2}$	0.629^{***}	0.232				
	(0.11)	(0.43)				
$LogIncome_{t-7}$			0.569^{***}	-0.666		
			(0.13)	(0.78)		
$LogIncome_{t-13}$					0.644^{***}	-0.684
					(0.12)	(1.00)
$LogIncomeNeighbor_{t-2}$	-1.162^{***}	-0.873				
	(0.3)	(0.82)				
$LogIncomeNeighbor_{t-7}$			-0.782	0.893		
			(0.41)	(1.01)		
$LogIncomeNeighbor_{t-13}$					-1.201^{***}	0.092
					(0.35)	(1.00)
Constant	3.078^{***}	3.823^{**}	1.989	1.755	3.033^{***}	4.336^{***}
	(0.78)	(1.14)	(1.07)	(0.97)	(0.81)	(1.17)
Observations	2850	2850	2660	2660	2432	2432
CISP & Time FE	No	Yes	No	Yes	No	Yes
F-Stat	63.3	6.9	69.8	12.7	52.1	5.36
R^2	0.0245	0.050	0.040	0.004	0.024	0.145

Robust SE in Parentheses

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table 3: OLS Regressions of Log Income on Homicide Rates

	Dependent Variable is $LogIncome_t$					
	Pooled OLS	Pooled OLS	Pooled OLS	CISP & Time FE	CISP & Time FE	CISP & Time FE
	(1)	(2)	(3)	(4)	(5)	(6)
$HomicideRate_{t-1}$	-0.000			-0.001^{**}		
	(0.00)			(0.00)		
$HomcideRate_{t-6}$		-0.000			-0.001	
		(0.00)			(0.00)	
$HomcideRate_{t-12}$			0.002			0.002^{*}
			(0.00)			(0.00)
$LogIncome_{t-2}$	0.911^{***}			0.345^{*}		
	(0.01)			(0.14)		
$LogIncome_{t-7}$		0.885^{***}			0.219^{*}	
		(0.01)			(0.09)	
$LogIncome_{t-13}$			0.873^{***}			0.136^{***}
			(0.01)			(0.03)
$LogIncomeNeighbor_{t-2}$	-0.186^{***}			-0.168		
	(0.02)			(0.13)		
$LogIncomeNeighbor_{t-7}$		-0.236^{***}			-0.226^{**}	
		(0.02)			(0.07)	
$LogIncomeNeighbor_{t-13}$			-0.192^{***}			-0.008
			(0.02)			(0.03)
Constant	0.950^{***}	1.219^{***}	1.130^{***}	3.227^{***}	3.93^{***}	3.586^{***}
	(0.06)	(0.07)	(0.07)	(0.16)	(0.18)	(0.05)
Observations	2850	2660	2432	2850	2660	2432
Time Fixed Effects	No	No	No	Yes	Yes	Yes
CISP Fixed Effects	No	No	No	Yes	Yes	Yes
R^2	0.822	0.785	0.795	0.078	0.027	0.022

Robust SE in Parentheses

$^{***}p < 0.01,\ ^{**}p < 0.05,\ ^{*}p < 0.10$

Table 4: 2SLS and IV Regressions of Log Income on Homicide Rates

	2SLS (1)	2SLS CISP & Time FE (2)	2SLS (3)	2SLS CISP & Time FE (4)	2SLS (5)	2SLS CISP & Time FE (6)
Panel A	Dependent Variable is $LogIncome_t$					
$HomicideRate_{t-1}$	-0.045^{*} (0.02)	-0.033 (0.06)				
$HomcideRate_{t-6}$			-0.061^{*} (0.02)	-0.031 (0.08)		
$HomcideRate_{t-12}$					-0.042^{*} (0.02)	0.101^{*} (0.05)
$LogIncome_{t-2}$	0.941^{***} (0.02)	0.296 (0.21)				
$LogIncome_{t-7}$			0.927^{***} (0.02)	0.192 (0.16)		
$LogIncome_{t-13}$					0.903^{***} (0.02)	0.337^{*} (0.17)
$LogIncomeNeighbor_{t-2}$	-0.276^{***} (0.04)	-0.166 (0.16)				
$LogIncomeNeighbor_{t-7}$			-0.359^{***} (0.06)	-0.250^{**} (0.09)		
$LogIncomeNeighbor_{t-13}$					-0.281^{***} (0.05)	-0.104 (0.08)
Constant	1.192^{***} (0.13)	3.670^{***} (0.48)	1.542^{***} (0.17)	4.626^{***} (0.61)	1.365^{***} (0.14)	4.272^{***} (0.62)
Panel B	First Stage $HomRate_{t-1}$		First Stage $HomRate_{t-6}$		First Stage $HomRate_{t-12}$	
$PostUPP_{t-1}$	-0.562^{**} (0.21)	-0.123 (0.07)				
$PostUPP_{t-6}$			-0.564^{**} (0.21)	-0.157 (0.09)		
$PostUPP_{t-12}$					-0.624^{**} (0.22)	-0.171^{*} (0.07)
$HomcideRate_{t-2}$	0.154 (0.1)	0.052^{***} (0.01)				
$HomcideRate_{t-7}$			0.093 (0.08)	-0.003 (0.01)		
$HomcideRate_{t-13}$					0.091 (0.08)	-0.013 (0.01)
$LogIncome_{t-2}$	0.398^{**} (0.15)	-1.682 (1.46)				
$LogIncome_{t-7}$			0.47^{***} (0.14)	-1.244 (1.09)		
$LogIncome_{t-13}$					0.468^{**} (0.14)	-1.689 (1.46)
$LogIncomeNeighbor_{t-2}$	-1.149^{***} (0.43)	0.404 (0.74)				
$LogIncomeNeighbor_{t-7}$			-1.284^{**} (0.41)	0.001 (0.56)		
$LogIncomeNeighbor_{t-13}$					-1.256^{**} (0.44)	0.441 (0.7)
Constant	3.832^{**} (1.34)	7.471 (4.41)	4.048^{**} (1.37)	7.025 (4.1)	4.003^{**} (1.47)	7.505 (4.94)
Observations	2850	2850	2660	2660	2432	2432
CISP & Time FE	No	Yes	No	Yes	No	Yes
Second Stage F-Stat	9.8	18.6	9.8	27.76	10.13	5.29
First Stage F-Stat	58.7	63.4	55.1	2.79	50.57	2.13
R^2 (Second Stage)	0.548	0.826	0.365	0.618	0.575	0.522
R^2 (First Stage)	0.040	0.013	0.0279	0.0386	0.028	0.067

Robust SE in Parentheses
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table 4 above depicts the two stage least squares instrumental variable regression. Utilizing the UPP occupation within a CISP, we hope to control for potential reverse causality issues between the mean nightlights and Homicide Rate. Looking at the fixed effect models in columns 2 and 4, we see that decreases in crime rate, at least initially, correspond with increase in mean income in CISPs. However, as the UPP program was intended to be a paradigm shift from previous *mano dura* policies, that focused only on short term outcomes, and instead long-lasting change in *favelas*, we primarily focus on the 12 month lagged crime rates, to understand the long run income effects that occur over a year after the implementation of the UPP. We find that there is statistically significant decrease to mean income in areas after the implementation of the UPP, which is counter to much of the literature on the relationship between crime and economic development. We provide a hypothesis for a rational behind these statistically significant outcomes below. Additional regression tables for deaths caused by police, rape rate, and thefts and robberies can be found in the Appendix.

5 Conclusion

In this paper we have described the landscape of violence and crime in Brazil through the past two decades and how this landscape changes throughout the *mano dura* approach with an emphasis on the implementation of the *Unidade de Polícia Pacificadora* in Rio de Janeiro. We then examined the potential effects of UPP policy upon homicides level and rates, and utilise this to identify the effect of crime and violence on spatial income distribution for Rio de Janeiro. The key results for this paper revolve around the sustainability of the effect of UPP on crime and the counterintuitive result in the effect of crime deterrence on income, possibly due to the confounding factor of services and income provided by drug gangs restrained by the pacification programme.

First, we confirmed the relationship between the decline in crime rates to having the UPPs implemented. The UPP was created with the intent of providing long term law enforcement and social services to counter the violence in Rio's *favelas*. Our results, however, suggested that while the effect of implementation of UPPs is negatively significant on homicide rates of CISPs at 6 months, after 12 months it still is negative but became insignificant.This could indicate that, for the contemporaneous and short-run effects, the UPPs have been successful in deterring crime and improving general security, though we have been unable to find evidence suggesting that this can be sustained in the long-run. Additionally, we recognised the differences in volume and seasonality of homicides between CISPs that contained UPPs and CISPs that regions that tended to have more violent crime were not the ones that received a UPP, suggesting that goals, such as performing a quick clean-up for the Olympics and World Cup may have been the primary motivation, over a long-term plan for protecting vulnerable communities.

Secondly, we estimated a positive correlation between the change in mean incomes, through our proxy of mean nightlights, and the local homicide rate, our proxy for overall violent crime. Our results indicated that mean income decreases as crime is deterred in an area, which is inconsistent in the crime economics literature (Bourguignon, 2000 is one of many examples) and the World Bank report (World Bank, 2007) which associates crime to have a negative effect on income and economic growth. A possible explanation for this could be that, before UPP intervention, government agencies typically underserviced *favelas* with education, health, and other public goods (Logan, 2015, Perlman, 2010) and there tends to be an absence of state presence (Soares, 2005). Instead, those living in *favelas* stole utilities such as electricity and water (Pilo', 2017), and there are even cases where drug lords and gang leaders invest profits into running mini-welfare states in the communities in which they operate, providing them with goods such as buses, medicine and food, and event concerts (Glenny, 2015, Faiola, 2002), activities which would be captured by the nightlights data. We hypothesise that the services provision by illegal drug cartels or theft of resources may have been eradicated as a result of the occupation by military and police forces through pacification programme. Though UPP implementation was supposed to be coupled with the timely provision of public services, it would seem that this has been limited to only a few of the occupied areas, or has not matched the services that were by provided by drug lords or illegally obtained.

Further research into this subject should make use of additional sources of data at the CISP level to improve on the biases and inconsistencies that are currently present in our model. Additional, there is great potential to explore the "theory of the second best" concerns that may arise when trying to eradicate crime from areas that have adapted to these conditions of having to provide their own public good illegally, and essentially being ruled by the local powerful gangs, without providing a timely replacement for the way these goods were obtained.

References

Ascher, W. and Mirovitskaya, N. 2012. Economic development patterns and the evolution of violence in Latin America. In Economic Development Strategies and the Evolution of Violence in Latin America (pp. 1-39). Palgrave Macmillan, New York. [Online]. [Accessed March 2019]. Available at: https://link.springer.com/chapter/10.1057/9781137272690˙1

Bannon, I. and Collier, P., 2003. Natural resources and violent conflict: Options and actions. The World Bank. [Accessed March 2019]. Available at: http://documents.worldbank.org/curated/en/578321468762592831/pdf/282450Natural0resources0violent0conflict.pdf.
Becker, G.S., 1968. Crime and punishment: An economic approach. In The economic dimensions of crime (pp. 13-68). Palgrave Macmillan, London. [Online]. [Accessed March 2019]. Available at: https://www.nber.org/chapters/c3625.pdf.

Bourguignon, F., 2000. Crime, violence and inequitable development. In Annual World Bank Conference on Development Economics 1999 (pp. 199-220). Available at: http://documents.worldbank.org/curated/en/448681468741326292/pdf/multi-page.pdf#page=205

Bowater, D. 2015. Rio favela still wracked with fear and violence as Olympics 2016 approaches, The Telegraph, April 4. [Online]. [Accessed March 2019]. Available at: https://www.telegraph.co.uk/news/worldnews/southamerica/brazil/11515531/Rio-favela-still-wracked-with-fear-and-violence-as-Olympics-2016-approaches.html.

Cano, I. 2006. Public security policies in Brazil: attempts to modernize and democratize versus the war on crime. International Journal of Human Rights, No.5, pp.132-149. [Online]. [Accessed January 2019]. Available at: http://www.scielo.br/scielo.php?script=sci˙arttext&pid=S1806-64452006000200007&lng=en&nrm=iso&tlng=en.

