University of New South Wales
School of Economics

Honours Thesis

The 19th Century 'Squattocracy' and Contemporary Inequality

Author:
Damoon Sadeghian
Student ID: 3414738

Supervisors:
A/Prof. Pauline Grosjean
Dr. Sarah Walker

Bachelor of Economics (Honours)
and
Bachelor of Commerce

Declaration

I hereby declare that this submission is my own work and that, to the best of my knowledge, it contains no material which has been written by another person or persons, except where acknowledgement has been made. This thesis has not been submitted for the award of any degree or diploma at the University of New South Wales, or at any other institute of higher education.

....

Damoon Sadeghian
28 October 2016

Acknowledgements

First and foremost I would like to thank my two amazing supervisors Sarah Walker and Pauline Grosjean. Sarah, I am incredibly grateful for how warm and welcoming you have been this past year. Thank you for all of the time and effort you spent teaching me how to use GIS software. On behalf of the Honours cohort I apologise for how noisy we have been. Pauline, you were already carrying one baby when we met, and I am sure you did not need the extra workload of another one. Your insights have been invaluable to this thesis, and I am grateful to have had the opportunity to learn from you.

A special thank you must also go to Carlos Pimienta, Gabriele Gratton, and Hongyi Li for their attempts to compete with me in Futsal every week. You tried your hardest, and I appreciate the effort. In all seriousness, I really enjoyed playing against all of you every week. Having observed my physical transformation over the past year, the counterfactual could have been truly horrendous.

To the Honours cohort of 2016, I would not have made it through this year without your companionship and solidarity. We suffered, but we suffered together, and for some reason that is slightly better. A special thanks must also go to Sébastien Printant, Calvin He, Cecilia Chang, Rohan Garga, Kai Zen, and Tran Nguyen, for always lifting my spirits whenever I felt down throughout the year.

To my mum, thank you for always making sure I stayed positive throughout the year. This has no doubt been a stressful period for you as well, and I appreciate the effort you make to always put me first. To Dena and Dave, thank you for all of the support, advice, and feedback you have provided over the past year, both academically and professionally. I always value your opinions, and I truly appreciate you making time to help me out.

Finally, to my incredible girlfriend Eloise, your love, warmth, and kindness has allowed me to get to this point. Being able to come home to you every night gave me the motivation to keep persisting throughout the year. Words cannot express how grateful I am for your undying support.

Contents

List of Tables

List of Figures

Abstract

This thesis investigates the historical origins of inequality in Australia by examining the relationship between squatting settlements in New South Wales and Victoria, and measures of inequality today. Controlling for different geophysical characteristics and other sources of observable difference, I find regions with more squatters historically have higher levels of inequality today. These results are robust to a range of different measures of inequality. I also explore several channels of persistence. I find that the persistence of inequality in these regions can be explained by weaker investment in education and the continuing focus on agricultural production. I propose two complementary explanations. Firstly, regions with large, agricultural landholders have an incentive to maintain the status quo. By limiting public investment in education, they ensure that there is a steady pool of low-skilled workers that they can draw upon. Secondly, historically coercive labour institutions weaken social cohesiveness, which limit collective action towards the provision of public goods. Weaker investment into education translates into inequality of opportunity and inequality of outcome.

Chapter 1

Introduction

Rising inequality is widely considered to be an urgent social problem. Notable leaders around the world, from Barack Obama, to the Managing Director of the International Monetary Fund, Christine Lagarde, have declared inequality to be a top priority (Atkinson, 2015; Mendolia and Siminski, 2016). In terms of economic costs, recent evidence from the OECD, drawing on thirty years of data from OECD countries, suggests that income inequality has a significantly negative impact on growth (Cingano, 2014).

From a theoretical perspective, individuals on the lower tail of the income distribution have a higher marginal propensity to consume than individuals on the upper tail of the income distribution. Therefore, a more equal distribution should lead to greater economy-wide consumption and GDP growth. Furthermore, marginal utility from income is increasing at a decreasing rate. So a more equal distribution of income should increase total social welfare. Inequality also has broader implications. Higher levels of inequality can create a sense of resentment and disenfranchisement in the hearts and minds of large portions of the population. This fragile state can then be exploited by populist leaders, who misinform the public about the causes of their frustration (Piketty, 2016).

Kuznets (1955) first hypothesised that the relationship between long-run development and inequality follows an inverse U-shaped pattern. He argues that in the early stages of development, increased economic growth and development causes an increase in inequality. This continues until a certain peak, at which point the gains from further economic growth and development are distributed more evenly throughout the population. Following this line of argument, highly developed nations today should display similar contemporary inequality outcomes. Yet, a simple comparison of the outcomes of the United States and Sweden destroys this logic. In 2013, the United States had a Gini coefficient of 0.408, while Sweden had a Gini coefficient of 0.25.

Clearly, the causes of differential inequality outcomes are highly complex. It is difficult to compare inequality across countries because there are a plethora of reasons why their contemporary outcomes could be different, including institutional differences, cultural differences, industry specialisation, and differences in education policy. Within-country analyses overcome this black box of institutional difference. In the ideal study, it would be possible to observe *tabula rasa*, and track differential outcomes subject to a shock in a particular region. In the absence of this perfect study, the Australian experience provides an interesting quasi-experiment to analyse.

Prior to the squatters' movement, land in rural New South Wales and Victoria was essentially vacant, aside from the nomadic Indigenous population. Once grassland in these areas was discovered, many settlers decided it would pay to make a *de facto* claim to land outside of the settlement boundaries, rather than to purchase land from the Crown within the settlement boundaries. The individuals that decided to occupy this unused Crown land became known as the squatters. Since the land was vacant, the squatters were able to claim very large plots of land, before any laws or regulations could prevent them from doing so. The highly unequal distribution of land ownership in the squatting regions was simply not possible within the settlement boundaries. The grassland they claimed became the foundation for the highly profitable wool industry.

Wool production increased exponentially, and was the primary driver of economic growth in Australia in the 19th Century (Stapledon, 2013). The squatters quickly became wealthy and highly influential members of the colony. Britain's reliance on Australian wool production resulted in the legal recognition of the squatters' rights to the land. Using their wealth, they were able to hire labourers to do their work, some of which were free men, and some of which were convicts holding tickets-of-leave (McLean, 2016). The labourers were treated harshly (Harris, 1847), creating a class divide between the wealthy large landowners, and their workers.

This thesis exploits the initial inequality in the squatting regions to empirically analyse the long-run persistence of inequality in Australia. Using 19th Century squatting maps, I generate a secondary dataset detailing the concentration of squatters in different regions of New South Wales and Victoria. This allows me to examine the relationship between squatting settlements in New South Wales and Victoria, and measures of inequality today. The empirical strategy follows a selection on observables design. First, I control for a range of geophysical characteristics, such as soil quality and ruggedness, in order to account for land-related factors that could drive inequality outcomes beyond the effect of squatters. In addition, I account for other factors that may affect contemporary inequality, such as mean income, the remoteness classification of the region, and area. The findings suggest that after controlling for geophysical and socio-economic factors, regions with a greater historical incidence of squatters have more inequality today. These results are robust across several measures of inequality including the Gini coefficient, percentile ratios, and top income shares. This suggests that inequality may have historical roots.

I also explore potential channels through which inequality could persist over time. Firstly, squatting regions are more specialised in agriculture today. I hypothesise that this is a result of the incentives the agricultural landed elite have to maintain the status quo in these regions. Secondly, the squatting regions have weaker access to public education. This impedes the opportunities of the working class in these regions.

It is likely that these channels are linked. Since physical and human capital complementarity is relatively low in agriculture, the wealthy have little incentive to lobby for investment in public education (Galor, Moav, and Vollrath, 2009). A complementary explanation for the weaker investment in public schools is that the coercive labour institutions in the squatting regions created an environment of weak social cohesiveness. Previous studies have shown that a divided population has limited potential for collective action towards provision of public goods (Banerjee and Iyer, 2005).

This thesis contributes to a growing body of research providing evidence that weak development outcomes are not just limited to developing countries (Ramcharan, 2010; Rajan and Ramcharan, 2011; Cinnirella and Hornung, 2016). While these other studies have explored the long-run effects of historical institutions on development outcomes in developed countries, they have not explicitly examined the long-run persistence of inequality, which is the primary focus of this thesis. Furthermore, the secondary dataset I generate improves the potential for using the Australian experience to study long-run economic development. This will be valuable for future research.

This thesis proceeds as follows. Chapter 2 reviews the literature relevant to this study. Chapter 3 provides a historical background. Chapter 4 discusses how I generate my secondary dataset, the contemporary datasets I use, and how I match the two together. In Chapter 5 I present results suggesting that there is greater inequality in the squatting regions. In Chapter 6 I explore the channels of persistence. Chapter 7 discusses the results from the previous two chapters, and Chapter 8 concludes.

Chapter 2

Literature Review

Two strands of literature are of particular importance for this study: (1) the origins of institutions, and (2) the channels through which institutions affect long-run outcomes. For the purposes of this thesis, I use the formal definition of 'institutions' provided by Hodgson (2006). He defines institutions as the "systems of established and embedded social rules that structure social interactions".

2.1 Origins of Institutions

Previous works in the literature have shown that early institutional arrangements can explain differences in long-run growth and development paths (Acemoglu, Johnson, and Robinson, 2001; Engerman and Sokoloff, 2002). Colonies that historically had growth-enhancing institutions that protect property rights and promote investment are wealthy today, while colonies that had extractive, rent-seeking institutions are among the poorer countries today. So why are growth-enhancing institutions observed in some regions, while extractive institutions are observed in others?

Acemoglu et al. (2001) argue that this is due to the disease climate of a region at the time of colonisation. Settlement in large numbers was less feasible in the colonial era in regions with a harsh disease climate. High European mortality gave colonial powers the incentive to extract resources from the colony, rather than to promote growth within the colony. Conversely, a forgiving disease climate allowed settlement in greater numbers. These settlers then had incentives to ensure the new colony prospered, through high quality institutions with well-protected property rights.

Engerman and Sokoloff (2002) instead argue that the type of institutions implemented depends on the factor endowments of the colony at the time of colonisation. For example, in Brazil and the Caribbean, the climate was suitable for the growth of stable crops. Stable crops are most efficiently produced in large scale plantations, which are conducive to slave labour. Widespread use of slave labour creates higher economic and political inequality. The resulting institutions in these regions protected the privileges of the landed elite, restricting the participation of the rest of the population. Conversely, in Canada, the climate was not suitable for the growth of stable crops, so slave labour was rare. The resulting institutional environment was characterised by relatively small income gaps, and greater prosperity.

Both studies argue that differences at the time of colonisation lead to differences in the ensuing institutional environment. Differences in institutional environment then have long-run implications for long-run economic growth and development. However, the cross-country nature of their analyses makes it difficult to identify why these institutions persist over time.

2.2 Persistence

Following the seminal works of Acemoglu et al. (2001) and Engerman and Sokoloff (2002), several influential works have sought to explain the channels of institutional persistence. These works generally focus on within-country variation, as it is difficult to explain channels of persistence in the presence of a black box of institutions. This is the approach I choose to take for this thesis.

Banerjee and Iyer (2005) exploit differences in land tenure systems within India to show that institutional overhang can lag development outcomes, even after the original source of institutional difference has been abolished. The British administration defined three alternate systems of land revenue collection in India: (1) a landlord-based system, (2) an individual cultivator-based system, and (3) a village-based system. This defined who was liable to pay taxes to the British. Following independence from the British, differences in land tenure systems disappeared.

Despite this, the authors find that landlord regions systematically underperform relative to non-landlord regions. These differences appear in agricultural investment and productivity, as well as in different measures of investment in public goods. The authors argue that differences in long-run outcomes are caused by the class-based antagonism created within the landlord regions, which persisted well into the post-independence period. This antagonism limited collective action, which resulted in less investment in public goods in the landlord regions.

Similarly, Dell (2010) uses a regression discontinuity approach to examine the long-run impacts of the *mita*, an extensive forced mining labour system in effect in Peru and Bolivia between 1573 and 1812. The author shows that all else being equal, former-*mita* regions have an average level of consumption that is 25% lower than non-*mita* regions. She also shows that incidences of stunted growth are 6 percentage points more prevalent in former-*mita* regions. The difference between regions on either side of the border is that the *mita* regions were historically subject to coercive labour institutions, while the non-*mita* regions were not. The author finds that a significant portion of this disparity in outcomes can be explained by lower levels of education and less developed road networks.

Interestingly, the conclusions from Dell (2010) contradict the hypothesis from Engerman and Sokoloff (2002). Engerman and Sokoloff (2002) argue that regions with high historical inequality tend to invest less in public goods, which leads to weaker long-run outcomes. The non-*mita* regions were characterised by large landowners, and high inequality, and yet it was this region that invested more in public goods. Dell (2010) notes that the implicit counterfactual to large landholders in Engerman and Sokoloff (2002) is secure, small landholders, which were predominant in parts of North America. She argues that this is not an appropriate counterfactual for Peru and many other places in Latin America due to the poor quality, extractive nature of the institutional structures in place prior to the formation of the landed elite. The landed elite in Peru may have actually shielded the lower class from exploitation by a highly extractive state.

The case of the squatters in Australia explored in this thesis exhibits characteristics of both the Indian and the Peruvian examples. The squatters were wealthy, highly influential and owned large plots of land. In this sense, they exhibit similar characteristics to the landlords in India and the large landholders in non-*mita* regions in Peru. While the implied counterfactual from Engerman and Sokoloff (2002) may not be appropriate in the Peruvian case, it does seem to apply in the Australian case. The institutional environment outside of the squatting regions was growth-promoting in nature, much like parts of North America. Therefore, the long-run effect of the coercive labour institutions present in the squatting regions may conform to the predictions from Engerman and Sokoloff (2002).

One potential concern about the above papers is that they all relate to developing countries. The conclusions from these papers may not have external validity to wealthy countries. While this thesis draws on the conclusions of these works, it contributes to a growing body of research suggesting that an elite class can lag the comparative development of a region, even in countries that are wealthy today.

Ramcharan (2010) finds that regions in the United States in which the agricultural elite had disproportionately large landholdings, had significantly less investment in education. Rajan and Ramcharan (2011) find that these regions also had significantly fewer banks per capita. This is consistent with the view that the elite use their influence to entrench their status in society, which manifests in policy through reduced investment in public goods. While these studies do find a negative long-run effect from an initial unequal distribution of wealth, they do not discuss the persistence of inequality outcomes over time.

Fundamentally, the literature suggests that early institutional arrangements can create an elite, landed class, who use their political influence to entrench their status in society. Weaker social cohesion leads to weaker collective action, which leads to weaker investment in public goods. Weaker investment in public goods such as public schools limits the opportunities of the working class to participate in society. This could then allow inequality to persist over time. In this thesis, I explore this possibility.

Chapter 3

Historical Background

This chapter identifies the conditions that allowed the squatters to become so powerful, and provides evidence of their influence in legal and institutional developments.

3.1 Initial Conditions

When European settlers arrived in Australia, they declared the land *terra nullius* (nobody's land), despite the presence of the Indigenous population. While the Aborigines had lived on the land for tens of thousands of years, their hunter-gatherer lifestyle was not recognised by the Europeans. In declaring the land *terra nullius*, it followed that all land in Australia was owned by the Crown.