Carrington, K. et al., 2018. The Palgrave handbook of criminology and the Global South, Palgrave Macmillan.

Central Intelligence Agency, 2019. The World Factbook: Brazil. [Online]. [Accessed March 2019]. Available at: https://www.cia.gov/library/publications/the-world-factbook/geos/br.html.

Costa, F.J., De Faria, J.S., Iachan, F.S. and Caballero, B., 2018. Homicides and the Age of Criminal Responsibility: A Density Discontinuity Approach. Economía, 19(1), pp.59-92.

[Online]. [Accessed March 2019]. Available at: https://papers.ssrn.com/sol3/papers.cfm ?abstract˙id=2670382.

DATASUS. DATASUS Tecnologia da Informação a Serviço do SUS: Mortalidade - Brasil. Ministério da Saúde. [Online]. [Accessed March 2019]. Available at: http://tabnet.datasus.gov.br/cgi/deftohtm.exe?sim/cnv/obt10uf.def

Detotto, C. & Otranto, E., 2010. Does Crime Affect Economic Growth? Kyklos, 63(3), pp.330–345. [Online]. [Accessed March 2019]. Available at: https://onlinelibrary.wiley.com /doi/epdf/10.1111/j.1467-6435.2010.00477.x.

Domanico, A., Malta M., C., Latkin, 2010. Crack Cocaine—Accelerating Use in Brazil, s.l.: National Institute on Drug Abuse. [Online]. [Accessed March 2019]. Available at: https://www.drugabuse.gov/international/abstracts/crack-cocaine-accelerating-use-in-br azil.

Ehrlich, I., 1973. Participation in illegitimate activities: A theoretical and empirical investigation. Journal of political Economy, 81(3), pp.521-565. [Online]. [Accessed March 2019]. Available at: http://www.jstor.org/stable/1831025.

Elonheimo, H., 2014. Evidence for the crime drop: survey findings from two Finnish cities between 1992 and 2013. Journal of Scandinavian Studies in Criminology and Crime Prevention, 15(2), pp. 209-217. [Online]. [Accessed March 2019]. Available at: https://www.tandfonline.com/doi/abs/10.1080/14043858.2014.939458.

Enamorado, López-Calva & Rodríguez-Castelán, 2014. Crime and growth convergence: Evidence from Mexico. Economics Letters, 125(1), pp.9–13. [Accessed March 2019]. Available at: https://www.sciencedirect.com/science/article/pii/S0165176514002900.

Faiola, A., 2002. Brazil's Benevolent Drug Lords. The Washington Post, September 19. [Online]. [Accessed March 2019]. Available at: https://www.washingtonpost.com/archive/politics/2002/09/19/brazils-benevolent-drug-lords/46de7df8-d1ac-4371-af65-396e8daa65f4 /?noredirect=on&utm˙term=.9bfb2baf78a7.

Fajnzylber, P., Lederman, D. & Loayza, N., 2002. Inequality and Violent Crime. The Journal of Law & Economics, 45(1), pp.1–39. [Accessed March 2019]. Available at: https://www.journals.uchicago.edu/doi/10.1086/338347.

Fajnzylber, P., Lederman, D., and Loayza, N., 2002. What causes violent crime? European Economic Review, 46(7):1323–1357. [Online]. [Accessed March 2019]. Available at: https://siteresources.worldbank.org/DEC/Resources/What˙Causes˙Crime.pdf

Farrell, G., Tseloni, A., Mailley, J. and Tilley, N., 2011. The crime drop and the security hypothesis. Journal of Research in Crime and Delinquency, 48(2), pp.147-175. [Online]. [Accessed March 2019]. Available at: https://journals.sagepub.com/doi/10.1177/0022427810391539.

Fleisher, B.M., 1966. The effect of income on delinquency. The American Economic Review, 56(1/2), pp.118-137. [Online]. [Accessed March 2019]. Available at: http://www.jstor.org/stable/1821199.

Forte, J., 2014, Residents Tell of Abuse by Pacification Force in Maré, The Rio Times, November 6. [Online]. [Accessed March 2019]. Available at: https://riotimesonline.com/brazil-news/rio-politics/residents-tell-of-abuse-by-pacification-force-in-complexo-da-mare/.

Forte, J., 2015, Pacification Force to Stay in Complexo da Maré 6 More Months, The Rio Times, January 2. [Online]. [Accessed March 2019]. Available at: https://riotimesonline.com/brazil-news/rio-politics/pacification-force-to-stay-in-complexo-da-mare-6-more-months/.

Gaviria, A., 2002. Assessing the effects of corruption and crime on firm performance: Evidence from Latin America. Emerging Markets Review, 3(3), pp.245–268. [Online]. [Accessed March 2019]. Available at: https://core.ac.uk/download/pdf/6266579.pdf.

Glenny, M. 2015. The day I met Rio's favela master: the drug lord who championed the poor. The Guardian, 2015. [Online]. [Accessed March 2019]. Available at: https://www.theguardian.com/world/2015/sep/13/rio-favel-master-nem-of-rocinha-drug-lord-champion-of-the-poor.

Government of the State of Rio de Janeiro, 1999. Lei 3329 99; Lei — no 3329 de 28 de dezembro de 1999. Jusbrasil. [Online]. [Accessed March 2019]. Available at: https://gov-rj.jusbrasil.combr/legislacao/22810/lei-3329-99

Government of the State of Rio de Janeiro, 2011. Lei 5890 11 I Lei n 5890, de 14 de janeiro de 2011. Jusbrasil. [Online]. [Accessed March 2019]. Available at: https://gov-rj.jusbrasil.com.br/legislacao/1026543/lei-5890-11#art-2.

Henderson, J.V., Storeygard, A. and Weil, D.N., 2012. Measuring economic growth from outer space. American economic review, 102(2), pp.994-1028. [Online]. [Accessed March 2019]. Available at: http://www.jstor.org/stable/23245442.

Islam, A., 2014. Economic growth and crime against small and medium sized enterprises in developing economies. Small Business Economics, 43(3), pp.677–695. [Online]. [Accessed March 2019]. Available at: http://documents.worldbank.org/curated/en/27316146818172 7530/pdf/WPS6768.pdf.

ISPDados. Open Data Sets. Instituto de Segurança Pública. [Online]. [Accessed March 2019]. Available at: http://www.ispdados.rj.gov.br/.

Janssen, A., 2015. The pacification of favelas in Rio de Janeiro. Radboud University. [Online]. [Accessed March 2019]. Available at: https://theses.ubn.ru.nl/bitstream/handle/123 456789/3956/Janssen

Kingstone, P., 2012. Economic exclusion and the shifting patterns of violence in Argentina and Brazil. In Economic Development Strategies and the Evolution of Violence in Latin America (pp. 213-250).

Palgrave Macmillan, New York. [Online]. [Accessed March 2019]. Available at: https://link .springer.com/chapter/10.1057/9781137272690˙8.

Leeds, E., 2007. Serving States and Serving Citizens: Halting Steps toward Police Reform in Brazil and Implications for Donor Intervention, Policing & Society, 17:1, pp.21-37. [Online]. [Accessed March 2019]. Available at: https://www.tandfonline.com/doi/pdf/10.108 0/10439460601124122?needAccess=true.

Logan, A., 2015. Fixing Favelas: Urban Housing Problems in Brazil. The Borgen Project. [Online]. [Accessed March 2019]. Available at: https://borgenproject.org/fixing-favelas-urban-housing-problems-brazil/.

Marquez-Padilla, F., Perez-Arce, F., and Rodriguez-Castelan, C., 2015. The (Non-) Effect of Violence on Education: Evidence from the "War on Drugs" in Mexico. Santa Monica, CA: RAND Corporation. [Online]. [Accessed March 2019]. Available at: https://www.rand .org/pubs/working˙papers/WR1082.html.

Miraglia, P., 2015. Drugs and drug trafficking in Brazil: Trends and policies. Center for 21st Century Security and Intelligence Latin America Initiative. [Online]. [Accessed March 2019]. Available at: https://www.brookings.edu/wp-content/uploads/2016/07/Miraglia-

Brazil-final.pdf.

Muggah, R., Garzon, J.C. and Suarez, M., 2018. Mano Dura: the costs and benefits of repressive criminal justice for young people in Latin America. Igarape Institute. In Youth, Peace and Security, pp.1-38. [Online]. [Accessed December 2018]. Available at: https://igarape.org.br/wp-content/uploads/2018/04/Mano-Dura-The-costs-and-benefits-of-repressive-criminal-justice-for-young-people-in-Latin-America.pdf.

NOAA. Version 1 VIIRS Day/Night Band Nighttime Lights. NOAA/NCEI. [Online]. [Accessed March 2019]. Available at: https://ngdc.noaa.gov/eog/viirs/download˙dnb˙composites.html.

NOAA. Version 4 DMSP-OLS Nighttime Lights Time Series. NOAA/NCEI. [Online]. [Accessed March 2019]. Available at: https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html.

Numbeo, 2019. Global Crime Index. [Online]. [Accessed March 2019]. Available at: https://www.numbeo.com/crime/rankings˙by˙country.jsp.

Perlman, J., 2010. Favela: Four decades of living on the edge in Rio de Janeiro. Oxford University Press. [Book].

Pilo', F., 2017. A socio-technical perspective to the right to the city: Regularizing electricity access in Rio de Janeiro's Favelas. International Journal of Urban and Regional Research, 41(3), pp.396-413. [Online]. [Accessed March 2019]. Available at: https://onlinelibrary.wiley.com/doi/full/10.1111/1468-2427.12489.

Pshiva, Rony, and Gustavo A. Suarez. 2010. "Capital Crimes: Kidnappings and Corporate Investment in Colombia" in The Economics of Crime: Lessons for and from Latin America, edited by Rafael Di Tella,

Sebastian Edwards and Ernesto Schargrodky, 63–100. Chicago: University of Chicago Press. [Online]. [Accessed March 2019]. Available at: https://ideas.repec.org/h/nbr/nberch/11833.html.

Reichenheim, M.E., De Souza, E.R., Moraes, C.L., de Mello Jorge, M.H.P., Da Silva, C.M.F.P. and de

Souza Minayo, M.C., 2011. Violence and injuries in Brazil: the effect, progress made, and challenges ahead. The Lancet, 377(9781), pp.1962-1975. [Online]. [Accessed March 2019].

Available at: https://www.ncbi.nlm.nih.gov/pubmed/21561649.

Rekow, L., 2016. Pacification & mega-events in Rio de Janeiro: urbanization, public security & accumulation by dispossession. Journal of Human Security, 12(1), pp. 4-34. [Online]. [Accessed March 2019]. Available at: https://ideas.repec.org/a/lib/00johs/v12y2016i1p4-34.html.

Reuters, 2019. Brazil minister proposes tougher laws to stem deadly crime wave. [Online]. [Accessed March 2019]. Available at: https://www.reuters.com/article/us-brazil-politics-crime/brazil-minister-proposes-tougher-laws-to-stem-deadly-crime-wave-idUSKCN1PT20U

Ríos, V., 2014. The Role of Drug-Related Violence and Extortion in Promoting Mex Ican Migration: Unexpected Consequences of a Drug War. Latin American Research Review, 49(3), pp.199–217. [Online]. [Accessed March 2019]. Available at: https://scholar.harvard.edu/files/vrios/files/riosv2014˙larr2014˙securityimmigration1.pdf.

Ríos, V., 2016. The impact of crime and violence on economic sector diversity. [Online]. [Accessed March 2019]. Available at: https://scholar.harvard.edu/files/vrios/files/riosv˙crimesectordiversitydec212015.pdf.

Ross, N. (2013). Crime: how to solve it and why so much of what we're told is wrong. London: Biteback Publishing. [Book]. [Accessed on February 25, 2019.]

Soares F., Soares Y., 2005. The Socio-Economic Impact of Favela-Bairro: What do the Data Say? IADB [Online]. [Accessed March 2019]. Available at: http://200.196.152.249/˙arquivosestaticos/FIS/pdf/the˙socio-economic˙impacto˙of˙favela-bairro˙what˙do˙the˙data˙say.pdf.