The Aboriginal economy possessed no stock of man-made reproducible capital to be extracted. Under other circumstances, European colonisation was accompanied by the taxation or diversion of output and trade flows by the occupying powers. Examples of this include Spanish rule in South America and British rule in India. In contrast, the Aboriginal economy did not generate surplus output or taxable reproducible wealth. The one asset of considerable economic potential to the European settlers was the land.

In the early days of New South Wales, settlement was concentrated in the immediate area around Sydney. In 1827 the British Colonial Office (BCO) in London attempted to concentrate settlement within the 'Nineteen Counties' (Alston, Harris, and Mueller, 2012) (see Figure 3.1). Roberts (1964) notes that "the bounds of the colony were more or less set. The Lachlan and the Goulburn plains seemed to impose a natural limit on the south; the Hunter region served for the north; while, to the west, the inhospitable lands beyond Bathurst indicated Wellington Valley, the offshoot since 1823, the outermost post.

Dissatisfaction with quit rent revenues provoked the 1831 Ripon Regulations, which provided for sales of Crown lands in New South Wales. The success of this reform was limited for a few reasons. Firstly, an increase in population led to greater land scarcity within the Nineteen Counties (Roberts, 1964). Secondly, the price of land within the limits was set too high at £1 per acre (Burroughs, 1967).

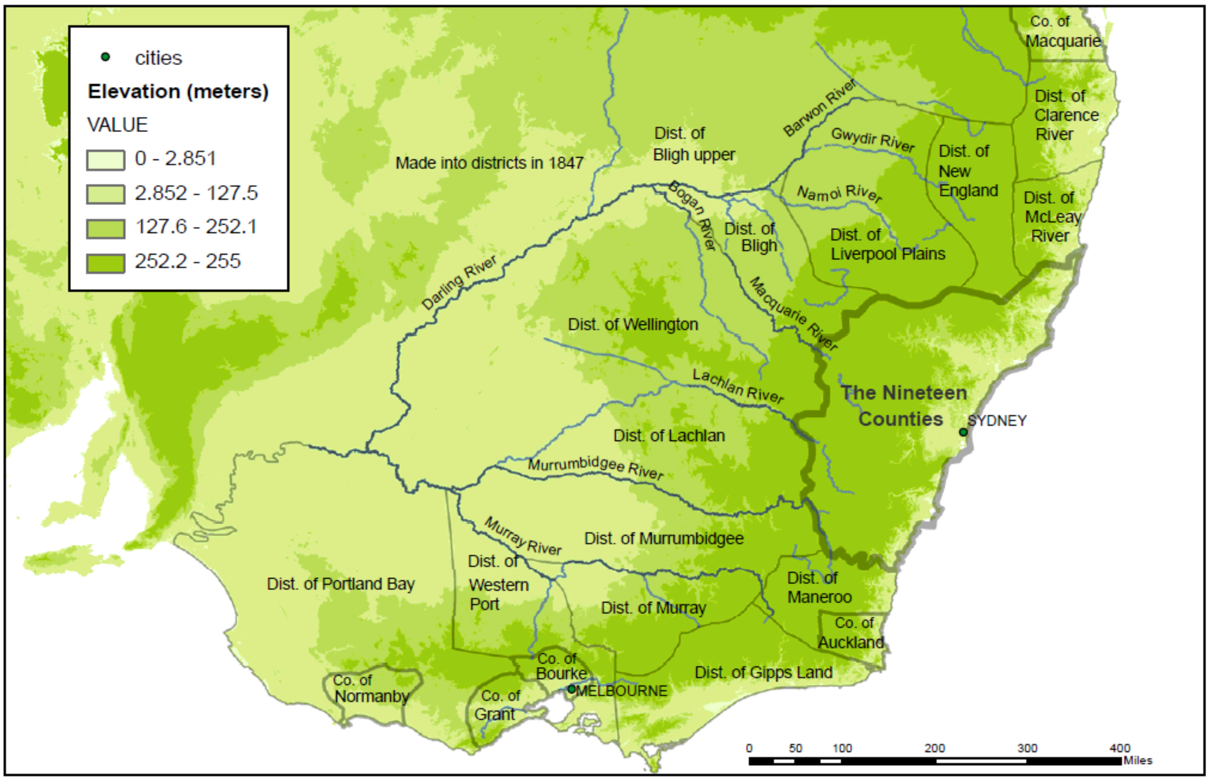

Figure 3.1: The Nineteen Counties and the squatters' districts established by 1844

Source: Dye and La Croix (2013), pp. 913.

Thirdly, the Aborigines engaged in the practice of 'firestick farming'. This involved selectively burning native vegetation to allow fresh plants to grow, which resulted in greater density of native animals (McLean, 2016). The reduced forestation may have improved the productivity of the land, but it may have indirectly helped the settlers, when exploratory expeditions beyond the frontier led to the discovery of vast grasslands suitable for sheep runs. This increased the expected value of land outside of the Nineteen Counties.

Meanwhile, the Aboriginal population declined drastically due to exposure to European diseases (Butlin, 1983). The diminished military power of the Aboriginal population decreased the value of *de jure* property rights. Finally, enforcement of the settlement boundaries was not credible due to inadequate police numbers. Therefore, many found that it would pay to make a *de facto* claim to the land beyond the frontier, rather than purchase *de jure* rights within the settlement boundary.

By the early 1830s a full-blown land rush had emerged. Roberts (1964) notes that "The Counties might spell recognized settlement, but it needed an irrevocable damning of Nature to set up actual limits". This was certainly the case, as a second land rush towards Victoria emerged in 1835.

3.2 The Emergence of the 'Squattocracy'

In the absence of *de jure* protections, institutions emerged to decrease the risk of property loss and cost of security. Incumbent squatters had incentive to protect each other's claims to land because any evidence of conflict increased the probability of all claimants being dispossessed by colonial administrators (Alston et al., 2012). Furthermore, by banding together, the squatters provided a stronger force to limit the threat of the Indigenous population.

These institutional developments made squatting viable, but several other factors contributed to the financial success of the squatters. Firstly, the labour and capital scarce, but land abundant Australian economy resulted in high returns from scarce factors of production. Land was essentially free for the squatters. The capital inputs necessary to commence producing wool were minimal. The one significant cost, sheep, were self-reproducing. The natural features of the land meant that there was no need for fencing. Sheep were allowed to wander with some monitoring. Therefore, the upfront costs of farm formation, that normally act as a barrier to entry, were minimal. Secondly, the British tariff on wool imports was eliminated in 1825, and the price of wool more than doubled between the mid-1820s and the mid-1830s (Dye and La Croix, 2013).

Thirdly, the industrial revolution had taken full effect in Britain. A booming textiles industry created strong demand for Australia's merino wool, which was considered to be premium quality. The entrepreneurial activity of John Macarthur, combined with the high value-to-bulk ratio of wool, proved that the distance between Australia and Britain was no boundary (McLean, 2016).

Finally, the unique characteristics of the early-Australian economy provided a cheap source of labour: convicts (I will discuss the labour conditions in more detail later). These factors combined drove the success of the squatters. Figure 3.3 shows how quickly the pastoral industry grew. In 1831, Britain imported wool primarily from Germany (71%) and Spain (11%). They only imported 8% of their wool from Australia (McLean, 2016). By 1850, Australian wool became dominant in the British market, accounting for 53% of their wool imports.

Furthermore, the production of wool in Australia increased exponentially. Per capita wool exports increased from 6 to 113 pounds between 1830 and 1850. Total wool exports are expressed per capita to account for the large growth in population during this period.

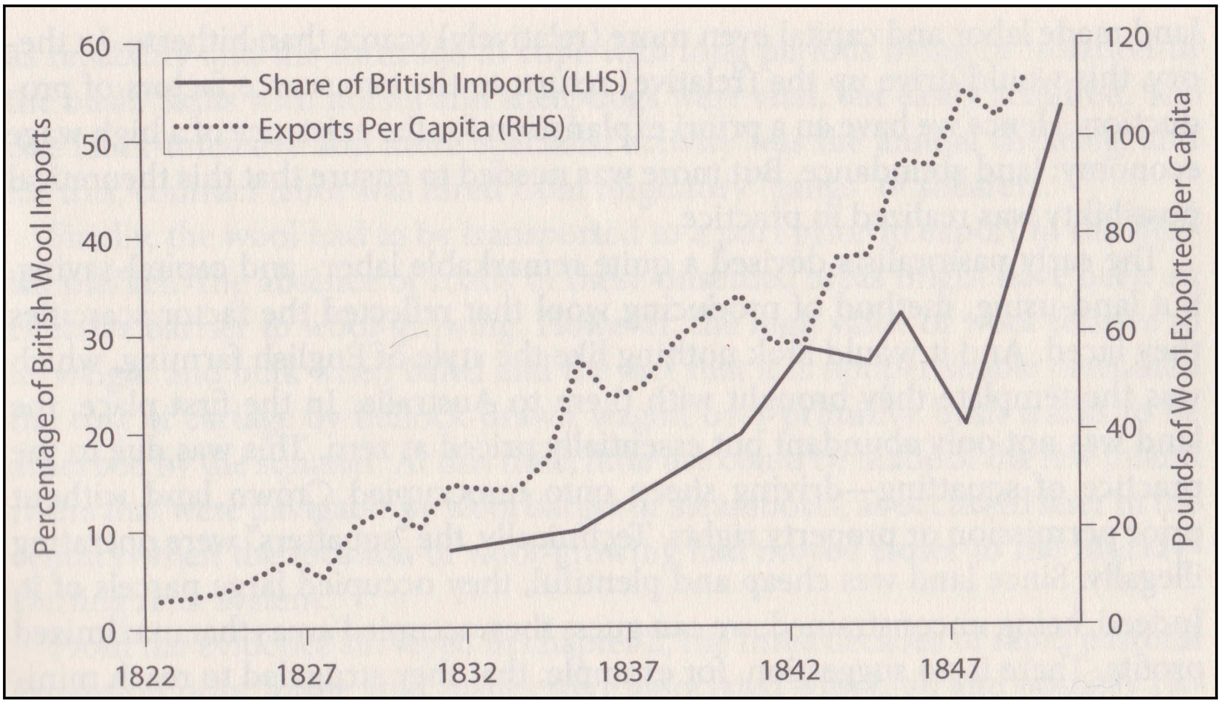

Figure 3.3: Wool trade 1822-1850

Source: McLean (2016) p.59

As the value of land beyond the settlement boundaries increased, there was greater incentive for a revenue-seeking government to extend the boundaries. Naturally existing claimants would mount heavy resistance to such changes. Pressures began to emerge in the 1830s as the Colonial Office, Edward Wakefield, and certain local interests pressed the NSW government to reassert their claim on what should have been public land. Despite resistance, in 1836 an Act was passed requiring squatters to purchase a license for an annual fee of £10 (Dye and La Croix, 2013). While this legitimised their claim to the land, it was not a lease. They did not have conventional *de jure* rights to the land specifying the boundaries and location of their claims. Therefore the licenses were of limited value to the squatters.

In 1839, another Act drastically diminished the rights of the squatters. It provided district commissioners with the power to "remove a squatter's license without appeal, to remove or destroy stock, to define the boundaries of the sheep run, to investigate charges of violence against Aborigines, and to collect license fees and newly imposed taxes on stock" (Dye and La Croix, 2013). Furthermore, the Act armed district commissioners with a mounted police force to enforce these changes.

There were two sides to this story. The squatters argued that the 1839 Act was unfair as it gave the government an annual option to terminate their licenses without compensation for improvements they had made to the land. They had settled an unknown region with no help from the government and all they had to show for it was an annual license. Therefore they demanded a longer license or lease, with some consideration for improvements made on the land (Roberts, 1964).

From the perspective of the government, three million acres of what should have been public land had been given up for a small annual payment. They had to reach some form of compromise. The governor of the time, Governor Gipps, recognised that without more secure rights, the squatters had no incentive to make permanent improvements to the land. He also understood the power and influence of the squatters when he wrote to Lord Stanley (Roberts, 1964): "It is most essential to bear in mind that the Squatters form by far the most powerful body in New South Wales - that in fact, almost anybody who has any property at all, is a Squatter. This, I would submit, is the chief ground on which the interference of Parliament might be called for, should the Squatters at any time seek to appropriate to themselves the Lands of the Crown".

In 1844, Governor Gipps proposed his Occupation Regulations. These regulations would have required squatters to purchase 320 acres of each sheep run at £1 per acre as a homestead to qualify for an eight-year license (Dye and La Croix, 2013). The license, which required an additional fee, would have given squatters the right to use the remainder of their sheep runs. As Gipps attempted to implement these regulations, he faced heavy opposition from not only the squatters, but also opposition in London.

The squatters and other stakeholders in the wool industry formed an alliance to oppose the regulations. They were supported by a coalition of English and Scottish woolens manufacturers, shipping companies, and other groups with linkages to Australian wool production. They joined forces to lobby members of Parliament. By mid-1846, they were rewarded for their efforts as the Australian Waste Lands Act was passed. This established *de jure* rights to *de facto* claims and provided security of tenure for the squatters. The Act gave squatters 14-year leases, and preemptive rights to purchase at least 160 acres at £1 per acre. Only the occupying tenant had rights to purchase the land for the duration of the lease (Dye and La Croix, 2013). Furthermore, rights of compensation for improvements were recognised.

The squatters had effectively stolen public land, become very wealthy through their technically illegal claim to land, and used their influence to have their rights to the land recognised. The formation of the 'Squattocracy' was complete, but new threats were on the horizon.

3.3 Squatters v Selectors

The discovery of gold in 1851 provided a shock to the Australian economy. Prior to this discovery, the Australian economy was primarily driven by the pastoral industry. The gold rush created mass migration to Australia, leading to a dramatic increase in population. In 1850 the non-indigenous population was approximately 400,000. By the end of the decade, it had reached almost 1.1 million (McLean, 2016). The immigrants were disproportionately male and of working age, which led to a reduction in the dependency ratio.

When the gold rush subsided towards the end of the decade, unemployment increased and there was out-migration following the discovery of gold in New Zealand. The colonial government needed to provide incentive to keep productive workers in Australia. They deemed the best way to retain the population was to make land available to potential farmers.

During this same period, political changes divided New South Wales into three distinct colonies: New South Wales, Victoria (1850) and Queensland (1859). Each colony had its own government, with the right to manage their own affairs, including land policy. Furthermore, the franchise was extended to ex-gold miners, reducing the political influence of the squatters.

The increasing political power of the new claimants led to land reform in New South Wales (Robertson's Act, 1861) and Victoria (Grant's Act, 1865). The new legislation allowed any individual to select up to 320 acres anywhere in the colony except land under leasehold. These individuals became known as the selectors. There were several conditions imposed on the selectors: a payment of a 25% deposit, one year of residence on the selected land, and improvements to the land of £1 per acre (Alston et al., 2012). Once leases on squatting land had expired, the government gave squatters preemption rights to secure 640 acres out of every 16,000 acres held under lease. The rest of the land was open to selectors. However, by using their preemption rights strategically, squatters could monopolise access to water (Roberts, 1964). This effectively made it impossible to claim any of the squatter's property.