Tardáguila, C., 2016. UPP da Maré: da promessa de oito unidades ao adiamento indefinido. Folha de S.Paulo March 21. [Online]. [Accessed March 2019]. Available at: http://cdn.static-economist.com/sites/default/files/imagecache/full-width/images/print-edition/2013/09/M/20130914˙AMM958˙0.png.

The Economist, 2015. The location of UPP's and Olympic zones in Rio de Janeiro [Photograph]. [Online]. [Accessed March 2019]. Available at: http://cdn.static-economist.com/sites/default/files/imagecache/full-width/images/print-edition/2013/09/M/20130914˙AMM958˙0.png.

The Economist, 2017. The world's most dangerous cities. The Economist, March, 31.

[Online]. [Accessed March 2019]. Available at: https://www.economist.com/graphic-detail/2017/03/31/the-worlds-most-dangerous-cities.

The Economist, 2018. Jair Bolsonaro, Latin America's latest menace. The Economist, September 20th. [Online]. [Accessed March 2019]. Available at: https://www.economist.com/leaders/2018/09/20/jair-bolsonaro-latin-americas-latest-menace.

United States Department of State Bureau of Diplomatic Security (OSAC), 2018. Brazil 2018 Crime & Safety Report: São Paulo, OSAC. [Online]. [Accessed March 2019]. Available at: https://www.osac.gov/Pages/ContentReportDetails.aspx?cid=23477.

Waiselfisz, J. J., 2014. Mapa da Violência: Os Jovens Do Brasil. Flacso Brasil. [Online]. [Accessed March 2019]. Available at: https://www.mapadaviolencia.org.br/pdf2014/Mapa2014˙JovensBrasil˙Preliminar.pdf.

Werling, E., 2014. Rio's Pacification: Paradigm shift or paradigm maintenance. Rio de Janeiro: Humanitarian Action in Situations other than War (HASOW). [Online]. [Accessed March 2019]. Available at: https://igarape.org.br/wp-content/uploads/2016/04/Rio%E2%80%99s-Pacification-.pdf.

World Bank, 2006. Brazil - crime, violence, and economic development in Brazil: elements for effective public policy (English). Poverty Reduction and Economic Management Unit: Latin America and the Caribbean Region. Washington, DC: World Bank. Report No. 36525-BR. [Online]. [Accessed January 2019].Available at: http://documents.worldbank.org/curated/en/104001468231894133/Brazil-Crime-violence-and-economic-development-in-Brazil-elements-for-effective-public-policy

World Bank; United Nations Office on Drugs and Crime. 2007. Crime, Violence, and Development : Trends, Costs, and Policy Options in the Caribbean. Washington, DC. © World Bank. Available at: https://openknowledge.worldbank.org/handle/10986/7687 License: CC BY 3.0 IGO.

WorldBank. 2019. GDP, PPP (constant 2011 international $) — Data. [Online]. [Accessed 15 Mar. 2019]. Available at: https://data.worldbank.org/indicator/NY.GDP.MKTP.PP.KD?end=2011&locations=BR&start=1993.

Yishay, Ariel Ben and Sarah Pearlman, "Homicide and Work: The Impact of Mexico's Drug War on Labor Market Participation," 2013. Working paper. [Online]. [Accessed 15 Mar. 2019].

Appendix

Pacifying Police Units by Occupation and Installation Dates

2008–2014

UPP	UPP Number	Located in CISP	Occupation Date	Installation Date
Santa Marta	1	10	11/19/2008	12/19/2008
Cidade de Deus	2	32/41	11/11/2008	2/16/2009
Batam	3	33	7/12/2008	2/18/2009
Chapéu Mangueira / Babilônia	4	12	5/11/2009	6/10/2009
Pavão-Pavãozinho	5	13/14	11/30/2009	12/23/2009
Tabajaras	6	10/12	12/26/2009	1/14/2010
Providência	7	4	3/22/2010	4/26/2010
Borel	8	19	4/28/2010	6/7/2010
Formiga	9	19	4/28/2010	7/1/2010
Andaraí	10	20	6/11/2010	7/28/2010
Salgueiro	11	19	7/30/2010	9/17/2010
Turano	12	6	8/10/2010	10/30/2010
São João / Matriz / Queto	13	25	1/6/2011	1/31/2011
Coroa / Fallet / Fogueteiro	14	6/7	1/6/2011	2/25/2011
Escondidinho / Prazeres	15	6/7	1/6/2011	2/25/2011
São Carlos	16	6	1/6/2011	5/17/2011
Mangueira	17	17	6/19/2011	11/3/2011
Macacos	18	20	10/14/2010	11/30/2011
Vidigal	19	15	12/13/2011	1/18/2012
Nova Brasília	20	45	11/28/2010	4/18/2012
Fazendinha	21	45	11/28/2010	4/18/2012
Adeus / Baiana	22	45	11/28/2010	5/11/2012
Alemão	23	45	11/28/2010	5/30/2012
Chatuba	24	22	6/27/2012	6/27/2012
Fé / Sereno	25	22/38	6/27/2012	6/27/2012
Parque Proletário	26	22	11/28/2010	8/28/2012
Vila Cruzeiro	27	22	11/28/2010	8/28/2012
Rocinha	28	11	12/13/2011	9/20/2012
Jacarezinho	29	25	10/14/2012	1/16/2013
Manguinhos	30	21	10/14/2012	1/16/2013
Barreira do Vasco / Tuiuti	31	17/21	3/3/2013	4/12/2013
Caju	32	17	3/3/2013	4/12/2013
Cerro-Corá	33	9	4/29/2013	6/3/2013
Arará / Mandela	34	21	10/13/2012	9/6/2013
Lins	35	25/26	10/6/2013	12/2/2013
Camarista Méier	36	26	10/6/2013	12/2/2013
Mangueirinha	37	59	8/5/2013	2/7/2014
Vila Kennedy	38	34	3/13/2014	5/23/2014
Maré[1]	39	21	3/30/2014	

Source: Instituto de Segurança Pública, (ISP) ISPDatos, accessed January, 30, 2019; "Arrests made After Occupation of Complexo de Maré," The Rio Times, 4/1/2014, available at https://riotimesonline.com/brazil-news/rio-politics/arrests-made-after-occupation-of-complexo-da-mare-in-rio/; "Rio favela still wracked with fear and violence as Olympics 2016 approaches," 4/1/2015, available at https://www.telegraph.co.uk/news/worldnews/southamerica/brazil/11515531/Rio-favela-still-wracked-with-fear-and-violence-as-Olympics-2016-approaches.html

Note:
[1] Though Maré was once occupied, the ISP no longer lists it as a favela where they have a unit installed. Occupation Date and Installation date were compiled from newspaper articles. The Telegraph has attributed this to Maré's size, territorial divisions, and its violence.

Figure 14: UPP by Occupation and Installation Dates

Table 5: 2SLS and IV Regressions of Log Income on Thefts and Robberies

	2SLS (1)	2SLS CISP & Time FE (2)	2SLS (3)	2SLS CISP & Time FE (4)	2SLS (5)	2SLS CISP & Time FE (6)
Panel A	Dependent Variable is $LogIncome_t$					
$TheftsRobberies_{t-1}$	-0.001^{**}	0.000				
	(0.00)	(0.00)				
$TheftsRobberies_{t-6}$			-0.001^{***}	0.000		
			(0.00)	(0.00)		
$TheftsRobberies_{t-12}$					-0.001^{**}	-0.002
					(0.00)	(0.00)
$LogIncome_{t-2}$	0.927^{***}	0.357^{*}				
	(0.01)	(0.14)				
$LogIncome_{t-7}$			0.904^{***}	0.239^{*}		
			(0.01)	(0.10)		
$LogIncome_{t-13}$					0.888^{***}	0.114
					(0.01)	(0.10)
$LogIncomeNeighbor_{t-2}$	-0.191^{***}	-0.197				
	(0.02)	(0.13)				
$LogIncomeNeighbor_{t-7}$			-0.243^{***}	-0.273^{**}		
			(0.02)	(0.10)		
$LogIncomeNeighbor_{t-13}$					-0.198^{***}	0.081
					(0.02)	(0.25)
Constant	1.119^{***}	3.775^{***}	1.435^{***}	4.819^{***}	1.292^{***}	2.498
	(0.09)	(0.72)	(0.10)	(1.16)	(0.10)	(4.22)
Panel B	First Stage $T\&R_{t-1}$		First Stage $T\&R_{t-6}$		First Stage $T\&R_{t-12}$	
$PostUPP_{t-1}$	-11.197^{*}	61.527^{*}				
	(4.94)	(24.61)				
$PostUPP_{t-6}$			-12.468^{*}	59.045^{*}		
			(5.22)	(22.68)		
$PostUPP_{t-12}$					-13.382^{*}	51.896^{*}
					(5.51)	(22,75)
$TheftsRobberies_{t-2}$	0.81^{***}	0.432^{***}				
	(0.03)	(0.08)				
$TheftsRobberies_{t-7}$			0.802^{***}	0.414^{***}		
			(0.04)	(0.09)		
$TheftsRobberies_{t-13}$					0.802^{***}	0.39^{***}
					(0.04)	(0.08)
$LogIncome_{t-2}$	1.739	-5.536				
	(4.18)	(16.49)				
$LogIncome_{t-7}$			1.668	-5.337		
			(4.38)	(16.59)		
$LogIncome_{t-13}$					1.645	-5.066
					(4.61)	(16.52)
$LogIncomeNeighbor_{t-2}$	14.274	12.614				
	(10.86)	(18.61)				
$LogIncomeNeighbor_{t-7}$			14.583	12.282		
			(11.27)	(18.42)		
$LogIncomeNeighbor_{t-13}$					16.743	25.165
					(11.74)	(17.87)
Constant	31.896	184.335	35.205	192.26^{***}	28.289	160.781^{***}
	(38.15)	(44.27)	(39.62)	(45.03)	(41.53)	(42.1)
Observations	2850	2850	2660	2660	2432	2432
CISP & Time FE	No	Yes	No	Yes	No	Yes
Second Stage F-Stat	36.81	32.56	34.982	43.82	33.50	12.07
First Stage F-Stat	182.24	12.40	159.18	11.74	128.29	13.31
R^2 (Second Stage)	0.752	0.857	0.667	0.633	0.729	0.0058
R^2 (First Stage)	0.660	0.838	0.646	0.836	0.635	0.836

Robust SE in Parentheses

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table 6: 2SLS and IV Regressions of Log Income on Death by Police

	2SLS (1)	2SLS CISP & Time FE (2)	2SLS (3)	2SLS CISP & Time FE (4)	2SLS (5)	2SLS CISP & Time FE (6)
Panel A	Dependent Variable is $LogIncome_t$					
$DeathbyPolice_{t-1}$	0.171^{*} (0.07)	-0.041 (0.07)				
$DeathbyPolice_{t-6}$			0.253^{*} (0.11)	-0.035 (0.08)		
$DeathbyPolice_{t-12}$					0.316 (0.22)	0.143 (0.09)
$LogIncome_{t-2}$	0.848^{***} (0.03)	0.336^{*} (0.15)				
$LogIncome_{t-7}$			0.794^{***} (0.04)	0.217^{*} (0.11)		
$LogIncome_{t-13}$					0.761^{***} (0.08)	0.212^{***} (0.03)
$LogIncomeNeighbor_{t-2}$	-0.083 (0.05)	-0.159 (0.15)				
$LogIncomeNeighbor_{t-7}$			-0.091 (0.07)	-0.231^{*} (0.10)		
$LogIncomeNeighbor_{t-13}$					-0.016 (0.13)	-0.131^{**} (0.04)
Constant	0.798^{***} (0.11)	3.318^{***} (0.43)	1.009^{***} (0.15)	4.285^{***} (0.49)	0.88^{***} (0.23)	5.500^{***} (0.39)
Panel B	First Stage $DeathbyP_{t-1}$		First Stage $DeathbyP_{t-6}$		First Stage $DeathbyP_{t-12}$	
$PostUPP_{t-1}$	0.141^{**} (0.05)	0.065 (0.04)				
$PostUPP_{t-6}$			0.121^{*} (0.06)	0.064 (0.04)		
$PostUPP_{t-12}$					0.074 (0.06)	0.049 (0.05)
$DeathbyPolice_{t-2}$	0.187^{***} (0.05)	0.051 (0.04)				
$DeathbyPolice_{t-7}$			0.190^{***} (0.05)	0.052 (0.05)		
$DeathbyPolice_{t-13}$					0.196^{***} (0.05)	0.042 (0.05)
$LogIncome_{t-2}$	0.338^{***} (0.06)	-0.285 (0.26)				
$LogIncome_{t-7}$			0.322^{***} (0.06)	-0.301 (0.27)		
$LogIncome_{t-13}$					0.304^{***} (0.07)	-0.217 (0.17)
$LogIncomeNeighbor_{t-2}$	-0.633^{***} (0.10)	0.306 (0.27)				
$LogIncomeNeighbor_{t-7}$			-0.590^{***} (0.10)	0.334 (0.29)		
$LogIncomeNeighbor_{t-13}$					-0.528^{***} (0.10)	0.345 (0.19)
Constant	0.953^{***} (0.26)	0.672 (0.43)	0.874^{**} (0.27)	0.634 (0.43)	0.756^{**} (0.27)	0.242 (0.38)
Observations	2850	2850	2660	2660	2432	2432
CISP & Time FE	No	Yes	No	Yes	No	Yes
Second Stage F-Stat	12.18	31.79	8.524	43.1	3.143	22.51
First Stage F-Stat	23.93	1.98	21.15	1.73	18.34	1.69
R^2 (Second Stage)	0.6016	0.9460	0.3051	0.7870	0.0674	0.6400
R^2 (First Stage)	0.0651	0.0508	0.0640	0.0528	0.0656	0.0571

Robust SE in Parentheses

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table 7: 2SLS and IV Regressions of Log Income on Rape Rate

	2SLS (1)	2SLS CISP & Time FE (2)	2SLS (3)	2SLS CISP & Time FE (4)	2SLS (5)	2SLS CISP & Time FE (6)
Panel A	Dependent Variable is $LogIncome_t$					
$RapeRate_{t-1}$	-0.021^{**} (0.01)	0.029 (0.04)				
$RapeRate_{t-6}$			-0.027^{**} (0.01)	0.024 (0.05)		
$RapeRate_{t-12}$					-0.022^{*} (0.01)	-0.079 (0.08)
$LogIncome_{t-2}$	0.923^{***} (0.01)	0.329^{*} (0.15)				
$LogIncome_{t-7}$			0.901^{***} (0.01)	0.213 (0.11)		
$LogIncome_{t-13}$					0.887^{***} (0.01)	0.246^{***} (0.07)
$LogIncomeNeighbor_{t-2}$	-0.189^{***} (0.02)	-0.180 (0.14)				
$LogIncomeNeighbor_{t-7}$			-0.240^{***} (0.02)	-0.253^{***} (0.07)		
$LogIncomeNeighbor_{t-13}$					-0.199^{***} (0.02)	-0.081 (0.05)
Constant	0.966^{***} (0.07)	3309^{***} (0.34)	1.231^{***} (0.08)	4.288^{***} (0.46)	1.151^{***} (0.08)	5.493^{***} (0.72)
Panel B	First Stage $RapeRate_{t-1}$		First Stage $RapeRate_{t-6}$		First Stage $RapeRate_{t-12}$	
$PostUPP_{t-1}$	-1.151^{***} (0.26)	-0.121 (0.26)				
$PostUPP_{t-6}$			-1.155^{***} (0.27)	-0.112 (0.25)		
$PostUPP_{t-12}$					-1.114^{***} (0.28)	-0.110 (0.26)
$RapeRate_{t-2}$	0.163 (0.09)	-0.002 (0.00)				
$RapeRate_{t-7}$			0.163 (0.09)	-0.008^{*} (0.00)		
$RapeRate_{t-13}$					0.148 (0.09)	-0.020^{***} (0.00)
$LogIncome_{t-2}$	0.161 (0.12)	0.662 (0.65)				
$LogIncome_{t-7}$			0.193 (0.12)	0.595 (0.59)		
$LogIncome_{t-13}$					0.262^{*} (0.12)	0.819 (0.78)
$LogIncomeNeighbor_{t-2}$	1.051^{*} (0.41)	0.282 (0.42)				
$LogIncomeNeighbor_{t-7}$			1.103^{**} (0.43)	0.449 (0.48)		
$LogIncomeNeighbor_{t-13}$					0.965^{*} (0.44)	0.003 (0.47)
Constant	-1.315 (1.23)	-1.070 (0.43)	-1.595 (1.27)	-1.281 (2.72)	-1.422 (1.34)	-0.758 (2.56)
Observations	2850	2850	2660	2660	2432	2432
CISP & Time FE	No	Yes	No	Yes	No	Yes
Second Stage F-Stat	31.095	17.03	29.127	26.78	23.579	4.50
First Stage F-Stat	11.06	0.63	10.07	1.48	9.54	6.89
R^2 (Second Stage)	0.7394	0.7213	0.6546	0.5383	0.6949	0.1418
R^2 (First Stage)	0.0464	0.0081	0.048	0.0002	0.0413	0.0039

Robust SE in Parentheses
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Anocracy in Ankara: Co-party Bias and the Turkish Housing Development Administration (TOKI), 2002 – 2016

Ethan Caspi[1], Christopher Dann[2], Lutfi Sun[3], Andre Zeromski[4], and Yihan Zhu[2]

[1]University of Southern California, California, USA
[2]The London School of Economics and Political Science, London, UK
[3]Trinity University, Texas, USA
[4]Washington and Lee University, Virginia, USA

Abstract

The 'Core Vote' and 'Swing Vote' are two different hypotheses explaining the distribution of government resources. In this paper, we test the Core Vote (Cox and McCubbins, 1986) versus Swing Vote (Lindbeck and Weibull, 1987; Dixit and Londregan, 1996) hypotheses and develop our theory of which strategy a government would adopt as the country moves from democracy towards autocracy. We develop a formal model to show the relationship between changes in democracy and its impact on which voting block(s) receives a higher distribution of public resources. Using the data of public housing between 2002 and 2016 and four general election results (2002, 2007, 2011, 2015) in Turkey, our OLS specifications show that our estimates of co-party bias follow the Polity IV trend very closely. As such, we find that the less democratic Turkey becomes, the more co-party bias there is. We expand on this initial discovery through the use of fixed effects regressions and further use an instrumental variable estimation as our identification strategy. Our results illustrate that the change of the incumbent Justice and Development Party's (AKP) vote share positively impacts the quantity of public housing built in a province which has a large AKP vote share.

1 Introduction

Following the third victory in the 2011 Turkish General Election for Recep Erdogan and his Justice and Development Party (AK Parti or AKP), there has been a gradual shift from democracy to autocracy. This shift is illustrated in events such as the recent constitutional referendum that strengthened the power of the executive branch; this is also captured by the drop in the Polity IV score, from its peak at 9, "Democracy" to the current score of 2, an "Open Anocracy". With that in mind, our research aims to analyse the subsequent effect of this shift toward autocracy. Specifically, we focus on the question of whether the fall in democracy leads to a rise in co-party bias, whereby the government channels more resources to the provinces that endorse the incumbent party.

To test our hypothesis, we use the number of TOKI ("Public Housing Development Administration") houses built in each province between 2002 to 2016 as our measure of the government's resource allocation decision. Firstly, we descriptively show that co-party bias has risen over time by running four baseline OLS regressions, including observable controls, for each election (2002, 2007, 2011, 2015), regressing AKP vote share share against the number of TOKI houses. These preliminary results show that our estimates of co-party bias actually follow the Polity IV trend very closely: the less democratic Turkey becomes, the more co-party bias there is.

Next, we employ panel data techniques; a fixed effects approach with instrumental variable is our main identification strategy, treating education levels as our instrument. We find a statistically significant relationship via our fixed effects specification that AKP vote share positively impacts TOKI houses when using education levels as an instrument, thereby confirming the trend between increasing co-party bias and decreasing democracy.

2 Literature Review

There are two competing models about the decision making behind the distributions of government funds and projects. Gary W. Cox and Mathew D. McCubbins view candidates as risk-averse and argue that the optimal strategy for reelection is to promise redistributions to "their reelection constituency and thereby to maintain existing political coalition" (1986). Lindbeck and Weibull (1987), on the other hand, find equilibrium redistribution to favor nonvoters and opposition supporters. Building on the Lindbeck and Weibull (1987) model, Dixit and Londregan (1996) claim that parties will target voters from different partisan preferences to buy their votes with large enough transfers that outweigh ideological attachments (1996). In this paper, we test the core vote (Cox and McCubbins, 1986) vs. swing vote (Lindbeck and Weibull, 1987; Dixit and Londregan, 1996) hypotheses and develop our theory of which strategy a government would prefer in a more autocratic setting.

Based on existing literature, we theorise that there is more evidence in favour of a core vote hypothesis in less democratic settings. Grau, Olle, and Navarro (2018) show that co-party favouritism is lower in Spanish local governments with competitive elections and significantly higher otherwise. There is evidence that political leaders favor their place of birth in countries with low education levels and weak political institutions (Hodler and Raschky 2014). We also see evidence that the governments of developing nations direct foreign aid funds toward political supporters and favor their strongholds for development projects (Atkinson, Hicken, and Ravanilla 2011, Miguna 2012, Jablonski 2014). Similarly, we can consider the co-ethnic favoritism in Sub-Saharan African countries to be evidence of core vote redistribution because of the accordance between political preferences and ethnic identities. Burgess, Jedwab, and Miguel (2015) show that road building in Kenya was more ethnically driven in times of autocracy than in democracy.

The evidence for the swing vote hypothesis, on the other hand, is stronger in more democratic countries. Ward and John find that marginal legislative constituencies in England received larger central government grants (1999). Dahlberg and Johansson show that the Swedish government directed intergovernmental grants to municipalities with more swing voters (2002). Modifying the Dixit and Londregan (1996) model, Stokes (2005) shows evidence from Argentina that weakly opposed swing voters receive patronage goods instead of loyal voters. In this paper, we want to see whether the Turkish government favors its own core or target constituencies with higher concentrations of swing voters. In Section I, we present our theory of redistributive politics. In Section II, we explain why the trajectory of Turkey from more to less democratic over the past decade makes it an ideal empirical setting for testing our model. In Section III, we test our hypothesis and show the government's strategy has changed over time. In Section IV, we present our conclusions and the questions we did not find an answer to in our study.

3 Theoretical Framework

In order to analyse both core vote and swing vote theories in a single scale, we use co-party bias as our main variable of interest. A positive co-party bias means that the government directs funds to its own voter base. Conversely, a negative co-party bias predicts that the government will favour people from different political ideologies in order to garner their votes. As mentioned above, we see more evidence in favour of the core vote hypothesis in less democratic settings. Therefore, we expect positive co-party bias in an autocracy. In a full autocracy where citizens have no power, however, the government will not need any support from the people and thus will have no incentive to favour certain groups over others. Although Besley and Persson (2011) develop theoretical models surrounding state capacity and government behaviour, especially with regards to the notion of group transfers, political openness, and executive constraints (being three dimensions of political conflict), we develop our own formal model to isolate the effect of swing voting behaviour on executive constraints.

We expect a negative co-party bias in more democratic settings. Since the ideological attachment of individuals is low and people have more say in the decision-making process, the net gains from swinging a voter are higher. In a full democracy, however, citizens will have absolute power, whereby goods and services will be distributed exactly based on their needs, leaving no room for political considerations. Combined with the relationship discussed above, we theorise that the function of co-party bias, based on democracy, is descriptively similar to a sinusoidal function.

3.1 Benchmark Model

Let N be the total population, while B (base) is the number of people who supported the incumbent in the previous election. We define ρ as the proportion of individuals who do not change their opinions, i.e. do not move across party lines. ρ can be interpreted as partisanship or, more directly, as 'ideological attachment'. In the next period, the government will receive ρB votes from its original base and $(1 - \rho)(N - B)$ votes from previously non-supporters.