By the mid-1880s, only 27% of selectors remained on their selected land. Selectors were successful in areas in which small-scale agriculture was more viable. In semi-arid and arid zones, only the pastoralists proved to be profitable (McLean, 2016). The land was simply not suitable for the selectors. Furthermore, the squatters were wealthier and had greater violence potential (Alston et al., 2012). It is easier for a wealthy ruling elite to defend their land than it is for small-scale farmers to usurp the wealthy. The resource advantage of the squatters allowed them to hire men to file claims on their behalf and to exploit legislative loopholes, which helped them entrench their wealth.

3.4 The Class Divide

Aside from the brief gold rush in the 1850's, the wool industry was the main driver of productivity and economic growth in Australia over the course of the 19th Century (Stapledon, 2013). Since they owned the means of production, the majority of the gains over this period were accrued by the squatters. Their wealth afforded them the luxury of employing workers to do the majority of the dirty work. This included shepherds, hut-keepers, watchmen, and labourers. These workers were either free men or convicts holding tickets-of-leave.

McLean (2016) notes that convict labour in Australia during this period fell in a spectrum of coercive labour institutions. It was not quite as extreme as slavery in the United States, but the conditions were still harsh. New South Wales had a dual labour market partly made up of a free labour market and partly from a government-directed labour market. In the free labour market, labour was supplied by free men and convicts (McLean, 2016).

Convicts received wages below the market rate for free labour. Those who were well regarded as workers were given a ticket-of-leave, which suspended their sentence subject to good behaviour and performance for the employer they are assigned to. A failure to meet these requirements resulted in a return to convict status. Therefore, convict labour was generally quite productive. This cheap source of efficient labour proved to be very profitable for the squatters.

Working conditions at the squatting stations were harsh, particularly for the convicts. The relationship between the squatters and their workers was akin to that of a master and a servant. For example, hutkeepers and watchmen were required to watch the sheep by night. If any sheep were robbed or attacked by dogs at night, the worker would lose their right to pay if they were a free man. If they complained, the 'master' would tell them to take their worries to court. They had confidence that the courts would rule in their favour if any such matter were to be presented in court. If the man were to be a convict, they would be given the present of a 'red shirt' (a scarified back) (Harris, 1847). It is no stretch to describe the labour institutions in these regions as coercive.

It should be clear by now that the squatters were wealthy and highly influential. They accrued the majority of the gains in productivity over the 19th Century. However, this did not extend to the working class labourers in these regions, which opened up a distinct class divide that was not present in other regions in Australia. As has been shown in the previous chapter, divisions in society are not conducive to strong collective action, which limits investment in public goods. For this reason, it is likely that there are long-run implications from the class divide in these squatting regions.

Chapter 4

Data

This section outlines the datasets used for this thesis. It also provides an explanation of the methodology I used to combine them.

4.1 Data Generation

To identify the locations where squatters were concentrated, I generate a secondary dataset using georeferenced historical squatting maps of New South Wales (Reuss, Browne, and Degotardi, 1860) and Victoria (Ham, 1854). Figure 4.1 below shows the squatting map of New South Wales.

Figure 4.1: Historical squatting map of New South Wales

Each polygon in the map represents the claim to land made by a squatter. Through correspondence with the National Library of Australia, I discovered that the size of a polygon does not necessarily depict the actual size of the plots of land the squatters claimed. To overcome this inaccuracy, I simply plot a point in the middle of each polygon using GIS software. The points represent squatting stations. The squatting stations can be thought of as the base of production for the squatters. By plotting each point, I record the latitude and longitude of each squatting station, which will be of use later. Following this process, I find 1601 unique squatting stations in total in New South Wales.

I follow a similar process for the historical squatters' map of Victoria. However the Victorian map denotes squatting stations in a different way, as seen in Figure 4.2.

Figure 4.2: Historical squatting map of Victoria

Each squatting station is denoted by a small circle with a number next to it. The red rectangle in Figure 4.2 isolates one particular squatting station to make this a bit more clear. To plot all the squatting stations in Victoria, I start by finding a squatting station with the number 1 next to it. I then follow the trail of numbers until I cannot find the next number. The trail of numbers resets for different regions within the map so there are duplicates of the same number. I follow each number trail until I find and plot every squatting station in Victoria. I find that there are 958 squatting stations in total in Victoria. Combining this with the points plotted in New South Wales, there are 2559 unique squatting stations in total in New South Wales and Victoria.

After plotting the locations of the squatting stations, it is necessary to transform the data into a form I can work with to merge with my contemporary datasets. I use shapefiles from the Australian Bureau of Statistics' (ABS) Australian Statistical Geography Standard (ASGS). The shapefiles provided in the ASGS provide statistical boundaries for the whole of Australia. They vary in their level of aggregation. The most aggregated level of the ASGS breaks Australia into states and territories. The next level down is Statistical Area 4, then Statistical Area 3, Statistical Area 2, Statistical Area 1, and Mesh Blocks. To put this more simply, each Statistical Area 3, for example, is comprised of several Statistical Area 2 regions. The ASGS also provides shapefiles for non-ABS structures that are commonly used for datasets, but are not defined and controlled by the ABS. This includes Postal Areas (POAs) and Local Government Areas (LGAs).

The level of aggregation I use is critical. The more aggregated the data is, the less informative the location of the squatting station becomes. So ideally I would merge the location of the squatting stations with the Mesh Blocks shapefile. However, the contemporary data I use is not available at such a level of aggregation. For this reason, I aggregate the data at Statistical Area 2 (SA2) and check the robustness of my results by also aggregating at the LGA level. Figure 4.3 and Figure 4.4 show maps of New South Wales and Victoria with the SA2 and LGA boundaries respectively. Figure 4.5 and Figure 4.6 show the same maps, but also include the plotted locations of all of the squatting stations.

SA2s generally have a population ranging from 3,000 to 25,000 persons, with an average population of approximately 10,000 persons. An example of a SA2 would be 'Kensington-Kingsford' or 'Paddington-Moore Park'. In urban areas SA2s are generally either large suburbs or a combination of small suburbs, whereas in rural areas, SA2s are much larger. LGAs include things like local councils, districts shires etc. Examples of LGAs include the 'City of Sydney' or 'City of Randwick'.

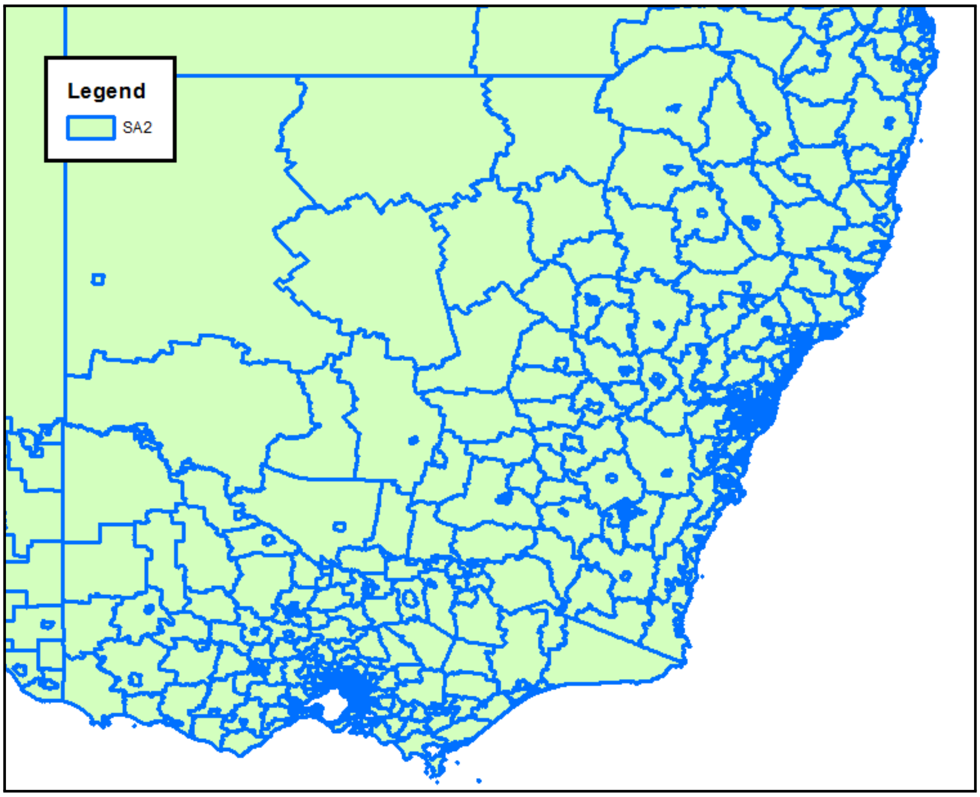

Figure 4.3: New South Wales and Victoria - Statistical Area 2

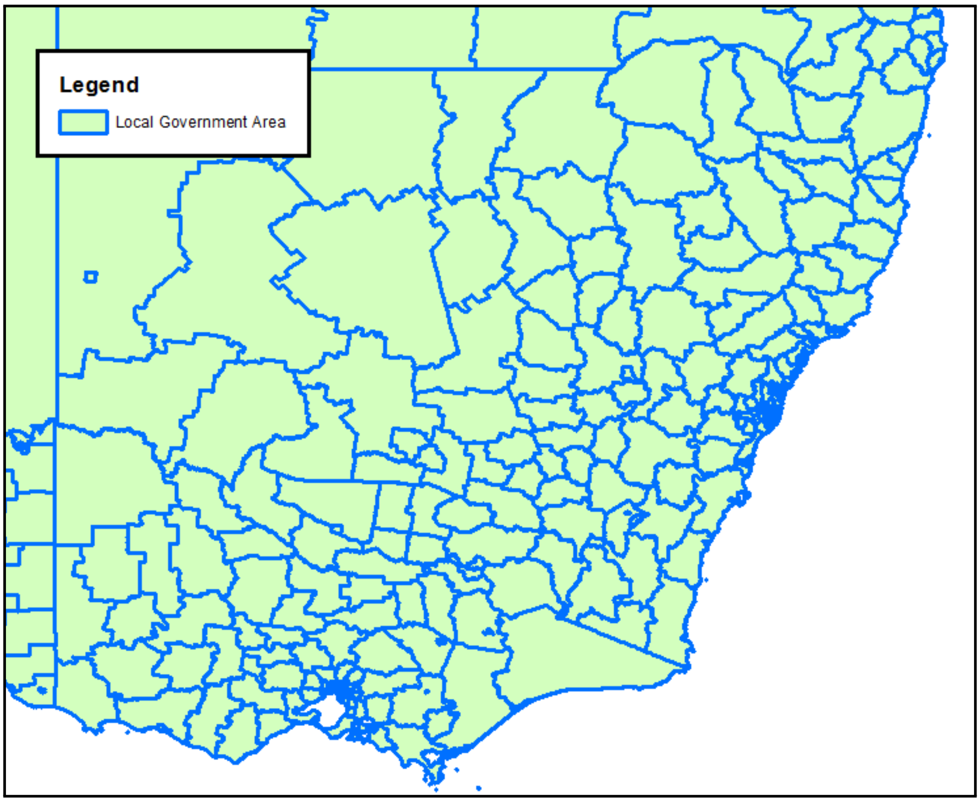

Figure 4.4: New South Wales and Victoria - Local Government Areas

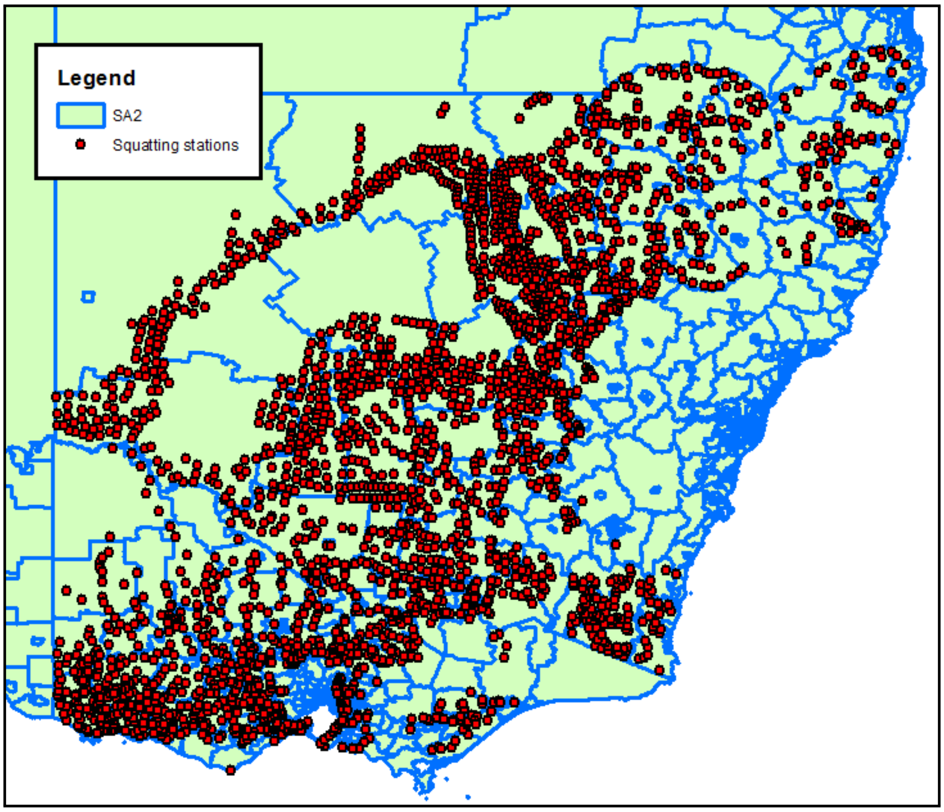

Figure 4.5: New South Wales and Victoria - Statistical Area 2 with location of squatters

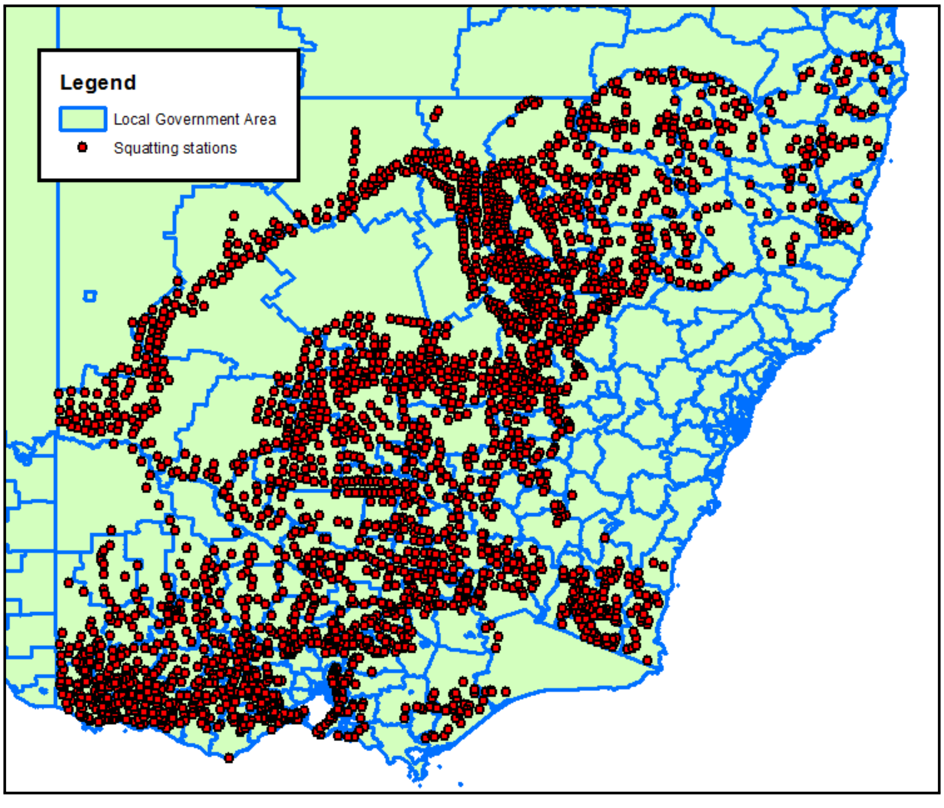

Figure 4.6: New South Wales and Victoria - Local Government Areas with location of squatters

4.2 Contemporary Data

After assigning the squatting station observations SA2 and LGA codes, I am able to merge it with contemporary datasets. I require several different datasets to investigate the possible historical origins of inequality of Australia. This section will explore the contemporary datasets I use and describe the characteristics of these datasets. Specifically, I use datasets relating to inequality, public good provision, soil quality, geophysical characteristics, geological features, and access to water.

4.2.1 Inequality

To measure inequality, I use the ABS Estimates of Personal Income for Small Areas, 2012-13 dataset, which is available at both the SA2 and LGA levels, and is derived from taxation data. The measures of inequality I extract from the data include the Gini coefficient, percentile ratios, and top income shares.

I primarily use the Gini coefficient for my analysis because it satisfies several desirable properties for a measure of inequality including the Pigou-Dalton Transfer Principle (Dalton, 1920; Pigou, 1912), Income Scale Independence, Principle of Population, and Anonymity (Litchfield, 1999) (see Appendix B for more information on this). The Gini coefficient takes on values between 0 and 1 with zero interpreted as perfect equality. It is the most commonly used and most easily recognisable measure of inequality, but it has its limitations. Piketty (2014) argues that it is impossible to summarise in a single numerical index all that a distribution can tell us about inequality. It is hard to tell where exactly along the distribution the inequality is most concerning. He also argues that synthetic indices like the Gini coefficient abstract the issue and are not easily interpretable. For example, how much more unequal is a society with a Gini coefficient of 0.430 compared to a society with a Gini coefficient of 0.395?

For this reason, I also use percentile ratios and top income shares. The percentile ratios are calculated by sorting all the incomes in ascending order. Each observation is then sorted into a percentile. To calculate the P80/P20 ratio for example, the top income of the 80th percentile is recorded in the numerator, and the top income of the 20th percentile is recorded in the denominator. So by this measure, the higher the P80/P20 ratio, the more unequal a given region is. The specific percentile ratios I use are the P80/P20, P80/P50, and P20/P50. It is easy to see that the P20/P50 ratio follows the opposite pattern to the P80/P20. The higher the P20/P50 ratio, the more equal a given region is.

The top income shares are simply the share of total income earned by the rich. So for example, the top 1% measure is the share of total income earned by the top 1% of people in a given region. Generally if the share of income earned by the very rich is higher, the region is more unequal, but it is important to note that this is only a small subset of the population, which ignores the rest of the distribution. The top income shares I use are the top 1%, top 5% and top 10%.

Notably, these measures of inequality are derived from the taxation statistics provided by the ATO. There are several advantages to using measures derived from tax data. Firstly, the tax data takes into account different sources of income including wages and salaries, investment, unincorporated business income, superannuation and annuities, and other income sources. Considering the wealthy tend to derive their total income from a variety of sources, and not just wages and salaries, it is important for a measure of inequality to take this into account.

However, the taxation statistics do have shortcomings that need to be recognised (Leigh, 2005). Firstly, in Australia the taxation unit is the individual. The distribution of income among individuals may not necessarily provide an accurate proxy for the distribution of income among households. Secondly, not everyone files a tax return. For example, individuals earning less than the tax-free threshold are not required to fill out a tax return. Therefore the taxation statistics may not accurately capture low incomes.

The alternative to using data derived from taxation statistics is to use data derived from the Census or HILDA. By definition, the Census is representative of the population, but the fact that incomes are grouped into brackets cannot be reconciled. HILDA is able to describe incomes more accurately than the Census but it suffers from a high degree of sampling error when working at the SA2 level due to the limited number of observations per SA2. For these reasons, I choose to use data derived from the taxation statistics.

4.2.2 Public Good Provision

I hypothesise that regions with historically more squatters will have less investment in public goods than regions with historically less squatters. To test for this, I require data on a variety of public goods. This data comes from the Regional Australia Institute (RAI) LGA Infrastructure and Essential Services Indicators 2011. From this dataset, I extract data that acts as a proxy for investment in education, health, roads, and rail. This dataset is only available at the LGA level.

The education data includes average distance from primary schools, and average distance from secondary schools. I use average distance to school as a measure of access to educational services, which represents the level of educational investment for a given region. The health data includes access to hospital services, access to GP services, and access to allied health services. Access to hospital services is measured as the number of people employed in hospitals in a region per resident population. Similarly access to allied health services is measured as the number of people employed in health services excluding hospitals in a region per resident population. Access to GP services is measured as the number of GP services per capita in 2011. Road infrastructure is measured as the distance from the region's business centre to the nearest major highway. Rail infrastructure is measured as the distance to the nearest railway station per service.

4.2.3 Soil Quality

Soil quality data is sourced from the Food and Agriculture Organisation of the United Nations, Global Suitability for Rain-fed Crops Excluding Forests. This dataset is available in raster form. Using grids, suitability scores are assigned to different regions based on land characteristics. There are 1,897,295 grid points around the world. Using GIS software, I merge the raster with the SA2 and LGA shapefiles. This assigns each LGA and SA2 a mean, median, mode, minimum and maximum score for the grids within the boundaries of the LGAs and SA2s. Table 4.1 summarises the different categories possible for the grids.

Table 4.1: Soil suitability for rain-fed crops categories

Category	Score	Rating
0	-	Undefined
1	$SI > 85$	Very high
2	$70 < SI \leq 85$	High
3	$55 < SI \leq 70$	Good
4	$40 < SI \leq 55$	Medium
5	$25 < SI \leq 40$	Moderate
6	$5 < SI \leq 25$	Marginal
7	$0 < SI \leq 5$	Very marginal
8	SI = 0	Not suitable
9	-	Water
10	-	Dominantly forest ($SI > 25$)
11	-	Dominantly forest ($SI < 25$)

For this thesis, I take the mode category to represent the soil quality for a given SA2 or LGA. Since the scale of soil suitability is only represented by categories 1-8, I cannot simply take the mean value for a given SA2 or LGA. This value could be skewed by the presence of water or forest in a region. After taking the mode, I create a dummy variable that is equal to 1 if the soil is good or 0 if the soil is bad. I classify any region with a mode category of between 1 and 5 as good soil. For every other value, I deem the soil in the region to be unsuitable.

4.2.4 Geophysical Characteristics

Data describing the geophysical characteristics of the land is obtained from Nunn and Puga (2012). The authors create a variable measuring the ruggedness of terrain. Like the soil data, this dataset is available as a raster. The authors calculate terrain ruggedness by taking a point on the earth's surface, like the filled black circle in Figure 4.7, and calculating the difference in elevation between this point and the point on the grid 30 arc-seconds north of it, represented by the hollow circle directly above it. This same calculation is carried out for each of the major compass directions (north, northeast, east, southeast, south, southwest, west, northwest). Terrain ruggedness is then calculated as the sum of the squared differences in elevation between the filled black circle and each of the eight adjacent points.

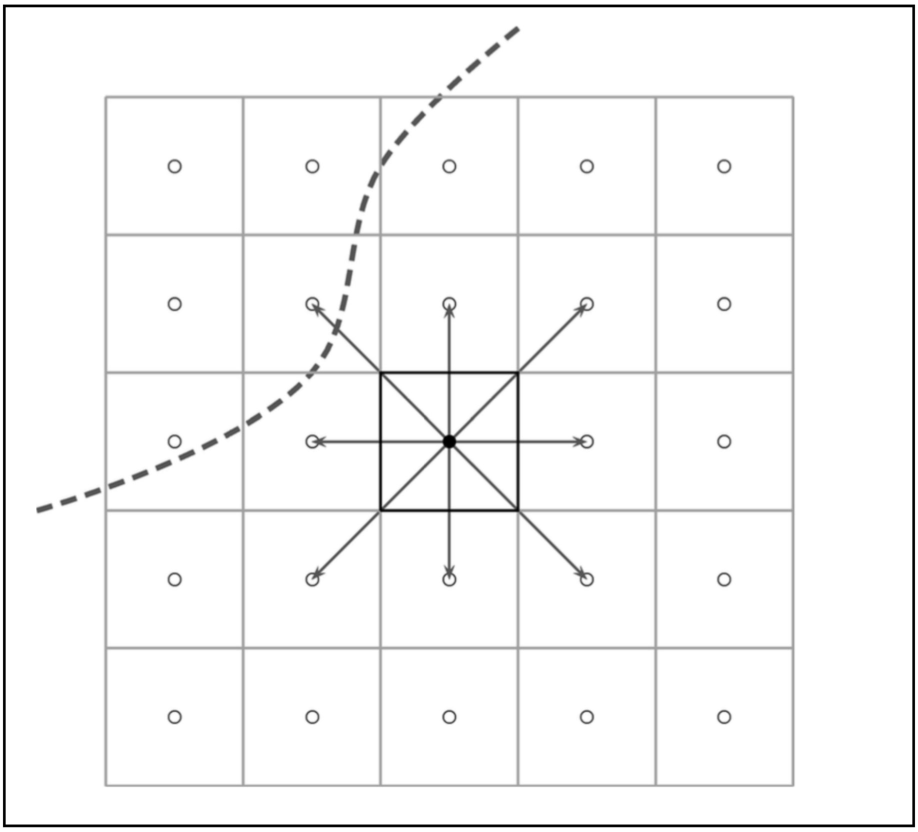

Figure 4.7: Schematic of the terrain ruggedness calculation

Source: Nunn and Puga (2012)

I use GIS software to merge the ruggedness data with the SA2 and LGA shapefiles. By merging the datasets, a given SA2 or LGA will be given a ruggedness value of the mean of all the grid points contained within its boundaries. This gives me a measure of the terrain ruggedness of each SA2 and LGA in New South Wales and Victoria.

Data characterising the gradient of terrain and the mineral endowments of a region is sourced from Grosjean and Khattar (2014). Gradient of terrain is split into five categories, with category 1 representing low gradient and category 5 representing high gradient. Mineral endowment categories are summarised in Table 4.2 below:

Table 4.2: Mineral endowment categories

Category	Mineral deposit
0	No minerals or traces
1	Minor coal
2	Minor others
3	Major coal
4	Major copper
5	Major gold
6	Major mineral sands
7	Major oil and gas
8	Major other

Each SA2 and LGA is assigned the terrain gradient and mineral endowment category that is most representative of the land within its boundary.

4.2.5 Access to Water

Data detailing the locations of bodies of water comes from two sources. The FAO's Global Suitability for Rain-fed Crops Excluding Forests data also includes bodies of water, as shown in Table 4.1. However, this data is limited to just lakes and other still bodies of water. I supplement this with Australia's River Basins, from Geoscience Australia. This details the locations of river streams around Australia.

For the purposes of my thesis, I create a dummy variable for whether or not there is a body of water within a digital boundary. I also account for the distance from the centroid of a digital boundary to the nearest river.

4.2.6 Balance of Covariates

A major concern for this study is that regions with squatters could be fundamentally different from regions without squatters. The regions the squatters decided to settle are likely to be subject to selection bias, based on the characteristics of the land. This means that the squatting regions were not determined randomly. This makes it difficult to isolate the long-run effect of the historical conditions in the squatting regions. However, controlling for differences in observable characteristics allows me to treat the decision to squat in these regions 'as good as randomly assigned'.

Table 4.3 reports the balance of covariates. Panel A shows that the squatters were generally concentrated in more remote regions, denoted by the signs of the estimates on Urban, Outer Regional, and Remote Australia. There were significantly more squatters in outer regional and remote Australia, but no significant difference in inner regional and very remote Australia.

Panel B reports differences in land characteristics. The estimates suggest that regions with squatters had significantly less instances of medium grade terrain. However, for the most part terrain in Australia is quite flat. There does seem to be some variation in mineral endowments. Squatting regions have significantly less coal, but more copper and mineral sands. Squatting regions also tend to have higher quality soil. This relationship is significant at the 1% level. This suggests that the squatters were able to select the most fertile regions.

Panel C shows that bodies of water are more common around squatting settlements, but squatting settlements are generally further away from rivers. This is likely due to the size of the SA2s in squatting regions. Finally, Panel D shows the squatting settlements have lower income and larger land on average.

Clearly there are differences between squatting regions and non-squatting regions. After controlling for these differences, I can be more confident that any residual effect is due to the historical unequal distribution of wealth in the squatting regions.

Table 4.3: SA2 - Balance of covariates

Variable	Number of squatting stations	Mean	Robust standard error	t-stat	Observations
Panel A: Remoteness and State					
Urban	-0.0142***	0.595	(0.002)	-8.5702	748
Inner Regional	-0.0007	0.277	(0.001)	-0.7502	748
Outer Regional	0.0103***	0.12	(0.002)	5.2032	748
Remote Australia	0.0029**	0.005	(0.001)	2.3130	748
Very Remote Australia	0.0016	0.003	(0.001)	1.5469	748
Victoria	-0.0033***	0.469	(0.001)	-2.9978	748
Panel B: Land characteristics					
Ruggedness	-0.1164	39.434	(0.125)	-0.9318	748
Low grade	0.0000	0.9479	(0.000)	0.1139	748
Low to medium grade	0.0004	0.0053	(0.000)	1.3017	748
Medium grade	-0.0005*	0.0468	(0.000)	-1.7167	748
No minerals or traces	-0.0001	0.004	(0.000)	-1.6201	748
Minor others	-0.0000	0.001	(0.000)	-0.9928	748
Major coal	-0.0079***	0.611	(0.002)	-3.9538	748
Major copper	0.0042***	0.0348	(0.001)	3.0379	748
Major gold	0.0007	0.3008	(0.001)	0.5453	748
Major mineral sands	0.0029**	0.0468	(0.001)	2.4204	748
Major oil and gas	0.0002	0.0013	(0.000)	0.9763	748
Soil Quality	0.0047***	0.659	(0.001)	3.8859	748
Panel C: Access to water					
Body of water	0.0078***	0.075	(0.001)	5.7185	748
Distance to river (km)	0.2968***	8.797	(0.057)	5.2105	748
Panel D					
Mean Income ($1000)	-0.2709***	55.5	(0.041)	-6.6630	743
Area (km^2)	394.0730***	1310.848	(95.835)	4.1120	748

Robust standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Chapter 5

Results

I combine all the datasets in Section 4 into two distinct datasets: one for analysis at the SA2 level, and one for analysis at the LGA level. With these datasets, I investigate if a relationship exists between historical concentration of squatters and contemporary inequality. I conduct my initial analysis using the data at the SA2 level as it provides more precision than the LGA data. I check the robustness of my results using data at the LGA level.