The government has three policy choices: Core, Swing, and None. Its goal is to maximise the number of expected supporters in the next term, and so we assume by convention that the government is office-seeking in nature. Politically motivated transfers from the government to people are used to impact ρ for the targeted population. The core vote strategy increases ρ for the base, i.e. makes it harder for government supporters to change their party associations. The swing vote strategy, on the other hand, has a negative impact on ρ for non-supporters, $N-B$. Let ε be the impact of either strategy on ρ. Table 1 shows the payoff matrix the government faces:

Table 1: Payoff Matrix

Time	Core	Swing	None
T_0	B	B	B
T_1	$(p+e)B+$ $(1-p)(N-B)$	$pB+$ $(1-p+e)(N-B)$	$pB+$ $(1-p)(N-B)$

Now, we modify this payoff matrix by considering the value of a vote from the government?s perspective and adding political costs of redistribution. Let $0 < \delta < 1$ be the influence of citizens? opinion on policy decisions. If $\delta = 0$, then the government has total control and no accountability, i.e. autocracy. As δ increases, citizens have more power in the government and so government policies better reflect individuals' preferences. That is why we treat δ as our democracy score. The incumbent will value every vote as much as δ. Then, the new goal for the government is not to maximise the number of votes it will receive in the next election, but rather to maximise its power in the next term. A perfect autocracy, for example, does not care about the votes it will receive because they do not augment the power of the government.

Assumption 1 *ρ is increasing with institutional constraints and political participation.*

We assume there are two inputs that go into one's attachment to an ideology. One is the institutional constraints denoted as ω. In an autocracy, the individuals have no choice and thus face information constraints. A government supporter will not easily change their opinion not because they sincerely agree with the government but because they lack the choices and information to disagree, which thus yields a high ρ. Similarly, if one opposes an autocratic government, their opposition must be strong enough to overcome the institutional constraints implying high ideological attachment, ρ. With increasing democracy, on the other hand, the institutional constraints will decrease: citizens will have more choices and be able to access more information. We formalise individual constraints as:

$$\omega = \left(\frac{1}{\delta+k}\right) m \tag{1}$$

where k is an arbitrary cap on constraints and m is a scaling factor. With zero democracy, $\frac{m}{k}$ is the maximum level of institutional constraints.

Another input in ρ is political participation denoted as θ. Ideological attachment has been linked to participation in campaign activities, turnout rates, interest in politics, and other forms of political participation (Dalton 1984, King 1969, Converse and Dupeux 1962, Verba et al 1978). There is also clear evidence that democracy enhances civic engagement and political participation (McNeal, Smith and Tolbert 2003). Moreover, letting people act as law-makers increases their interest in politics (Schmidt 1989; Zimmerman 1999). It thus follows that greater democracy brings about higher participation, which is related to ideological attachment. For simplicity, we assume a linear relationship between democracy and participation, $\theta = \delta$.

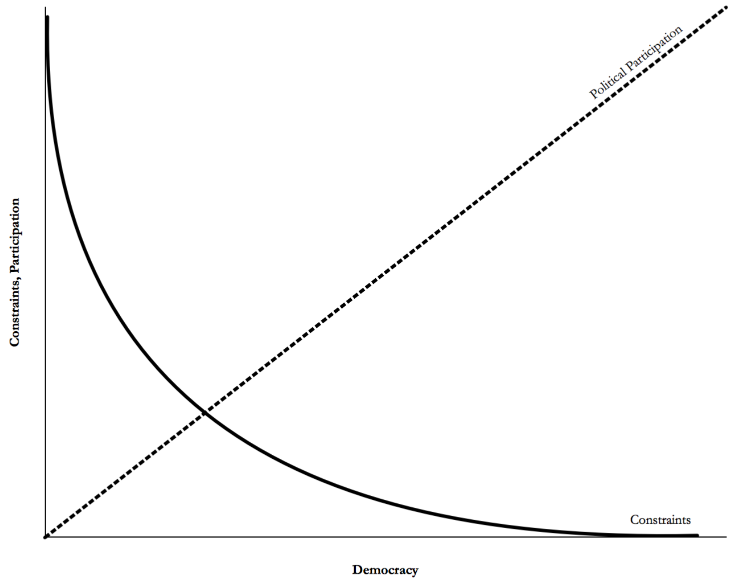


Figure 1: Institutional Constraints and Political Participation in Relation to Democracy

Proposition 1 *Suppose assumption 1 holds, then ρ is convex and initially decreasing in δ*

$$\rho(\omega, \theta) = \frac{((\omega + \theta) \ + s)}{\frac{m}{k} + s} \tag{2}$$

where s is the fraction of 'sticky voters' whose party identity does not change with or without constraints and/or participation. This relationship gives us the convex and initially decreasing ρ in δ. The relationship is scaled by $\frac{m}{k} + s$ in order to satisfy $\rho \in (0, 1)$.

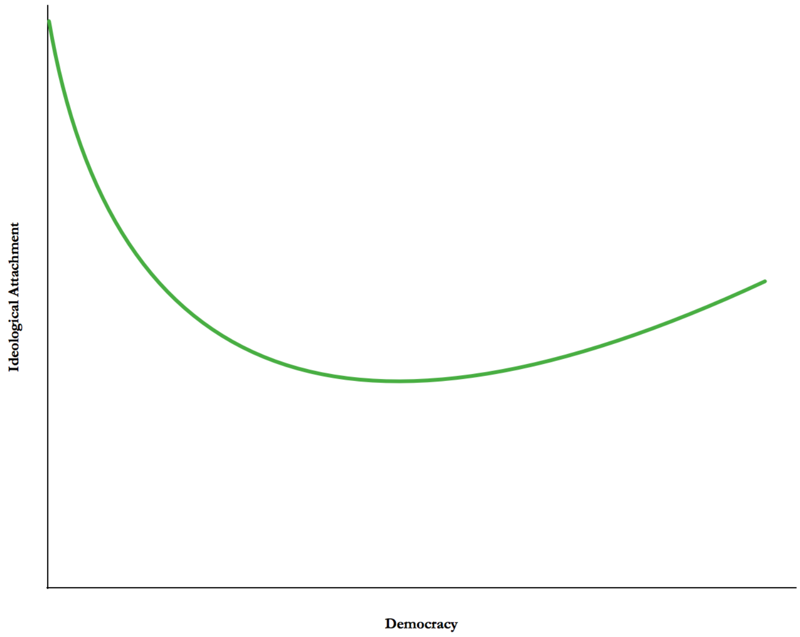


Figure 2: Ideological Attachment in Relation to Democracy

Our function of ρ is consistent with the findings of Karp and Banducci (2007) whereby party attachments are weak in newly established democracies.

Assumption 2 *The minimum cost of moving a person across party lines is increasing in ρ, and the minimum cost of keeping a voter's support is decreasing in ρ (increasing in $1 - \rho$).*

In their redistribution model, Dixit and Londregan (1995, 1996) claim that to convince someone to swing to another party, the transfer must be at least as strong as the target's ideological attachment. By that logic, the transfer needed to maintain a citizen's support must be small if their ideological attachment is high since they are likely to support the incumbent in the next election anyway. Thus, the minimum monetary cost of redistribution that is successful in its goal of maintaining a supporter will be related to $1 - \rho$, where the minimum monetary cost of swinging a supporter is related to ρ. As ρ is a maximum when $\delta = 0$, the cost of the core vote strategy will be at its lowest for an autocracy, whereas the cost of the swing vote strategy is at its highest.

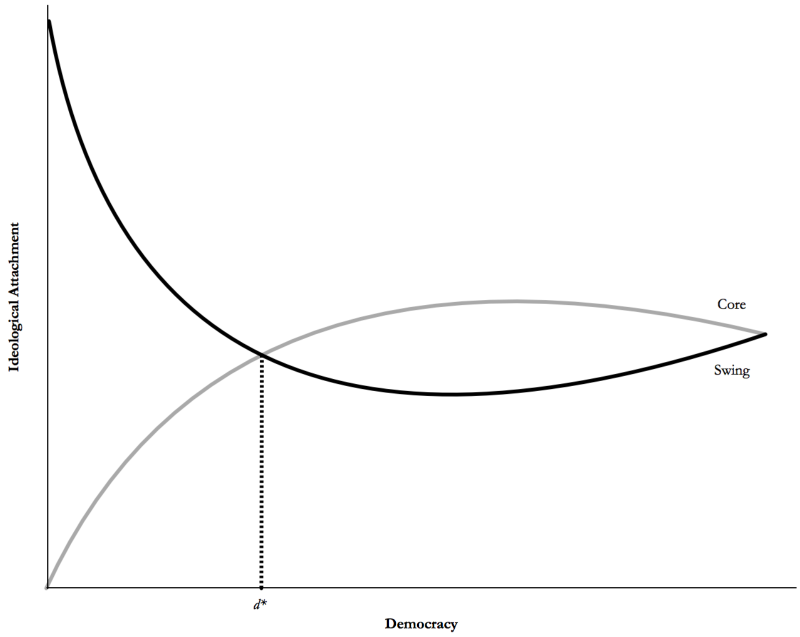


Figure 3: Monetary Cost of Redistribution

Assumption 3 *The reaction of the public to politically motivated redistribution is proportional to the size of redistribution and the public's participation in the politics.*

The intuition behind this relationship is that a politically motivated redistribution will not receive any backlash from the public in a country where citizens are not participating in politics, whereas it will be placed under more scrutiny as political participation increases. Based on Assumptions 2 and 3, the political cost of the core vote strategy is related to $(1 - \rho)\delta$ and the political cost of the swing vote strategy is related to $\rho\delta$. The reason we are interested in the political cost(s) is because we cannot compare the gains from redistribution to monetary costs. The lost popularity based on redistribution, however, can be compared to the gained support.

Proposition 2 *Co-party bias is positive for δ close to zero and negative when δ is close to 1.*

Including the political gains and losses from redistribution, the government faces the following payoff matrix in Table 2.

Table 2: Payoff Matrix

	Core	Swing	None
T_0	B	B	B
T_1	$(\rho+\varepsilon)\delta B + (1-\rho)\delta(N-B)$	$\rho\delta B + (1-\rho+\varepsilon)\delta(N-B)$	$\rho\delta B + (1-\rho)\delta(N-B)$
	$-(1-\rho)\delta B$	$-\rho\delta(N-B)$	

Subtracting the payoff of policy option None from both Core and Swing strategies, the new payoff matrix is as follows in Table 3.

Table 3: Payoff Matrix

Core	Swing
$\varepsilon\delta B - (1-\rho)\delta B$	$\varepsilon\delta(N-B) - \rho\delta(N-B)$

For simplicity, assume $B = N - B$. This is especially the case if the incumbent has a minimum majority, whereby supporters are equal to non-supporters, or if the government is targeting individuals instead of populations, whereby $B = N - B = 1$. In the latter, ρ can be interpreted as the 'probability that an individual switches party lines' instead of a vote share proportion. Also, via individually targeted redistributions, we can still assume ρ to be the same for all individuals, as the government can observe ρ for a population based on previous elections but not for an individual. Nevertheless, we are interested in population targeting within the scope of this paper.

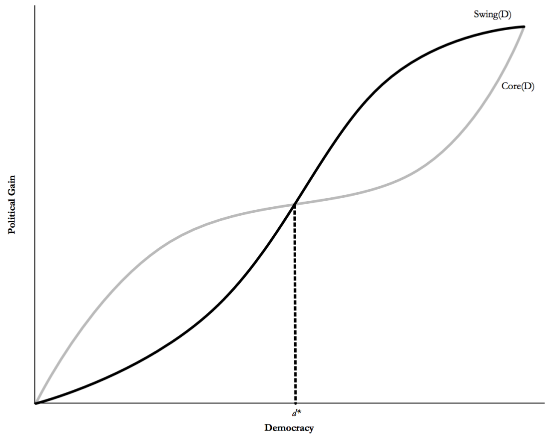


Figure 4: Political Gains from Redistribution

Figure 4 shows that when plotted in relation to δ, we see that the gains from core vote policies surpasses the gains from swing vote policies for low values of δ, and vice versa for high values of δ.