Figure 5.1 illustrates the relationship between the number of squatters historically in a SA2 and the Gini coefficient. Since there are many observations equal to zero in this plot, I also include a correlation plot omitting these observations, as they may skew the relationship. This is shown in Figure 5.2. Each plot represents the Gini coefficient for a given SA2. The positive relationship suggests that regions with historically more squatters tend to be more unequal today.

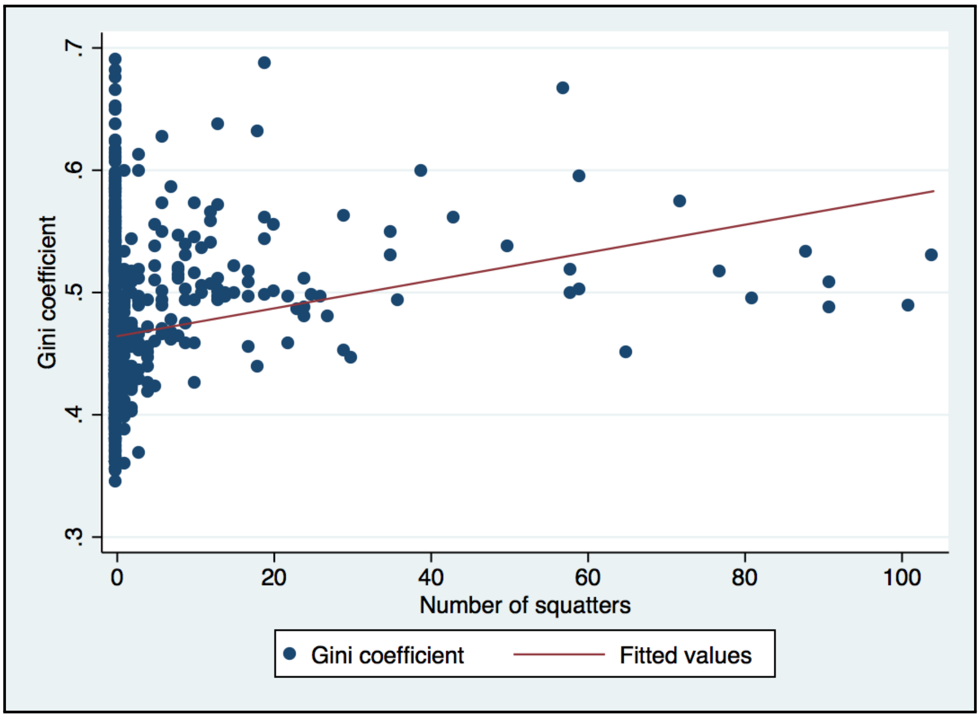

Figure 5.1: Correlation plot: Gini coefficient and historical number of squatters in SA2

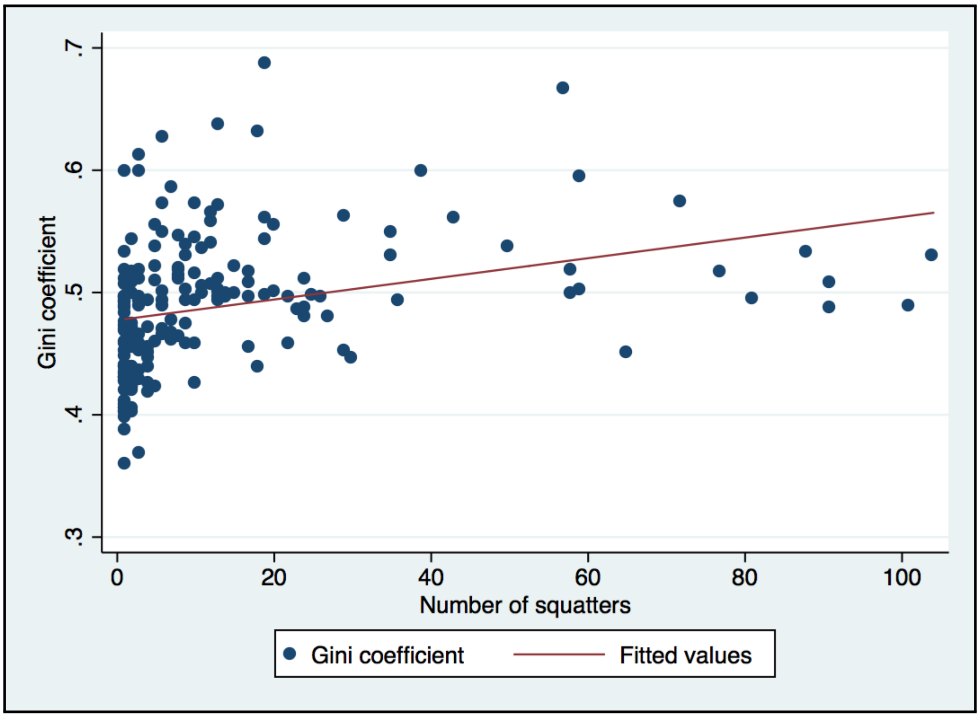

Figure 5.2: Correlation plot: Gini coefficient and historical number of squatters in SA2 - omitting regions with no squatters

Similarly, Figure 5.3 and Figure 5.4 show the same relationships at the LGA level.

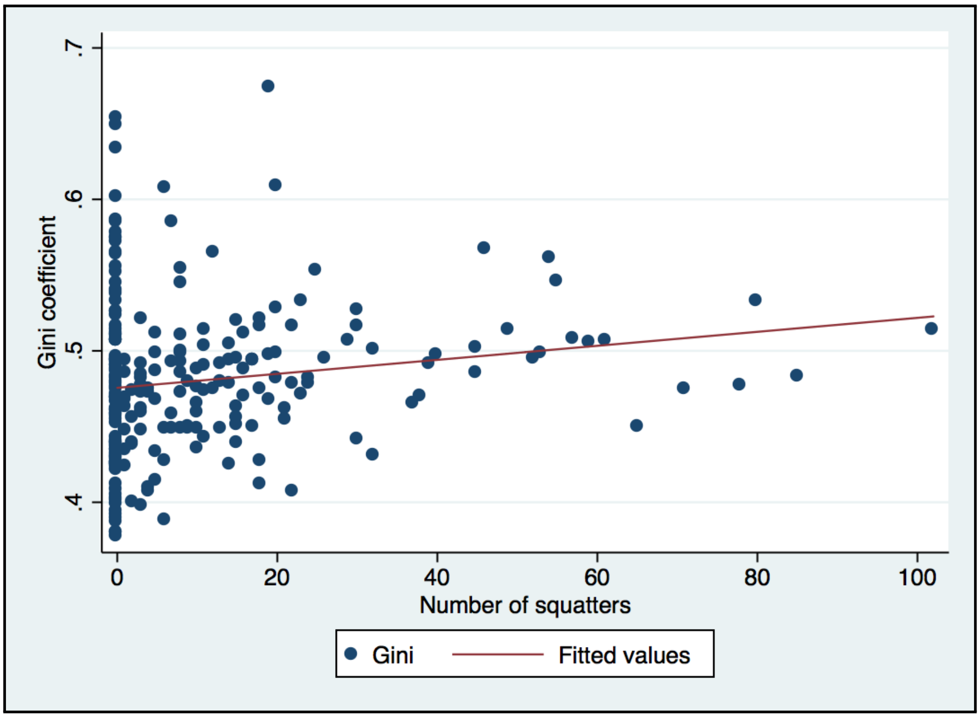

Figure 5.3: Correlation plot: Gini coefficient and historical number of squatters in LGA

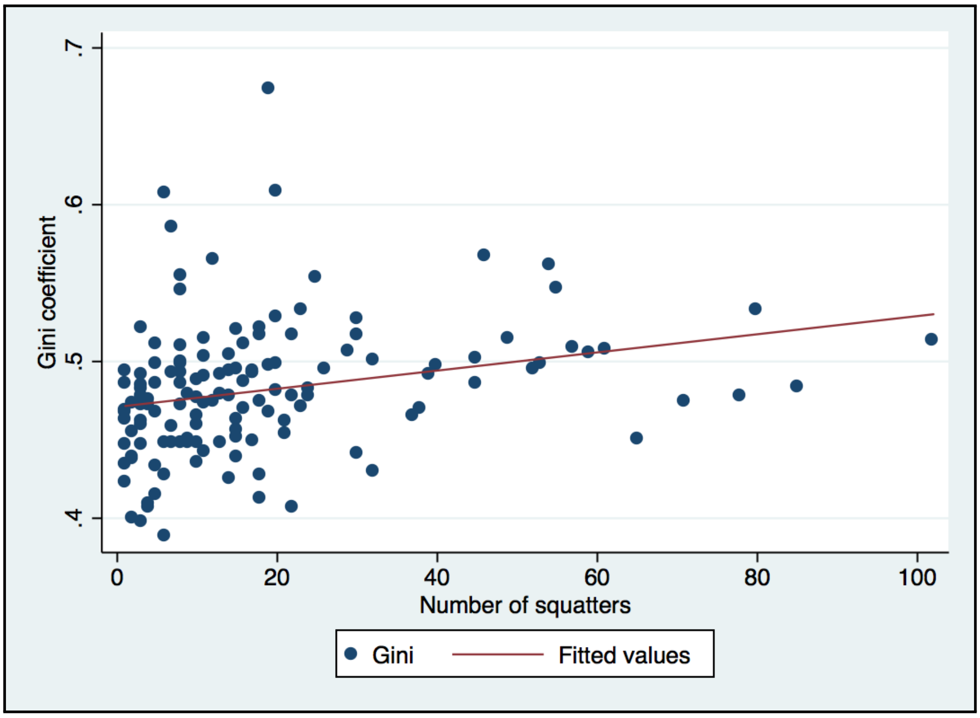

Figure 5.4: Correlation plot: Gini coefficient and historical number of squatters in LGA - omitting regions with no squatters

While the relationship between the historical number of squatters in a LGA and the Gini coefficient is positive, Figure 5.3 and Figure 5.4 seem to show that the relationship is less positive than the relationship at the SA2 level. However this is to be expected. The greater the level of aggregation, the more information regarding the location of squatting stations is lost. The relationship should become more blurred as the data becomes more aggregated.

To formally test these relationships I use OLS based on the following specification:

$$y_i = \alpha + \beta S_i + \boldsymbol{\gamma} \boldsymbol{X_i} + \epsilon_i \tag{5.1}$$

The outcome of interest y_i is some contemporary measure of inequality for a region i. The primary measure I use is the Gini coefficient, but I test if the results are robust to the P80/P20 ratio, P80/P50 ratio, P20/P50 ratio, top 1% income share, top 5% income share, and top 10% income share. The variable of interest S_i is the number of squatting stations historically in region i. This captures the concentration of the squatters in a region. $\boldsymbol{X_i}$ is a vector of controls including mean income, area, geophysical characteristics, soil quality, access to water, remoteness dummies, and a state dummy. To account for heteroskedasticity, I use Huber-White heteroskedasticity robust standard errors for all specifications. Descriptive statistics for the SA2 dataset are reported in Table 5.1.

Table 5.1: SA2 - Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Inequality measures					
Gini coefficient	743	.468	.061	.344	.689
P80/P20	743	5.12	1.186	3.16	12.7
P80/P50	743	1.873	.187	1.52	2.98
P20/P50	743	.377	.05	.15	.49
Top 1%	743	7.448	3.029	3.6	22.9
Top 5%	743	20.271	4.913	13.2	44.7
Top 10%	743	31.421	5.445	22.8	57.5
Variable of interest					
Number of squatters	748	3.353	11.74	0	104
Controls					
Mean Income ($1000)	743	55.5	16.556	30.877	159.029
Area (km^2)	748	1310.848	6763.423	1.267	146691
Ruggedness	748	39.434	60.051	0	836.32
Grade	748	1.099	.428	1	3
Mineral endowment	748	3.769	1.062	0	7
Soil quality	748	.659	-	0	1
Body of water	748	.075	-	0	1
Distance to river (km)	748	8.797	10.893	0	86.805
Urban	748	.595	-	0	1
Inner Regional	748	.277	-	0	1
Outer Regional	748	.12	-	0	1
Remote Australia	748	.005	-	0	1
Very Remote Australia	748	.003	-	0	1
New South Wales	748	.531	-	0	1
Victoria	748	.469	-	0	1

OLS estimates are reported in Table 5.2. Column 1 shows there is a positive relationship between the number of squatting stations historically in a SA2, and the Gini coefficient. This relationship is significant at the 5% level. The results suggest that an additional ten squatting stations historically is associated with a 0.009 increase in the Gini coefficient. Therefore, the results suggest there is a meaningful relationship between the prominence of squatters historically and the level of inequality today.

Table 5.2: SA2 - OLS Estimates

	Parsimonious specifications				Baseline specification
	(1)	(2)	(3)	(4)	(5)
Variable	Gini coefficient	Gini coefficient	Gini coefficient	Gini coefficient	Gini coefficient
Number of squatters	0.0009**	0.0011***	0.0011***	0.0011***	0.0012***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Controlling for:					
Mean Income	✓	✓	✓	✓	✓
Area	✓	✓	✓	✓	✓
Remoteness dummies	✓	✓	✓	✓	✓
State dummy	✓	✓	✓	✓	✓
Terrain ruggedness	-	✓	✓	✓	✓
Terrain grade	-	✓	✓	✓	✓
Mineral endowment	-	-	✓	✓	✓
Soil quality	-	-	-	✓	✓
Access to water	-	-	-	-	✓
Observations	743	743	743	743	743
R-squared	0.4910	0.5203	0.5324	0.5371	0.5425

Robust standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

However, it is likely that this estimate is subject to omitted variable bias. One concern is that the geophysical characteristics of the land drive differences in outcomes over time. It is possible that there is greater variation in the potential productivity of land when the terrain is more rugged. This could then influence the differences in inequality. Similarly, differences in the gradient of terrain could drive differences in productivity, which then influences the level of inequality.

Column 2 reports the relationship between the prominence of squatters and the Gini coefficient when controlling for terrain ruggedness and terrain grade. The estimate suggests that an additional ten squatting stations historically is associated with a 0.011 increase in the Gini coefficient.

A second potential concern is that differences in mineral endowments within a region could drive differences in outcomes. Some regions could have greater deposits of coal or gold for example. Having access to valuable minerals could have a strong influence on the incomes of a select few, which could drive the differences in inequality. If this is correlated with the locations of the squatting settlements, the results will be biased. Column 3 reports the estimate controlling for differences in mineral endowments in different regions to account for this.

A third concern is that differential outcomes can be explained by variation in soil quality. Soil quality is an important determinant for the productivity of agriculture, which will have an impact on the level of inequality within a region. Furthermore, as mentioned in Section 3, historically the selectors were more successful on land that was suitable for agriculture. The squatters exerted less influence in these areas than areas in which the selectors struggled to be productive. For this reason, I control for soil quality in Column 4.

Similarly, it is likely that the regions the squatters decided to settle were highly dependent on access to water. Therefore, to account for this self-selection bias, I control for access to water in column 5, which is the baseline specification. Column 5 reports my baseline specification, controlling for observables. The results suggest that an additional ten squatting stations historically is associated with a 0.012 increase in the Gini coefficient. Notably, the estimate remains stable across all specifications. I also test to see if the results are robust to using different measures of inequality as my dependent variable. These specifications are reported in Table 5.3.

Column 1 of Table 5.3 is equivalent to Column 5 of Table 5.2, for a frame of reference. Columns 2-4 report the results with percentile ratios as the dependent variable.

Table 5.3: SA2 - OLS Estimates - Checking for robustness across inequality measures

Variable	(1) Gini coefficient	(2) P80/P20	(3) P80/P50	(4) P20/P50	(5) Top 1%	(6) Top 5%	(7) Top 10%
Number of squatters	0.0012***	0.0147*	0.0003	-0.0009***	0.0240**	0.0469**	0.0547**
	(0.000)	(0.008)	(0.001)	(0.000)	(0.011)	(0.019)	(0.021)
Controlling for:							
Mean Income	✓	✓	✓	✓	✓	✓	✓
Area	✓	✓	✓	✓	✓	✓	✓
Remoteness dummies	✓	✓	✓	✓	✓	✓	✓
State dummy	✓	✓	✓	✓	✓	✓	✓
Terrain ruggedness	✓	✓	✓	✓	✓	✓	✓
Terrain grade	✓	✓	✓	✓	✓	✓	✓
Mineral endowment	✓	✓	✓	✓	✓	✓	✓
Soil quality	✓	✓	✓	✓	✓	✓	✓
Access to water	✓	✓	✓	✓	✓	✓	✓
Observations	743	743	743	743	743	743	743
R-squared	0.5425	0.3905	0.5060	0.3935	0.6091	0.6674	0.6583

Robust standard errors in parentheses

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

An additional ten squatting stations historically within a SA2 is associated with a 0.147 increase in the P80/P20 ratio, which is significant at the 10% level, and a 0.009 decrease in the P20/P50 ratio, which is significant at the 1% level. There is no significant difference in the P80/P50 ratio. These results help to identify the point in the distribution of income where differences in inequality appear. There seems to be no real difference in the middle to upper-middle end of the income distribution. However, there does seem to be some difference across the lower tail of the distribution. The results suggest that individuals on the lower tail of the income distribution seem to earn relatively less in regions with greater prominence of squatters historically.

Columns 5-7 report the results with top income shares as the dependent variable. An additional ten squatting stations is associated with a 0.24 percentage point increase in the share of income held by the top 1%, a 0.47 percentage point increase in the share of income held by the top 5% and a 0.55 percentage point increase in the share of income held by the top 10%. These relationships are all significant at the 5% level.

The results in Table 5.3 suggest that there is a robust relationship between the historical prominence of squatters in SA2s and the level of inequality today at the SA2 level. These results are also robust to other specifications of the baseline model. In particular I consider using an inverse hyperbolic sine transformation for the number of squatters: $ln(S_i + \sqrt{S_i^2 + 1})$. This helps to give the distribution of the variable of interest a 'more normal' distribution by approximating a log transformation, without discarding the observations equal to zero (see Table A.1 in Appendix A). The results are also robust to using a quadratic term for the number of squatters (see Table A.2 in Appendix A). Furthermore, since there are many observations in my dataset with no squatters, I also show that the results are robust to using a truncated regression for all of the above specifications (see Tables A.3, A.4, and A.5 in Appendix A).

I also test to see if these results are robust at the LGA level. Descriptive statistics at the LGA level are reported in Table 5.4. OLS estimates using the baseline specification are reported in Table 5.5. Aggregating the data results in a loss of significance across the majority of the specifications. There is still a positive, and significant relationship between the historical prominence of squatters and the Gini coefficient. An additional ten squatting stations is associated with a 0.004 increase in the Gini coefficient, which is still notable.

I attribute the loss of significance to two key factors. Firstly, aggregating the data to the LGA level rather than the SA2 level results in a loss of precision describing the historical location of the squatters. A higher level of aggregation means having to take an average across a larger area, which dilutes the influence of the squatters. Secondly, I lose a significant number of observations using a higher level of aggregation. While the direction of all the estimates remain consistent, the standard errors are larger, reducing the significance of the results.

Table 5.4: LGA - Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Inequality measures					
Gini coefficient	232	.481	.052	.377	.674
P80/P20	232	5.33	1.142	3.45	11.25
P80/P50	232	1.875	.153	1.59	2.52
P20/P50	232	.362	.051	.17	.47
Top 1%	229	7.546	2.572	4.3	20.9
Top 5%	229	20.48	4.179	14.9	41.7
Top 10%	232	31.713	4.543	24.9	53.8
Variable of interest					
Number of squatters	232	11.03	17.512	0	102
Controls					
Mean Income ($1000)	232	52.502	15.723	34.237	143.328
Area (km^2)	232	4430.971	8904.849	5.717	93213.45
Ruggedness	232	53.228	59.788	0	319.973
Grade	230	1.187	.524	1	3
Mineral endowment	230	4.109	1.045	2	6
Soil quality	232	.767	-	0	1
Body of water	232	.172	.379	0	1
Distance to river (km)	232	14.894	15.046	.035	72.966
Urban	232	.349	-	0	1
Inner Regional	232	.375	-	0	1
Outer Regional	232	.233	-	0	1
Remote Australia	232	.026	-	0	1
Very Remote Australia	232	.017	-	0	1
New South Wales	232	.659	-	0	1
Victoria	232	.341	-	0	1

Table 5.5: LGA - OLS Estimates

Variable	(1) Gini coefficient	(2) P80/P20	(3) P80/P50	(4) P20/P50	(5) Top 1%	(6) Top 5%	(7) Top 10%
Number of squatters	0.0004*	0.0019	0.0001	-0.0002	0.0154	0.0255	0.0273
	(0.000)	(0.006)	(0.001)	(0.000)	(0.011)	(0.016)	(0.017)
Controlling for:							
Mean Income	✓	✓	✓	✓	✓	✓	✓
Area	✓	✓	✓	✓	✓	✓	✓
Remoteness dummies	✓	✓	✓	✓	✓	✓	✓
State dummy	✓	✓	✓	✓	✓	✓	✓
Terrain ruggedness	✓	✓	✓	✓	✓	✓	✓
Terrain grade	✓	✓	✓	✓	✓	✓	✓
Mineral endowment	✓	✓	✓	✓	✓	✓	✓
Soil quality	✓	✓	✓	✓	✓	✓	✓
Access to water	✓	✓	✓	✓	✓	✓	✓
Observations	230	230	230	230	229	229	230
R-squared	0.4233	0.2868	0.5483	0.3417	0.6496	0.6561	0.6237

Robust standard errors in parentheses

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

I also run different specifications of my baseline model, as I did at the SA2 level. These include using an inverse hyperbolic sine transformation of the number of squatters (reported in Table A.6), a quadratic parameterisation of the number of squatters (reported in Table A.7), and a truncated regression form of all of the above specifications (reported in Tables A.8, A.9, and A.10). The results are robust for all of these specfications.

Chapter 6

Channels of Persistence

The previous chapter established a positive relationship between the historical prominence of squatters in a region with the level of inequality today. This chapter provides evidence for the transmission mechanisms through which inequality persists over time.

I focus on two main channels of persistence. The first relates to the possibility that the industries prominent in the squatting regions may attract unskilled workers, resulting in higher inequality today. It is likely that these regions are overspecialised in agriculture today, and this may explain the persistence over time.

The second channel relates to weaker investment in public goods. Weaker investment in public goods could be a result of weaker collective action that comes from the historically coercive labour institutions in these regions.

6.1 Agriculture

I expect agriculture to be more dominant in the squatting regions, as the landed elite have an incentive to maintain the status quo over time. Regions with higher levels of agricultural employment are correlated with higher levels of inequality, as shown in Figure 6.1. Therefore, if regions with more squatters historically are still dominated by agricultural employment, this could explain why inequality is higher in these regions.

I use OLS to explore this possibility, using the specification outlined in Equation 6.1:

$$y_i = \alpha + \beta S_i + \boldsymbol{\gamma X_i} + \epsilon_i \tag{6.1}$$

Aside from the dependent variable, this specification is identical to the model outlined in the previous chapter. The outcome of interest y_i here is the proportion of the working population in region i employed in agriculture, forestry and fishing. $\boldsymbol{X_i}$ is the vector of baseline controls, which includes mean income, area, remoteness dummies, state dummy, geophysical characteristics, soil quality, and access to water. The results of this regression are reported in Table 6.1.

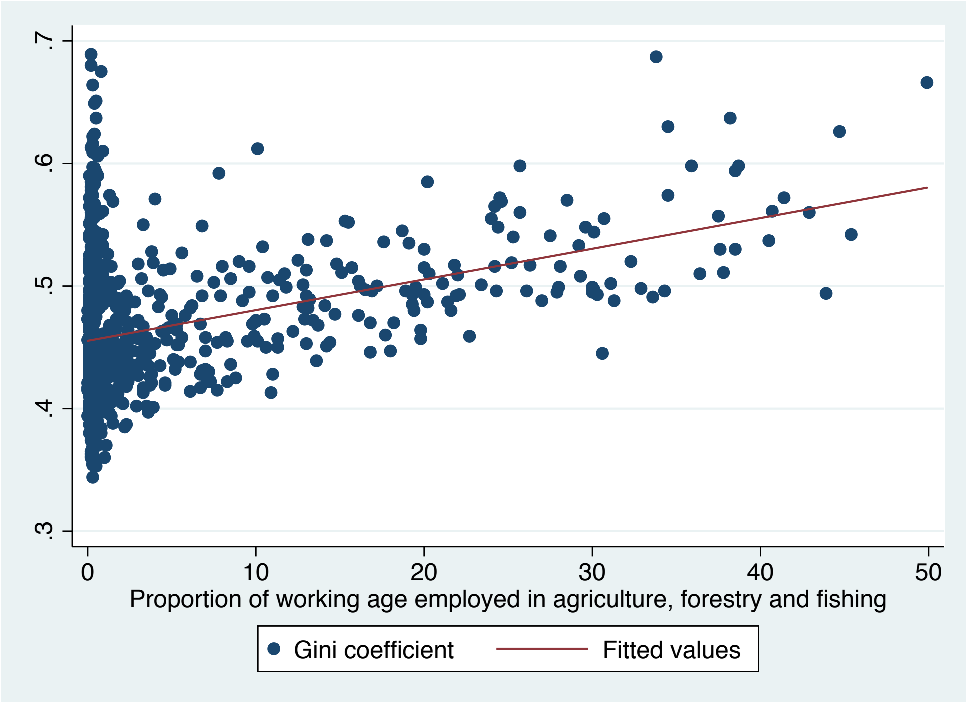

Figure 6.1: Correlation plot - Proportion of working population employed in Agriculture, Forestry and Fishing, and Gini coefficient

Column 1 of Table 6.1 reports the estimate aggregating at the SA2 level and Column 1 reports the estimate aggregating at the LGA level.

Table 6.1: Channels - Agriculture

Variable	(1) SA2 Agriculture	(2) LGA Agriculture
Number of squatters	0.2891***	0.1837***
	(0.069)	(0.057)
Observations	743	230
R-squared	0.7109	0.6831

Robust standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Controls include: mean income, area, remoteness, state, ruggedness, terrain grade, mineral endowment, soil quality, and access to water.

Both estimates show a strong positive relationship between the number of squatting stations historically and the proportion of workers employed in agriculture. Predictably, the magnitude of the LGA estimate is less than the SA2 estimate. The SA2 estimate suggests that an additional ten squatting stations historically is associated with a 2.9 percentage point increase in the proportion of workers employed in agriculture. The LGA estimate suggests a 1.8 percentage point increase. Therefore, it is clear that agriculture remains particularly important for these regions today. This in itself is not too surprising. However, the fact that these regions are overspecialised in agriculture relative to other regions, even after controlling for land characteristics is notable.

So higher levels of agricultural employment are positively correlated with inequality, and the squatting regions tend to focus their production efforts on agriculture. This suggests that the higher levels of inequality in the squatting regions could be explained by the dominance of agriculture.

However, it is possible that low-skilled workers self-select into these regions to work in agriculture. It is also possible that agriculture has remained dominant in these regions because the landed elite have a stake in the status quo. The literature suggests that the landed elite may prevent investment in schools and other public goods to preserve their lofty status in society (Engerman and Sokoloff, 2002). If workers are uneducated they are less likely to migrate to urban centres. This ensures the elite have a pool of cheap labour to draw upon. This could then drive the inequality in these regions.

6.2 Public Goods

A complementary explanation for the persistence of inequality is that the historical coercive labour institutions in the squatting regions limited the possibility of collective action towards greater investment in public goods. I will explore this possibility, focusing on three broad categories of investment: education, health, and transport. I hypothesise that regions in which the squatters were more prominent will have weaker investment into these public goods.

The wealthy tend to have less stake in the quality of public goods. For example, the wealthy elite in regional areas tend to send their children to private boarding schools. Any investment into improving the quality of education in regional areas only threatens their lofty status in society. Furthermore, the Indian case described by Banerjee and Iyer (2005) gives credence to the idea that the squatters may have wasted resources trying to preserve the status quo, while the working class may have wasted their resources trying to create a more equal distribution of wealth. This distracts policymakers from implementing broad-based growth promoting policies such as investments into education, health and transport.

I use OLS to explore these channels of persistence, using the specification outlined in Equation (6.2):

$$y_i = \alpha + \beta S_i + \boldsymbol{\gamma X_i} + \epsilon_i \tag{6.2}$$

The dependent variable y_i is some measure of investment over time in education, health, and transport. $\boldsymbol{X_i}$ is a vector of controls, similar to those presented in Section 5. The only difference is that I now control for population density rather than area. Descriptive statistics for the dependent variables are outlined in Table 6.2. Importantly, the data I use to estimate the relationships between historical prominence of squatters and public good provision is aggregated to the LGA level because the public goods data is not available at the SA2 level.

Table 6.2: Public goods - Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Distance from primary school	231	4.227	5.718	.6	60.7
Distance from secondary school	231	12.449	16.222	.6	147.6
Access to hospital services	231	.013	.016	0	.177
Access to GP services	231	5.469	1.041	1.7	8.5
Access to allied health services	231	.108	.022	.049	.166
Distance from major highway	231	4.689	7.054	.5	69.5
Distance to nearest railway station	231	17.516	26.093	1.8	174.4

6.2.1 Education

I hypothesise that a channel through which inequality persists in regions with a historically greater prominence of squatters is investment into education, or lack thereof. Greater investment into the public provision of education improves the opportunities available to those that cannot afford to enter the private education system.

To estimate the relationship between the historical prominence of squatters and investment in education, I use average distance from primary and secondary schools, measured in kilometres, as my dependent variable. A region that has a larger average distance to primary or secondary schools, controlling for population density, is assumed to have lower investment in education. For the same reasons outlined in Section 5, I also control for other observable sources of bias, including mean income remoteness, state, terrain ruggedness, terrain grade, mineral endowments, soil quality, and access to water.