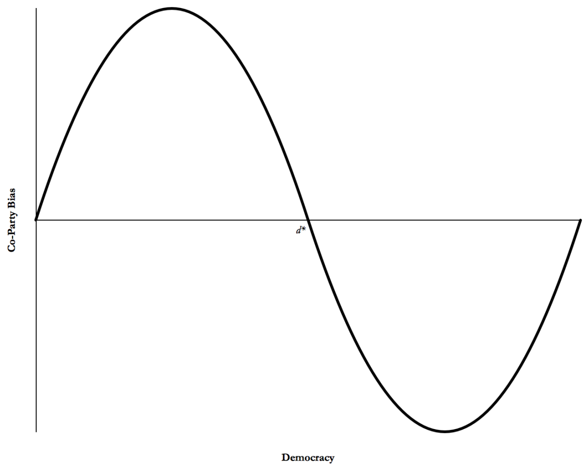


Figure 5: Co-Party Bias in Relation to Democracy

Consequently, the dominant strategy for an autocratic government is to distribute transfers to its supporters, whereas the dominant strategy for a more democratic government is to target swing voters, as is evident in Figure 5.

4 Institutional Background and Data

4.1 Institutional Background

Despite its current hegemonic tendencies, the AKP was not formed with authoritarian or autocratic intentions, nor did it continue the practices of old authoritarianism characterised by excessive use of armed forces to maintain status quo and political stability (Somer, 2016). Influenced by private Turkish entrepreneurs who prioritised global markets, EU membership, limited government, and democratisation, the AKP was created in 2002 by businessmen and politicians who were part of the disbanded Welfare Party and Virtue Party (Keyman and Gumuscu, 2014). As the AKP helped heal the economic damage caused by the 2001 economic crisis, committed to full EU membership in 2005, and encouraged inter-party collaboration for legal-political reforms, it appeared as if Turkey was on the path towards becoming a full liberal democracy.

Turkey's turn to autocracy began in 2007 when the government abolished any parliamentary quorum requirements for impending presidential elections, relying only on popular vote. Although this referendum seemed to place more power into the hands of the citizens, presidential elections could bypass parliament and become more susceptible to government manipulation (Somer, 2019). In early 2008, the Dogan Media Group uncovered corruption scandals within the AKP, to which the AKP immediately launched a controversial campaign against them, encouraging others to undermine the DMG's credibility as a news outlet. Furthermore, the AKP government in 2009 began mass imprisonment of journalists on dubious charges. The 2010 Referendum brought the High Council of Judges and Prosecutors under the government's control, and it weakened the military check on the AKP's executive power. In 2013, the Gezi Park Protests erupted over issues surrounding the freedom of press, expression, assembly, and government secularism, which resulted in 11 deaths and over 8,000 injuries. With the government dissatisfied with the parliamentary results of the June 2015 election, a snap November election was held, further compounding the notion of a faux competitive political landscape (Sayari, 2016). Finally, Turkey's constant state of emergency, established in 2016 after the failed coup attempt, has brought further restrictions of the media and resulted in the suspension of the European Convention on Human Rights. Although by no means a comprehensive timeline of Turkey's transition to autocracy, these events convey the trend the AKP has gravitated towards, becoming a phenomenon political scientists refer to as "Erdoganism" (Yilmaz, 2018). The AKP's shift from democracy to autocracy has been captured by the drop in its Polity IV score, from its peak score at 9, "Democracy" to its current score of 2, an "Open Anocracy". AKP hegemony arose as the party enjoyed persistent and overwhelming electoral victories over a long period of time after having a significant influence over the political landscape. In this vein, the AKP also draws in its support by simultaneously pushing out much of the political competition and appealing to social conservatives with traditionalist Islamic values. Regional favouritism is commonly used by the AKP as a reward system in exchange for loyal voting behaviour (Cinar, 2016). With the irrefutable dominance of the AKP in the last two decades, a new relationship between the state and its constituents has emerged, whereby democratic institutions are weakened in exchange for greater state power. This has generated the belief amongst Turkish constituents that the continuation of the economic benefits enjoyed relies on the AKP under President Recep Tayyip Erdo?an alone (Somer, 2016). Even so, the margin between the AKP and the closest leading party is wide enough that Turkey's constituents believe there is indeed no alternative (Keyman and Gumusku, 2014).

The fluctuations in Turkey's political institutions over the past two decades makes it a perfect empirical setting for testing our theory of co-party bias as a function of democracy. As is clear from the AKP's political behaviour since 2007, institutional restraints have increased year to year, and the commitment to democracy has conversely decreased. Thus, connecting these events back to our theoretical framework, it seems that δ is approaching 0, as Turkey's constituents have a diminishing effect on political decisions, and ρ is close to 1. Thus, with proposition 2 in mind, we expect positive co-party bias in Turkey. Because the cost of maintaining a supporter is lower than winning a swing voter in an autocracy, we expect the AKP to maximise their votes by providing higher redistribution, or TOKI housing, to AKP loyal provinces.

4.2 Data

Our units of observation are organised as time clusters between election years, where the independent variables are lagged to the time period prior to an election, and our dependent variable extends to the period between national elections. We use lags because evidence suggests that the results of the AKP Vote Share affect the stock of TOKI construction projects in the following rather than the immediate period. The intuition is that building proposals normally take a substantial amount of time to be planned and approved by the incumbent party before they are executed. We lagged our confounders to the period prior to an election because voting behaviour is influenced by these factors. To ensure we control for systematic differences between provinces, we include data for average provincial GDP per capita, the number of natural disasters (categorised by whether a province experienced an earthquake of magnitude 6 or greater in a given year), which thus subsequently damages infrastructure in specific provinces, net migration in a province, and housing stock.

The TOKI housing data was obtained from the government's Public Housing Development Administration (TOKI). TOKI is the government-backed public housing program established in 1984 that aims for disaster relief and strengthening urbanisation in lower and middle-income areas. Data for each project is recorded by consultant firms and the TOKI itself. As TOKI projects are our dependent variable of interest, we focus on the core-vote hypothesis to test whether Turkey's gradual shift towards autocracy has made it such that provinces with stronger incumbent (AKP) support and authority garner preferential treatment in terms of social housing. For all 81 provinces, this variable, $Houses$, is derived as the yearly average of the number of houses constructed per thousand residents during a time cluster following an election. Figures 6 and Figure 7 demonstrate the variation in the data regarding our dependent and independent variables of interest.

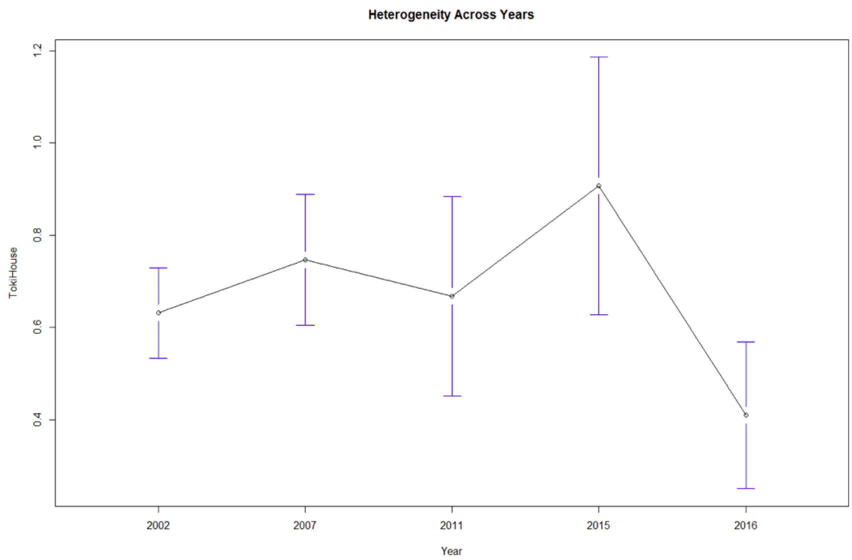


Figure 6: Heterogeneity Across Years

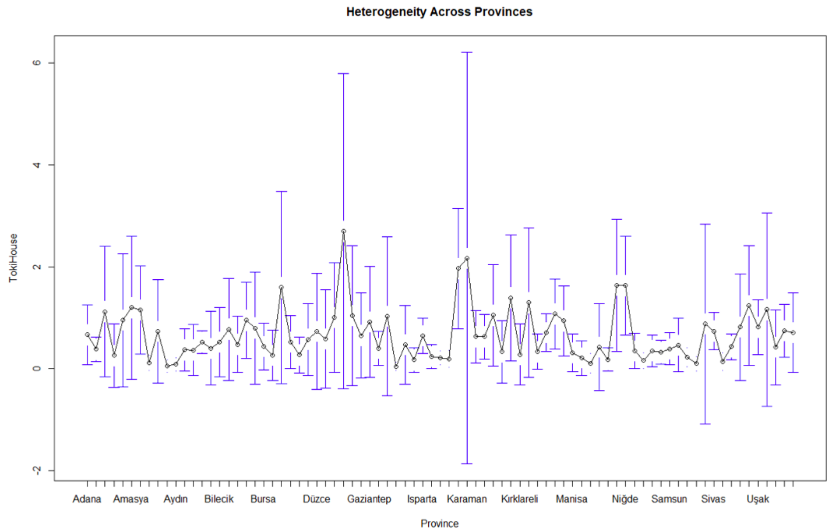


Figure 7: Heterogeneity Across Provinces

The Turkish Statistical Institute is a government agency which publishes national figures in the fields of economy, social issues, demography, culture, and environment. Our main independent variable, the percent of parliamentary votes for the AKP during a given election year are reported for the national election years from 2002-2015. Data for *education* levels, the instrumental variable in our regression, was also taken from this database and is defined as the percentage of citizens in the provinces who hold a college degree or above in a given election year. Our confounding variables, $logGDP$, $Migration$, and $HousingStock$, are defined as the logarithmic value of average provincial GDP per capita, net provincial migration per thousand, and the number of housing units per thousand in each province, respectively. Each variable is observed in our specified time clusters from 2000 to 2015 and are taken from the Turkish Statistical Institute. Measures for our final confounder, $Disasters$ (natural disasters in each province), come from the Disaster and Emergency Management Authority (AFAD) and are defined as earthquakes above magnitude six from 2003 to 2016.

Figure 8 plots TOKI Houses and AKP support, separated by year. We can see a clear trend that shows an increase in housing projects that peaks around the 50% mark across all years. Note that the 2015 election has the highest quantity of housing projects near the 50% mark, which bodes well for our core vote hypothesis.

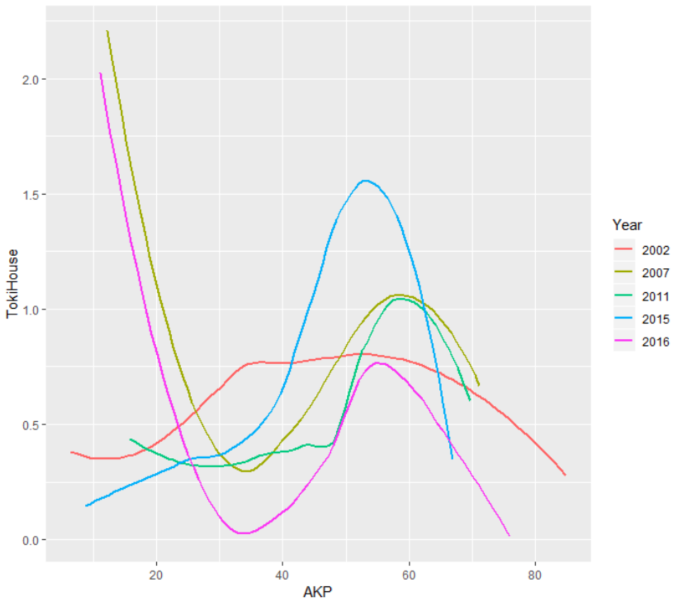


Figure 8: TOKI Houses versus AKP Vote Share by Year

Figure 9 plots the trend for AKP Vote Share and TOKI Houses, broken down by province (the axes for both vote share and housing projects have been normalised). Some provinces exhibit dually increasing trends in vote share and housing, such as Amasya, Cankiri, and Kutahya.