Table 6.3: LGA - Education investment

Variable	Distance from primary school (1)	(2)	(3)	(4)	(5)	Distance from secondary school (6)	(7)	(8)	(9)	(10)
Number of squatting stations	0.0938***	0.1059***	0.1172***	0.1183***	0.1147***	0.1331	0.1569*	0.1508*	0.1496*	0.1608*
	(0.030)	(0.034)	(0.035)	(0.035)	(0.035)	(0.084)	(0.084)	(0.079)	(0.079)	(0.083)
Controlling for:										
Mean income ($1000)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Population density	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Remoteness dummies	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
State dummy	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Terrain ruggedness	-	✓	✓	✓	✓	-	✓	✓	✓	✓
Terrain grade	-	✓	✓	✓	✓	-	✓	✓	✓	✓
Mineral endowment	-	-	✓	✓	✓	-	-	✓	✓	✓
Soil quality	-	-	-	✓	✓	-	-	-	✓	✓
Access to water	-	-	-	-	✓	-	-	-	-	✓
Observations	231	230	230	230	230	231	230	230	230	230
R-squared	0.5392	0.5615	0.5778	0.5785	0.5833	0.4319	0.4386	0.4454	0.4455	0.4469

Robust standard errors in parentheses

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

OLS estimates are provided in Table 6.3. Columns 1-4 report the parsimonious specifications for the relationship between the historical prominence of squatters and the average distance from primary schools. Column 5 reports the baseline specification. The estimate in Column 5 suggests that an additional ten squatting stations in a LGA is associated with an additional 1.1km in average distance to primary schools.

This relationship is significant at the 1% level. Furthermore, this estimate is relatively stable across the five specifications. Similarly, Columns 6-9 report the parsimonious specifications for the relationship between the historical prominence of squatters and the average distance from secondary schools. Column 10 reports the baseline specification. The results suggest that an additional ten squatting stations historically is associated with an increase in the average distance from secondary schools by 1.6km. This result is significant at the 10% level and is relatively stable across the various specifications.

The results in Table 6.3 overall seem to indicate that the squatting regions have seen weaker investment in education relative to non-squatting regions.

6.2.2 Health

I also investigate the potential long-run implications for health investment. The dependent variables I use to proxy for health investment are access to allied health services, which is measured as the proportion of the population employed in health services excluding hospitals, access to hospital services, which is measured as the proportion of the population employed in hospital services, and access to GP services, which is the number of GP services per capita. Higher values for all of these variables indicate greater levels of development with respect to health. Table 6.4 reports the relationship between historical prominence of squatters and access to allied health services. Similarly, Table 6.5 reports the relationship between historical prominence of squatters and access to hospital services (Columns 1-5) and access to GP services (Columns 6-10).

Column 5 of Table 6.4 reports the estimate relating the prominence of squatters to access to allied health services using the baseline specification. It suggests that a one standard deviation increase in the number of squatting stations is associated with a decrease in access to allied health services by 0.19 standard deviations. This result is significant at the 5% level and is stable across the various specifications.

Columns 1-5 of Table 6.5 suggest that the presence of squatters in these regions had no effect on access to hospital services. Columns 6-10 suggest a negative relationship between the presence of squatters and access to GP services. However, using the baseline specification, this result is not significant. Overall there does not seem to be a an economically significant difference in investment in health services in the squatting regions.

Table 6.4: LGA - Health investment (a)

	Access to allied health services				
Variables	(1)	(2)	(3)	(4)	(5)
Number of squatters	-0.0002**	-0.0002*	-0.0002**	-0.0002**	-0.0002**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
<u>Controlling for:</u>					
Mean Income	✓	✓	✓	✓	✓
Population density	✓	✓	✓	✓	✓
Remoteness dummies	✓	✓	✓	✓	✓
State dummy	✓	✓	✓	✓	✓
Terrain ruggedness	-	✓	✓	✓	✓
Terrain grade	-	✓	✓	✓	✓
Mineral endowment	-	-	✓	✓	✓
Soil quality	-	-	-	✓	✓
Access to water	-	-	-	-	✓
Observations	231	230	230	230	230
R-squared	0.1424	0.1715	0.1811	0.2022	0.2029

Robust standard errors in parentheses

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

Table 6.5: LGA - Health investment (b)

	Access to hospital services					Access to GP services				
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Number of squatters	-0.0000	-0.0000	-0.0000	-0.0000	0.0000	-0.0059	-0.0091*	-0.0089	-0.0094*	-0.0076
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
Controlling for:										
Mean income ($1000)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Population density	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Remoteness dummies	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
State dummy	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Terrain ruggedness	-	✓	✓	✓	✓	-	✓	✓	✓	✓
Terrain grade	-	✓	✓	✓	✓	-	✓	✓	✓	✓
Mineral endowment	-	-	✓	✓	✓	-	-	✓	✓	✓
Soil quality	-	-	-	✓	✓	-	-	-	✓	✓
Access to water	-	-	-	-	✓	-	-	-	-	✓
Observations	231	230	230	230	230	231	230	230	230	230
R-squared	0.1078	0.1072	0.1097	0.1189	0.1194	0.2072	0.2306	0.2581	0.2622	0.2740

Robust standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

6.3 Transport

To examine the long-run effect of the squatters on transport infrastructure, I use proxies for investment in roads and rail. I proxy for investment in roads using distance from the business centre of the LGA to the nearest highway. I proxy for rail investment using average distance to the nearest railway station per service. Stronger transport infrastructure increases the mobility of the rural population to access employment in other regions. Therefore, I hypothesise there to be a negative relationship between investment in transport infrastructure and the prominence of squatters in a region.

Columns 1-5 of Table 6.6 suggest that there is a positive but statistically insignificant relationship between the historical prominence of squatters and the distance from the business centre of a LGA from a major highway. Column 10 suggests that an additional ten squatting stations is associated with an increase in the distance to the nearest railway station by 2.4km. This estimate is significant at the 5% level, but it does not seem very stable across the various specifications.

Overall, the transport results suggest there could be a positive relationship between the prominence of squatters and investment into transport infrastructure. However, for the most part, these results are not significant.

Table 6.6: LGA - Transport infrastructure

	Distance from major highway					Distance to nearest railway station				
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Number of squatters	0.0112	0.0280	0.0341	0.0329	0.0368	0.1059	0.2013*	0.1982	0.2010	0.2465**
	(0.032)	(0.034)	(0.035)	(0.035)	(0.034)	(0.131)	(0.120)	(0.127)	(0.127)	(0.124)
Controlling for:										
Mean income ($1000)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Population density	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Remoteness dummies	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
State dummy	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Terrain ruggedness	-	✓	✓	✓	✓	-	✓	✓	✓	✓
Terrain grade	-	✓	✓	✓	✓	-	✓	✓	✓	✓
Mineral endowment	-	-	✓	✓	✓	-	-	✓	✓	✓
Soil quality	-	-	-	✓	✓	-	-	-	✓	✓
Access to water	-	-	-	-	✓	-	-	-	-	✓
Observations	231	230	230	230	230	231	230	230	230	230
R-squared	0.2506	0.2886	0.3015	0.3020	0.3047	0.2926	0.4216	0.4342	0.4344	0.4659

Robust standard errors in parentheses

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

6.4 Testing the Channels

The first section of this chapter showed that there is a positive relationship between employment in agriculture and inequality, and that regions with squatters have a more agricultural focus. The second section of this chapter showed that regions with greater historical prominence of squatters have lagged development in terms of investment in education. It is possible that these are the channels through which inequality persists over time.

To test this hypothesis, I include these channels into the baseline model from Chapter 5. The results from including these variables are reported in Table 6.7. Column 1 reports the estimated effect of the historical prominence of squatters on inequality at the SA2 level, after controlling for the proportion of the population employed in agriculture. Once agriculture is controlled for, the significant effect found in Chapter 5 disappears. The persistence of inequality may be attributed to the dominance of agriculture in the squatting regions. It is important to note that the proportion of the population employed in agriculture is an outcome, which could be the result of the landed elite preserving the status quo in the squatting regions. Column 2 reports the same specification at the LGA level, and provides further robustness to the findings from Column 1.

The estimate in Column 3 suggests that the persistence of inequality in squatting regions could also be explained by weaker investment in primary education. After including average distance from primary school as a covariate, the coefficient on the variable of interest loses it statistical significance. The estimate in Column 4 suggests that it is not the weaker investment in secondary education that explains the persistence of inequality, as the coefficient on the number of squatters is still statistically significant. Similarly, since the estimate in Column 5 is similar to the estimate in Column 3, it does not seem likely that weaker investment in secondary education is the driver of the persistence of inequality. Column 6 reports the estimate after including all of the channels of persistence in the regression. Predictably, the estimate on the variable of interest is statistically insignificant.

The results in Table 6.7 suggest that the channels identified in the previous section are able to explain the long-run persistence of inequality in the squatting regions.

Table 6.7: Testing the channels of persistence

	Dependent variable: Gini coefficient					
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	SA2	LGA with agriculture	LGA with primary	LGA with secondary	LGA with all education	LGA with agriculture and education
Number of squatting stations	0.0000	-0.0001	0.0003	0.0004*	0.0003	-0.0001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Agriculture, Forestry and Fishing	0.0039***	0.0030***	-	-	-	0.0031***
	(0.000)	(0.000)	-	-	-	(0.000)
Distance from primary school	-	-	0.0014	-	0.0016	-0.0005
	-	-	(0.001)	-	(0.001)	(0.001)
Distance from secondary school	-	-	-	0.0004	0.0005	-0.0000
	-	-	-	(0.000)	(0.000)	(0.000)
Observations	743	230	230	230	230	230
R-squared	0.6453	0.5709	0.4330	0.4302	0.4432	0.5721

Robust standard errors in parentheses

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

Baseline controls included in every specification. Controls include mean income, area, remoteness, state, terrain ruggedness, terrain grade, mineral endowment, soil quality, and access to water.

Chapter 7

Discussion

In Chapter 5 I investigated the potential long-run persistence of inequality in squatting regions. The results suggest that there is greater inequality in these regions today, even after controlling for geophysical characteristics. The leftover variation is presumed to be a human residual created by the large landholders in these regions. The results are somewhat robust to aggregation, but the loss of significance is not surprising. This could simply be explained by weaker precision and the smaller sample size.

In Chapter 6 I proposed and tested the hypothesis that the persistence of inequality in the squatting regions can be explained by the dominance of agriculture in these regions, and weaker investment in public goods. I found evidence suggesting that the persistence of inequality can be explained by the agricultural sector and weaker investment in education. Furthermore, the results suggest that the dominance of agriculture in the squatting regions can be explained by the strong incentives the large landholders in these regions have to preserve the status quo of agricultural production. It is also possible that the two channels I have identified are linked.

In this chapter, I discuss the links between the findings from the previous two chapters and use them to explain the historical origins of inequality.

7.1 Preserving the Status Quo

The massive economic growth over the 19th Century can largely be attributed to the productivity of the wool industry. The squatters became very wealthy and highly influential as a result. As the landed elite, they owned the means of production, so they accrued the majority of the benefits from this economic growth. This created incentives to maintain the status quo.

One way to do this is to limit investment in education. By preventing investment in education, or at least not promoting investment in education, the landed elite were able to keep a reserve army of low-skilled labour for their own benefit. This limits the working class' ability to accumulate human capital, and entrenches the elite class' lofty status in society. Furthermore, the working class are less likely to migrate to urban areas if they are unskilled.

Figure 7.1 shows that workers in agriculture, forestry and fishing in Australia have a relatively low level of educational attainment. 46.6% of individuals involved in this industry did not complete high school, compared with 22.8% in other industries. This outcome is not unique to Australia. Data from the US Department of Agriculture (1998) shows that 56.9% of agricultural employment consists of workers who did not complete high school, in contrast to an average of 13.7% in the economy as a whole. Furthermore, 16.6% of agricultural workers in the US completed thirteen or more years of schooling, relative to the economy-wide average of 54.5%.

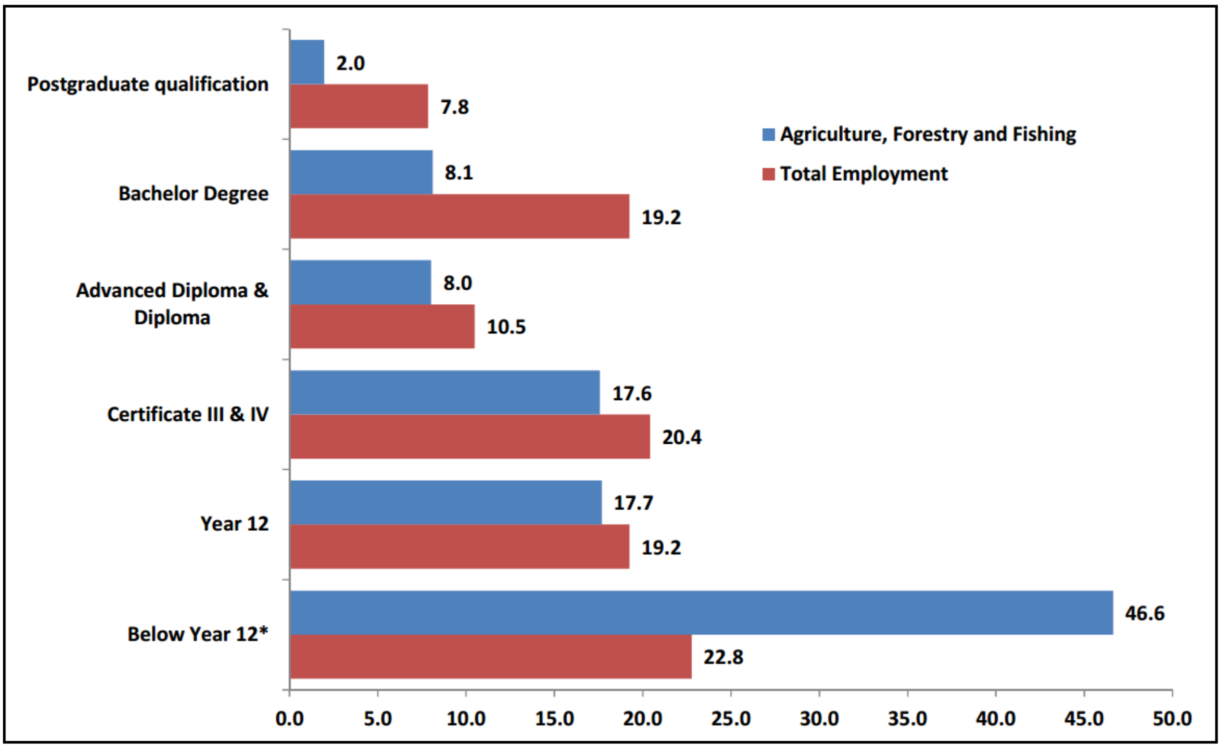

Figure 7.1: Educational attainment - Agriculture compared with other industries

Source: ABS 2011 *Census of Population and Housing.*

It is possible that low-skilled workers self-select into agriculture due to a lack of opportunities, but the results in the previous chapter show there seems to be a general lack of intent to improve educational attainment in these regions. This is symptomatic of what Easterly (2007) calls structural inequality (as opposed to market-based inequality). Consistent with Engerman and Sokoloff (2002), the author finds that large agricultural endowments concentrated in the hands of the few leads to high structural inequality. While the effect of market inequality is ambiguous, high structural inequality leads to poor development outcomes.

This is consistent with the findings from Galor et al. (2009). The authors find that where there is some level of complementarity between physical and human capital in production, the wealthy have incentive to support investment in public education. In these cases, the wealthy receive a benefit from the accumulation of human capital of their workers. This is akin to the market inequality described by Easterly (2007). However, the complementarity between physical and human capital in agricultural production is limited.