Table 4: Summary Statistics

	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
TokiHouse	405	0.672	0.869	0.000	0.121	0.895	7.903
AKP Vote Share	405	45.224	15.517	6.500	35.500	57.100	84.820
Education	405	8.517	4.294	0.030	4.840	11.730	22.250
GDPK	405	6,891.76	3,169.336	1,752.630	4,509.750	8,698.750	19,562.250
Housing Stock	405	39.237	38.906	0.000	5.992	60.880	219.982
Migration	405	-2.548	9.050	-29.030	-6.550	2.130	47.380

Problems with our data arose when measuring GDP per capita from 2000-2003 because the Turkish Statistical Institute began recording these figures in 2004. Thus, we took national GDP growth trends from 2000-2003 and extrapolated figures recorded in 2004 backward in time to estimate the nominal value during these missing years. Although these estimated figures are inaccurate, we believe it is not too large of an empirical drawback as it was only for years preceding the 2002 election. Furthermore, Turkey experienced a drastic recession in 2001, so actual figures of GDP per capita during this time are likely to be volatile and, if officially measured, would have potentially been recorded inaccurately.

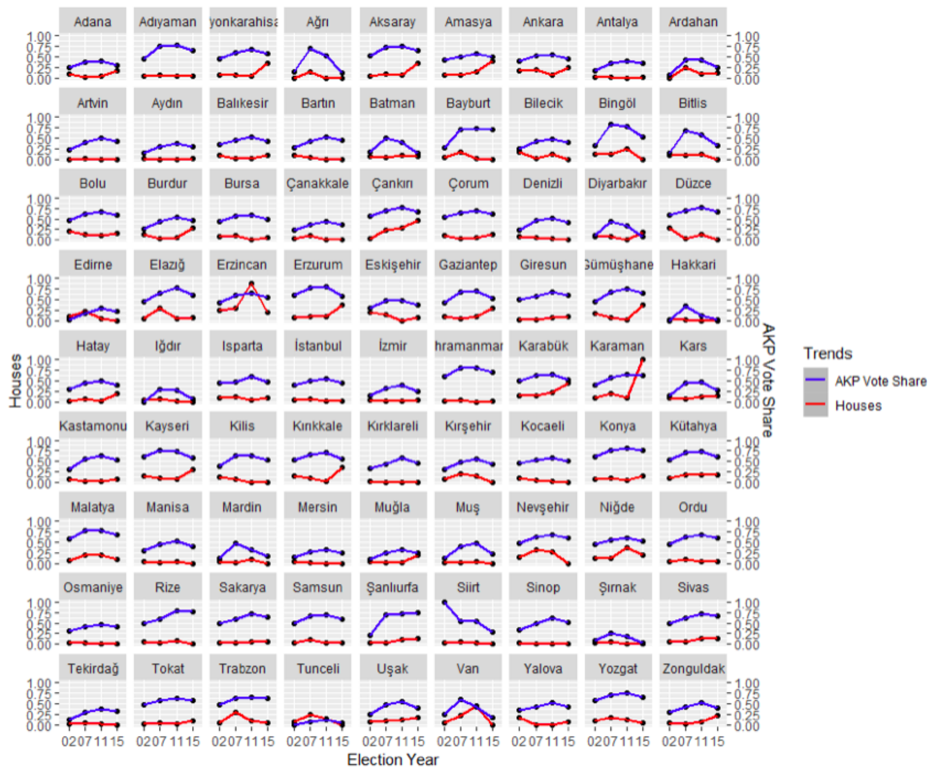


Figure 9: TOKI Houses per Province

Similarly, net migration data immediately prior to 2008 was absent on the Turkish Statistical Institute's database. To address this problem, we averaged net migration from the 2002 election and from 2008 to use as an estimate for the 2007 election. Since we do not have any information on migration trends for each province during this period, averaging was our best estimation strategy.

5 Methods and Results

Our goal is to estimate how changes in vote share for the AKP influence public expenditure, denoted through the average number of public housing projects (TOKI) allocated to a province in the years between national elections. We used the period 2002-2015, a period where Erdogan maintained power whilst Turkey's polity score started to rapidly decrease around 2011. We are primarily interested in how the distribution of public housing projects changed as Turkey became more autocratic. The hypothesis is that as Turkey became more autocratic over time, Erdogan's government began to allocate more public housing projects to provinces that had an AKP majority vote share in the last national election.

To build a baseline understanding of the general trend, we ran a cross-sectional OLS for each election year and plotted the coefficients for $AKPVoteShare$ against the polity score trend. Our OLS specification is:

$$\begin{aligned} Houses \ = \ \alpha \ + \ \beta_1 AKPVoteShare \ + \ \beta_2 logGDP \ + \ \beta_3 Disaster \\ + \ \beta_4 HousingStock \ + \ \beta_5 Migration \ + \ \varepsilon \end{aligned} \tag{3}$$

The dependent variable, $Houses$, is calculated by taking the amount of housing projects built in the years between national elections in a province and averaging the values, then transforming it so that it is in units per 1000 residents. The averaging was done to eliminate the occurrence of provinces assigned numerous projects in one year but zero projects in the following year. $AKPVoteShare$ is the percentage that AKP gets in a province during a national election. Note that in the initial cross-sectional analysis we pooled the provinces, so that the independent variables aren't indexed by province or by year. $logGDP$ is the GDP per capita by province per year, transformed logarithmically to allow for easier interpretation. $Disaster$ is an indicator variable for whether a province experienced an earthquake with magnitude greater than 6 during the period used for the analysis. Note that for our initial cross sections the model for the 2015 national election has $Disaster$ omitted because there was no variance in the observations for this period (there were no earthquakes with a magnitude equal to or greater than 6). $HousingStock$ is the existing housing stock of a province per 1000 residents, by year. $Migration$ is the net annual average of people that move into a province, also standardised per 1000 residents of a province. For each election year (2002, 2007, 2011, 2015) all covariates are lagged a year behind our dependent variable. Lagging the covariates is done to maintain the intuition of these factors influencing future housing projects. Since the public housing projects are decided well in advance, it is more likely that current trends in the covariates influence housing projects than future, unknown values of the covariates.

Table 5: Cross Section by Year

	(1)	(2)	(3)	(4)
Variables	2002	2007	2011	2015
AKP Vote Share	0.00932***	0.0103*	0.0172**	0.0263***
	(0.00338)	(0.00577)	(0.00809)	(0.00951)
log(GDP)	0.293*	-0.0149	0.0805	-0.146
	(0.158)	(0.272)	(0.492)	(0.702)
Disaster	0.276	1.562**	3.120***	
	(0.250)	(0.602)	(0.971)	
Housing Stock	0.00201	0.0143*	-0.00197	0.00666
	(0.0119)	(0.00792)	(0.00609)	(0.00568)
Migration	0.00422	-0.0175	-0.00241	-0.0139
	(0.0150)	(0.0115)	(0.0165)	(0.0222)
Constant	-2.102	0.0956	-0.907	0.640
	(1.299)	(2.314)	(4.337)	(6.187)
Observations	81	81	81	81
R^2	0.170	0.192	0.166	0.119

Standard errors in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 5 shows the results from our initial cross sections. AKP Vote Share has a statistically significant coefficient to the 100% level for the years 2002 and 2015, whereas 2007 is significant at the 90% level and 2011 is significant at the 95%. The interpretation of the coefficients for AKP Vote Share is that for an $X\%$ vote share for AKP in some province, Houses increases by $\beta_1 X$. For instance, an increase in AKP Vote Share by 1% in 2015, would lead to an increase in 0.0263 housing projects per thousand residents, or 2.63 housing projects per 100,000 residents. As we expected from our hypothesis, the coefficients steadily increase each year, signalling a growing allegiance from Erdogan to provinces that vote more heavily for AKP.

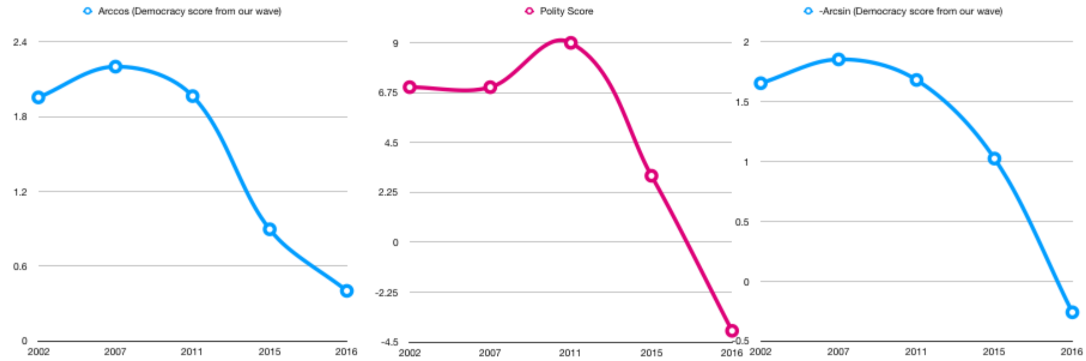


Figure 10: Turkish Democracy from 2002 to 2018

Figure 10 plots the coefficients for $AKPVoteShare$ against the polity scores showing a similar, yet more rapidly decreasing trend toward autocratic behaviour. Note that in Figure 10 our estimator for co-party bias shows a more rapid decrease in democracy compared to the change in polity score. One reason for this may be the case, other than variations in data used for estimating the level of democracy, is that Polity's methodology is slow in recognising the development of autocratic behaviour.

Post-coup TOKI projects (2016-2018) fall on the increasing part of the wave. Zero party bias in projects after 2016 further supports our point about close to extreme autocracies. As a country becomes more autocratic, it reaches a critical turning point where there is not much to gain from rewarding citizens because they have high ideological attachments and their opinion no longer matters for determining power.

5.1 Instrumental Variable Approach

The results from the initial OLS show promising evidence in support of our hypothesis. However, establishing a causal story requires further specifications to capture any omitted variables that may be contributing to incorrect estimates for the $AKPVoteShare$ coefficients. One of the biggest concerns was solving the issue of simultaneity. Previous literature has found a relationship between public expenditure and electoral outcomes but looked at public expenditure as an explanatory variable for election results. Instrumental variable estimation is a regression specification that addresses simultaneity. We chose the instrument education even though it is not a completely exogenous regressor (i.e. a random event or trial). We included controls for education to maintain the exogeneity condition necessary for instrumental variable estimation to remain valid. Using $education_{it}$ as an instrument, we run a two stage least squares regression, with state and time fixed effects:

$$\begin{aligned} Houses_{it} = \beta_0 + \sum_{j=2}^{81} \gamma_j Province_i + \delta_1 2007 + \delta_2 2011 + \delta_3 2015 + \beta_1 \widehat{AKP}_{it} \\ + \beta_2 logGDP_{it} + \beta_3 Disaster_{it} + \beta_4 HousingStock_{it} + \beta_5 Migration_{it} + \varepsilon_{it} \end{aligned} \tag{4}$$

The new coefficient β_1 uses the fitted values from the first stage regression:

$$\begin{aligned} \widehat{AKP}_{it} = \beta_0 + \sum_{j=2}^{81} \gamma_j Province_i + \delta_1 2007 + \delta_2 2011 + \delta_3 2015 + \beta_1 education_{it} \\ + \beta_2 logGDP_{it} + \beta_3 Disaster_{it} + \beta_4 HousingStock_{it} + \beta_5 Migration_{it} + \varepsilon_{it} \end{aligned} \tag{5}$$

The dependent variable $Houses_{it}$ is the quantity of houses per 1000 residents in province i in time t. Each of the 81 provinces in Turkey was assigned an index number, and the index t denotes one of four periods associated with a national election. For example, using $t = 1$ would take vote share values from the 2002 national election, housing projects averaged between 2003-2007, and the remaining covariates from 2000-2002. We also include time dummies to control for time variant trends that may have impacted the distribution of housing projects. The variable $\sum_{j=2}^{81} Province_i$ is a vector of state fixed effects corresponding to each of the 81 provinces in Turkey. The first province, Adana, is omitted from the vector and β_0 is the common intercept for both time and state fixed effects. Another way of writing the second stage specification is:

$$\begin{aligned} Houses_{it} = \alpha_i + \lambda_t + \beta_1 \widehat{AKP}_{it} + \beta_2 logGDP_{it} + \beta_3 Disaster_{it} \\ + \beta_4 HousingStock_{it} + \beta_5 Migration_{it} + \varepsilon_{it} \end{aligned} \tag{6}$$

Where α_i is an intercept of state fixed effects for each province and λ_t is the intercept for time fixed effects for each period.

Table 6: Fixed Effects Results and Two-Stage Least Squares with Education

	(1) Pooled OLS	(2) Province FE	(3) Time FE	(4) Two-Way
Panel A	Dependent variable is $Houses_{it}$			
$AKPVoteShare$	0.0213	-0.00208	0.0306	0.0800**
	(0.0227)	(0.0131)	(0.0280)	(0.0425)
$log(GDP)$	-0.119	0.266	0.124	1.351
	(0.238)	(0.275)	(0.145)	(1.957)
$Disaster$	1.150**	0.836	1.209**	1.119**
	(0.531)	(0.598)	(0.554)	(0.482)
$HousingStock$	0.00383	-3.18e-05	0.00199	-0.00674
	(0.00299)	(0.00305)	(0.00330)	(0.00422)
$Migration$	-0.00111	-0.000485	-0.00554	0.00799
	(0.00948)	(0.0143)	(0.0112)	(0.0139)
Panel B	First stage for $AKPVoteShare_{it}$			
$Education$	-1.794**	-4.601*	-1.888*	-2.161**
	(0.541)	(0.592)	(0.565)	(0.762
$log(GDP)$	15.20***	47.80**	7.216*	-25.06
	(3.647)	(3.860)	(3.453)	(15.27)
$Disaster$	-4.409	-3.697	-3.877	-3.711
	(7.694)	(3.030)	(6.263)	(3.755)
$HousingStock$	0.114**	0.125**	0.129*	0.0664*
	(0.0436)	(0.0466)	(0.0448)	(0.0375)
$Migration$	-0.521***	0.125	-0.255*	-0.0269
	(0.121)	(0.0973)	(0.137)	(0.0895)
Province Fixed-Effects	No	Yes	No	Yes
Time Fixed-Effects	No	No	Yes	Yes
Observations	324	324	324	324
Number of Provinces	81	81	81	81

Robust standard errors in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

To overcome endogeneity in determining the causal relationship between AKP vote share and public housing provision, we decided to include education as our instrumental variable. In our dataset, *education* is defined as the percentage of citizens in the provinces who hold a college degree or above in the year of which the election happened (t). The data is obtained from the government statistic dataset of Turkey, TUIK. Numerous papers that analysed the formation of AKP supporters showed that there is a negative correlation between education level and supporting rate of the Justice and Development Party (AKP). AKP has attracted voters from a variety of parties across the political spectrum, including voters from the pro-Islamic party (Virtue Party, FP), centre-left party (Democratic Left Party, DSP), centre-right party (True Path Party, DYP and Motherland Party, ANAP) and far-right conservative party (Nationalist Movement Party, MHP) (Baslevent and Akarca, 2008). Yet across different previous political backgrounds, the supporters of AKP tends to be less educated which is summarised from different surveys. Since then, such trends did not vary a lot as Marscall (2016) expected that districts with more educated populations tend to support parties like the left-centre Republican People's Party (CHP).

On the other hand, we cannot build a direct link between education and public housing supply, which is a requirement for satisfying the exclusion restriction in instrumental variable estimation. Although both factors are linked with the local income and economic growth, the nature of the TOKI fund means that it is more centrally controlled and is less constrained by the local finance. Given the aforementioned literature on the relationship between education levels and the AKP's vote share, we thus use education as a valid instrument in our 2SLS specifications. In terms of the exclusion restriction, we feel this is not violated. Despite being a 'blunt' instrument, we see no perceivable chain of reasoning as to why provincial education levels would directly impact the number of TOKI houses provided by the government.

The high statistical significance of the coefficient for education in the first stage regression from Table 6 provides a strong defence of the validity for our choice of instrument. However, education is potentially and highly likely to be endogenous. Another concern is the issue of reverse causality between education and $logGDP$, which could be remedied either by instrumenting some new, exogenous variable on $logGDP$ or using a new, completely exogenous instrument other than education. We assume that some of the covariates used for controlling for $Houses$ in the second stage also influence education in the first stage, maintaining the exogeneity condition needed for instrumental variable estimation.

The results from the second-stage fixed effects echo the results presented earlier, albeit with a smaller coefficient for $\widehat{AKP_{it}}$ (also only statistically significant to the 90% level) and a statistically significant coefficient for disaster. For a one percent increase in the vote share for the AKP in a province, the number of housing projects allocated to that province increases by 0.08 per 1000 residents, or 8 houses per 100,000 residents per vote share percent. This means that for a province with a vote share of fifty percent, there was an average of 400 more projects built per 100,000 residents.

6 Conclusion

Our current results point to a normatively significant effect of co-party bias on the distribution of public expenditure. Over time, shifts toward a Turkish government exhibiting more autocratic behaviour seemingly led it to move more of its public housing projects to provinces that voted heavily in favour for AKP. At first glance this sort of behaviour sounds intuitively ordinary for a government that has become more centralised in its decision-making capabilities. Erdogan's shift to fuel more resources to provinces that support AKP follows the core voter hypothesis framework that we developed. As Turkey became more autocratic, co-party biased increased, altering the distribution of government expenditure. Theoretically, if Turkey continues the path toward full autocracy, we would eventually see the fall-off effect from co-party bias where the distribution of public expenditure becomes "fairer", i.e. not motivated by maintaining the core vote or swinging voters from other parties to support AKP.

All this being said, the results from our current research are by no means completely robust. The methodology employed should be considered more of an introductory approach to a problem that requires far more time and careful consideration to draw causal conclusions. We were limited not just by our own time but also econometric training. There were a variety of notable variables that may have been important to include as confounders, such as religiosity by province. Future research should investigate utilising nested models such as mixed effects by including data on the city level for election data. Unlike the provincial elections, which are proportional representation, city-level elections are first-past-the-post, and the same parties that run for Turkish parliament run for city mayor. The addition of city-level data markedly increases sample size and can further break down the possible effects of democratic backsliding in Turkey by looking at the redistribution of housing projects to cities with mayors from the AKP. The inclusion of city-level data would also increase the amount of time periods, as city-level elections are held in different years than national elections. Another important aspect of our empirical design is the endogeneity of our instrument, education. The high likelihood of education being endogenous damages the validity instrumental variable estimates, as well as the issue of reverse causality between education and GDP per capita. Any developments made to our methods should include either finding a new instrument that is certainly exogenous, like a random event (weather, natural disasters, etc.) or utilising a different method of instrumental variable estimation, such as three-stage least squares (also known as “seemingly unrelated regression”) to establish the exogeneity condition necessary to use education as a valid instrument.

References

Atkinson, J., Hicken, A., & Ravanilla, N. (2011). Pork and typhoons: The political economy of disaster assistance in the Philippines. Manuscript, Ann Arbor, Mich.: University of Michigan.

Baslevent, C., & Akarca, A. (2008). Micro evidence on inter-party vote movements in turkey: Who voted for AKP in 2002?.

Besley, T., & Persson, T. (2011). Pillars of Prosperity. Princeton: Princeton University Press.

Burgess, R., Jedwab, R., Miguel, E., Morjaria, A., & Padró i Miquel, G. (2015). The value of democracy: evidence from road building in Kenya. American Economic Review, 105(6), 1817-51.

Cinar, K. (2016). Local determinants of an emerging electoral hegemony: The case of Justice and Development Party (AKP) in Turkey. Democratization, 23(7), 1216-1235.
Converse, P. E., & Dupeux, G. (1962). Politicization of the Electorate in France and the United States. Public Opinion Quarterly, 26(1), 1-23.

Cox, G. W., & McCubbins, M. D. (1986). Electoral politics as a redistributive game. The Journal of Politics, 48(2), 370-389.

Curto-Grau, M., Solé-Ollé, A., & Sorribas-Navarro, P. (2018). Does electoral competition curb party favoritism?. American Economic Journal: Applied Economics, 10(4), 378-407.

Dahlberg, M., & Johansson, E. (2002). On the vote-purchasing behavior of incumbent governments. American political Science review, 96(1), 27-40.

Dalton, R. J. (1984). Cognitive mobilization and partisan dealignment in advanced industrial democracies. The Journal of Politics, 46(1), 264-284.

Dixit, A., & Londregan, J. (1995). Redistributive politics and economic efficiency. American political science Review, 89(4), 856-866.

Dixit, A., & Londregan, J. (1996). The determinants of success of special interests in redistributive politics. the Journal of Politics, 58(4), 1132-1155.

Hodler, R., & Raschky, P. A. (2014). Regional favoritism. The Quarterly Journal of Economics, 129(2), 995-1033.

Jablonski, R. S. (2014). How aid targets votes: the impact of electoral incentives on foreign aid distribution. World Politics, 66(2), 293-330.

Keyman, E. F., & Gumuscu, S. (2014). Constructing Hegemony: The AKP Rule. In Democracy, Identity, and Foreign Policy in Turkey (Islam and Nationalism Series, pp. 29-43). London: Palgrave Macmillan.

King, A. (1969). Political parties in western democracies: some sceptical reflections. Polity, 2(2), 111-141.

Lasswell, H. D. (1936). Politics: Who gets what, when, how. New York: P. Smith.

Lindbeck, A., & Weibull, J. W. (1987). Balanced-budget redistribution as the outcome of political competition. Public choice, 52(3), 273-297.

Marschall, M., Aydogan, A., & Bulut, A. (2016). Does housing create votes? Explaining the electoral success of the AKP in Turkey. Electoral Studies, 42, 201-212.

Miguna, M. (2012). Peeling Back the Mask: a quest for justice in Kenya. Gilgamesh Africa.

Sayari, S. (2016). Back to a Predominant Party System: The November 2015 Snap Election in Turkey. South European Society and Politics, 21(2), 263-280.

Somer, M. (2016). Understanding Turkey's democratic breakdown: old vs. new and indigenous vs. global authoritarianism. Southeast European and Black Sea Studies, 16(4), 481-503.

Somer, M. (2019). Turkey: The Slippery Slope from Reformist to Revolutionary Polarization and Democratic Breakdown. The ANNALS of the American Academy of Political and Social Science, 681(1), 42-61.

Stokes, S. C. (2005). Perverse accountability: A formal model of machine politics with evidence from Argentina. American political science review, 99(3), 315-325.

Tolbert, C. J., McNeal, R. S., & Smith, D. A. (2003). Enhancing civic engagement: The effect of direct democracy on political participation and knowledge. State Politics & Policy Quarterly, 3(1), 23-41.

Ward, H., & John, P. (1999). Targeting benefits for electoral gain: Constituency marginality and the distribution of grants to English local authorities. Political Studies, 47(1), 32-52.

Yilmaz, I., & Bashirov, G. (2018). The AKP after 15 years: emergence of Erdoganism in Turkey. Third World Quarterly, 39(9), 1812-1830.

Zimmerman, J. F. (1999). Citizen Law-making: The Initiative. Graduate School of Public Affairs, State University of New York at Albany.

Contacts

Health Economics Working Group

Anmol Kanotra (a.kanotra@lse.ac.uk)
Kerry Neitzel (k.neitzel@lse.ac.uk)
Maithili Modi (m.modi1@lse.ac.uk)
Abi Thomas (a.j.thomas@lse.ac.uk)
Xuyi Yang (x.yang46@lse.ac.uk)

Urban Economics Working Group

Christopher Dann (c.m.dann@lse.ac.uk)
Jakub Grohmann (j.grohmann@lse.ac.uk)
Amy Ho (a.ho2@lse.ac.uk)
Sherman Khoo (j.khoo3@lse.ac.uk)
Anna Kim (j.kim57@lse.ac.uk)
Swapnil Lal (s.n.lal@lse.ac.uk)

Development Economics Working Group

Savannah Kochinke (s.kochinke@lse.ac.uk)
Edgar Akopyan (e.akopyan@lse.ac.uk)
Somjeeta Chatterjee (s.chatterjee5@lse.ac.uk)
Ardi Janjeva (a.janjeva@lse.ac.uk)
Napat Matchaseth (n.matchaseth@lse.ac.uk)

Political Economy Working Group

Ethan Caspi (ecaspi@usc.edu)
Lutfi Sun (lsun@trinity.edu)
Andre Zeromski (zeromskia20@mail.wlu.edu)
Yihan Zhu (y.zhu39@lse.ac.uk)