This suggests that it is not the wealthy per sé that limit investment in education. As long as the incentives of the wealthy align with providing public education to the workers, investment in public education will be strong. However, the large landholders in the squatting regions do not have as strong an incentive to lobby for stronger investment in education, and this lags the development of the workers in these regions.

The results in Chapter 5 suggest that the differences in inequality in the squatting regions relative to the non-squatting regions appear more strongly in the lower end of the distribution of income. Considering that production today in these regions is still dominated by agriculture, and that large agricultural landholders have weak incentives to invest in public education, this outcome seems to make sense. However, rapid technological change may increase the returns from human capital in the agricultural sector, which could help boost incentives to promote investment in public education.

7.2 Coercive Labour Institutions

An alternate but complementary explanation for the persistence of inequality in the squatting regions is that coercive labour institutions created a class divide limiting the collective action necessary for strong provision of public goods. The workers in the squatting regions were treated harshly. They were overworked, and they were not remunerated if a mistake was made. Mistakes were particularly costly for convicts. They would be presented with 'red shirts'. The squatters knew that if any particular situation was taken to court, their resources and influence would acquit them of any wrongdoing. This gap between the classes resulted in a lack of social cohesiveness in these regions.

The limiting effects of a divided society have long been documented in the literature. "No society can surely be flourishing and happy, of which the far greater part of the members are poor and miserable" (Smith, 1776). Lagerlöf (2005) shows that regions where slavery was more prominent in the U.S. have weaker development outcomes today. Nunn (2008) argues that weak social cohesiveness caused by the African slave trade can explain the weak development outcomes of parts of Africa.

In the Indian case explored by Banerjee and Iyer (2005), the authors explain that the class divide created in the landlord regions prevented collective action towards public good provision. They argue that the non-landlord regions were able to focus on growth promoting policies, while the landlord regions were focused on fixing the skewed distribution of wealth. A similar story could be true for the long-run development outcomes in the squatting regions. In the case on Peru, Dell (2010) shows that the negative long-run effects of coercive labour institutions in the *mita* region dominate the negative long-run effects of land inequality in the non-*mita* regions. Dell (2010) argues that the large landholders protected the population in the non-*mita* regions from the state.

However, the squatting regions in New South Wales and Victoria were characterised by large land inequality and coercive labour institutions. While large landholders helped shield the working class from an extractive state in Peru, this was not necessary in Australia. It was not a matter of picking your poison. It is likely that the coercive labour institutions resulted in a lack of social cohesiveness in the squatting regions which limited the potential for collective action towards stronger public education investment.

7.3 Implications

The status quo hypothesis and the coercive labour institutions hypothesis are likely linked. Both hypotheses predict weaker investment in public education, but how does this affect inequality? It is well established in the literature that, *ceteris paribus*, a worker's income is positively related to the level and quality of their education. This is a desirable property of a well-functioning economy, as it provides incentives to accumulate human capital, improving long-run growth and development outcomes. In reality, individuals have varying levels of access to education. Even if they do have access to education, naturally there is heterogeneity in the quality of education provided.

The results in this thesis suggest that there is a negative relationship between the presence of large landholders in the squatting regions and access to schools. Individuals born into a working class family in these regions will not have the same opportunities as individuals born into a wealthy family. It is common for children in wealthy families in regional areas to go to expensive private boarding schools. Private schools in Australia receive relatively generous funding from the state and federal government.

Meanwhile, it has recently been revealed that public schools in New South Wales are funded at 86% of their entitlements, and public schools in Victoria are funded at 83% of their entitlements (Knott and Hunter, 2016). So the working class in the squatting regions have weaker access to public schools, public schools in general are underfunded, and the wealthy elite tend to send their children to boarding schools. This creates a gap in the level and quality of educational attainment between the rural elite and the rural working class. This gap in opportunities restricts the expected earning potential of the rural working class, relative to the rural elite. This is how inequality of opportunity results in inequality of outcomes.

Therefore, any policy that aims to reduce long-run inequality in the squatting regions must focus on improving equality of opportunity. We can learn a lot from previous policy responses. In the landlord districts of India, the policy focus was on inequality of outcomes. Meanwhile, the non-landlord districts focused on growth-generating, productivity-enhancing reforms, and it is these regions that have greater development outcomes. Following this line of thinking, the government needs to make a concerted effort to improve access to quality public education in the squatting regions.

For many working class rural families, distance education is the only option. More recently, technological advancements have made virtual education more viable. In 2015, Australia's first virtual secondary school, Aurora College opened for students in New South Wales. Having more virtual schools would go a long way to remove the distance barrier to a quality education. This should improve the potential for human capital accumulation for the working class in rural areas, which should then reduce the opportunity gap between the working class and the elite. However, virtual education is still a recent initiative. While it sounds like a good idea in theory, it is too soon to tell how successful it will be in reducing the gap in opportunities between the working class and the wealthy in rural areas.

Chapter 8

Conclusion

Overall, this thesis provides evidence suggesting that inequality in Australia has historical origins. The discovery of grasslands outside the settlement boundaries created a land rush, allowing individuals to claim large plots of public land. From this land, Australia's wool industry flourished, and the squatters became very influential and wealthy. Once wealthy, they were able to employ labourers to do their work. Anecdotal evidence suggests that these workers were treated poorly, particularly convicts holding tickets-of-leave.

Controlling for geophysical characteristics and other key sources of imbalance, I find that regions with more squatters have higher levels of inequality and weaker investment in education. Furthermore, these regions are relatively more specialised in agriculture. I explain the persistence of inequality using two complementary hypotheses.

Firstly, the large landholders in the squatting regions had incentives to maintain the status quo. They had become wealthy through agricultural production. To maintain the dominance of agricultural production in the region, the elite prevented (or at least did not promote) investment in education. By limiting investment in education, the landed-elite maintained a reserve army of low-skilled workers. Secondly, the coercive labour institutions in the squatting regions limited the potential for collective actions towards provision of public schools. Since the wealthy tend to send their children to elite boarding schools, the shortfall of public schools has a larger negative effect on the rural working class.

Weaker access to quality education facilities limits the expected earning potential of the working class. However, rapid technological advancement has improved physical and human capital complementarity in agriculture. It is possible that this will now provide incentives for the landed elite to lobby for greater provision of public schools in the squatting regions. This could be a slow process, and policy should be proactive not reactive. There is potential for further investment in virtual education services, similar to Aurora College in New South Wales. This could help eliminate the distance barrier to quality education, which should help improve equality of opportunity in rural areas. However evaluating the effectiveness of such a policy is outside the scope of this thesis, and is left for future research.

While this thesis has provided some evidence suggesting that historical institutions can allow inequality to persist over long periods, there are limitations to my conclusions. My empirical strategy only controls for observable characteristics. However there are likely to be unobservable sources of bias. Therefore the results from this thesis should only be interpreted as suggestive. Furthermore, I do not make any claims that the historical institutional difference is the primary reason for differences in long run outcomes. I merely suggest that it could be one of many potential reasons.

Possible extensions of this research could include investigating the squatters' influence using a stronger empirical strategy. Future research could employ an instrumental variable approach to account for unobservable differences. One potential instrument that is relevant to the density of squatting, and could plausibly meet the exclusion restriction, is the historical density of the Indigenous population in a region. However, the Indigenous population has only been counted in the Census since the 1971 Census, and estimates of historical Indigenous population vary wildly.

Alternatively, the secondary dataset I generated from 19th Century squatting maps could potentially be used in a spatial regression discontinuity framework, as the dataset provides precise locations of the squatting settlements. Ideally, I would have liked to have used a spatial regression discontinuity design to identify 'the squatting effect' as the cause of differences in long-run outcomes. However, I was not able to acquire sufficient contemporary data at such a micro level, so this was not feasible. Such an approach could account for unobservable differences between squatting and non-squatting regions, as it is likely that other characteristics vary smoothly at the border of squatting regions.

While I focus on income inequality in this thesis, as this is an important determinant of wealth inequality (Piketty, 2016), future studies could extend this line of research by estimating measures of land inequality. Since the original source of inequality in the squatting regions was land-based, it would be interesting to identify if this persists over the very long-run.

Chapter A

Alternative Specifications

Table A.1: SA2 - Inverse hyperbolic sine transformation

VARIABLES	(1) Gini coefficient	(2) P80/P20	(3) P80/P50	(4) P20/P50	(5) Top 1%	(6) Top 5%	(7) Top 10%
Log number of squatters	0.0173***	0.2926***	0.0104*	-0.0160***	0.2977***	0.6445***	0.8021***
	(0.002)	(0.057)	(0.006)	(0.002)	(0.084)	(0.127)	(0.142)
Controlling for:							
Mean Income	✓	✓	✓	✓	✓	✓	✓
Area	✓	✓	✓	✓	✓	✓	✓
Remoteness dummies	✓	✓	✓	✓	✓	✓	✓
State dummy	✓	✓	✓	✓	✓	✓	✓
Terrain ruggedness	✓	✓	✓	✓	✓	✓	✓
Terrain grade	✓	✓	✓	✓	✓	✓	✓
Mineral endowment	✓	✓	✓	✓	✓	✓	✓
Soil quality	✓	✓	✓	✓	✓	✓	✓
Access to water	✓	✓	✓	✓	✓	✓	✓
Observations	743	743	743	743	743	743	743
R-squared	0.5741	0.4199	0.5078	0.4378	0.6121	0.6736	0.6668

Robust standard errors in parentheses

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

Table A.2: SA2 - Quadratic specification

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	Gini coefficient	P80/P20	P80/P50	P20/P50	Top 1%	Top 5%	Top 10%
Number of squatters	0.0036***	0.0614***	0.0028**	-0.0033***	0.0581***	0.1272***	0.1587***
	(0.001)	(0.014)	(0.001)	(0.001)	(0.017)	(0.027)	(0.030)
Number of squatters squared	-0.0000***	-0.0006***	-0.0000***	0.0000***	-0.0005***	-0.0011***	-0.0014***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Controlling for:							
Mean Income	✓	✓	✓	✓	✓	✓	✓
Area	✓	✓	✓	✓	✓	✓	✓
Remoteness dummies	✓	✓	✓	✓	✓	✓	✓
State dummy	✓	✓	✓	✓	✓	✓	✓
Terrain ruggedness	✓	✓	✓	✓	✓	✓	✓
Terrain grade	✓	✓	✓	✓	✓	✓	✓
Mineral endowment	✓	✓	✓	✓	✓	✓	✓
Soil quality	✓	✓	✓	✓	✓	✓	✓
Access to water	✓	✓	✓	✓	✓	✓	✓
Observations	743	743	743	743	743	743	743
R-squared	0.5664	0.4144	0.5089	0.4261	0.6111	0.6715	0.6639

Robust standard errors in parentheses

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

Table A.3: SA2 - Truncated regression

VARIABLES	(1) Gini coefficient	(2) P80/P20	(3) P80/P50	(4) P20/P50	(5) Top 1%	(6) Top 5%	(7) Top 10%
Number of squatters	0.0012***	0.0147*	0.0003	-0.0009***	0.0240**	0.0469**	0.0547***
	(0.000)	(0.008)	(0.001)	(0.000)	(0.011)	(0.019)	(0.021)
Controlling for:							
Mean Income	✓	✓	✓	✓	✓	✓	✓
Area	✓	✓	✓	✓	✓	✓	✓
Remoteness dummies	✓	✓	✓	✓	✓	✓	✓
State dummy	✓	✓	✓	✓	✓	✓	✓
Terrain ruggedness	✓	✓	✓	✓	✓	✓	✓
Terrain grade	✓	✓	✓	✓	✓	✓	✓
Mineral endowment	✓	✓	✓	✓	✓	✓	✓
Soil quality	✓	✓	✓	✓	✓	✓	✓
Access to water	✓	✓	✓	✓	✓	✓	✓
Observations	743	743	743	743	743	743	743

Robust standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.4: SA2 - Truncated regression (inverse hyperbolic sine transformation)

VARIABLES	(1) Gini coefficient	(2) P80/P20	(3) P80/P50	(4) P20/P50	(5) Top 1%	(6) Top 5%	(7) Top 10%
Log number of squatters	0.0173***	0.2926***	0.0104*	-0.0160***	0.2977***	0.6445***	0.8021***
	(0.002)	(0.056)	(0.006)	(0.002)	(0.083)	(0.125)	(0.140)
Controlling for:							
Mean Income	✓	✓	✓	✓	✓	✓	✓
Area	✓	✓	✓	✓	✓	✓	✓
Remoteness dummies	✓	✓	✓	✓	✓	✓	✓
State dummy	✓	✓	✓	✓	✓	✓	✓
Terrain ruggedness	✓	✓	✓	✓	✓	✓	✓
Terrain grade	✓	✓	✓	✓	✓	✓	✓
Mineral endowment	✓	✓	✓	✓	✓	✓	✓
Soil quality	✓	✓	✓	✓	✓	✓	✓
Access to water	✓	✓	✓	✓	✓	✓	✓
Observations	743	743	743	743	743	743	743

Robust standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.5: SA2 - Truncated regression (quadratic specification)

VARIABLES	(1) Gini coefficient	(2) P80/P20	(3) P80/P50	(4) P20/P50	(5) Top 1%	(6) Top 5%	(7) Top 10%
Number of squatters	0.0036***	0.0614***	0.0028**	-0.0033***	0.0581***	0.1272***	0.1587***
	(0.001)	(0.014)	(0.001)	(0.001)	(0.017)	(0.027)	(0.030)
Number of squatters squared	-0.0000***	-0.0006***	-0.0000***	0.0000***	-0.0005***	-0.0011***	-0.0014***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Controlling for:							
Mean Income	✓	✓	✓	✓	✓	✓	✓
Area	✓	✓	✓	✓	✓	✓	✓
Remoteness dummies	✓	✓	✓	✓	✓	✓	✓
State dummy	✓	✓	✓	✓	✓	✓	✓
Terrain ruggedness	✓	✓	✓	✓	✓	✓	✓
Terrain grade	✓	✓	✓	✓	✓	✓	✓
Mineral endowment	✓	✓	✓	✓	✓	✓	✓
Soil quality	✓	✓	✓	✓	✓	✓	✓
Access to water	✓	✓	✓	✓	✓	✓	✓
Observations	743	743	743	743	743	743	743

Robust standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.6: LGA - Inverse hyperbolic sine transformation

VARIABLES	(1) Gini coefficient	(2) P80/P20	(3) P80/P50	(4) P20/P50	(5) Top 1%	(6) Top 5%	(7) Top 10%
Log number of squatters	0.0070**	0.0594	-0.0066	-0.0053	0.1560	0.3371	0.3892
	(0.003)	(0.079)	(0.009)	(0.003)	(0.131)	(0.210)	(0.238)
Controlling for:							
Mean Income	✓	✓	✓	✓	✓	✓	✓
Area	✓	✓	✓	✓	✓	✓	✓
Remoteness dummies	✓	✓	✓	✓	✓	✓	✓
State dummy	✓	✓	✓	✓	✓	✓	✓
Terrain ruggedness	✓	✓	✓	✓	✓	✓	✓
Terrain grade	✓	✓	✓	✓	✓	✓	✓
Mineral endowment	✓	✓	✓	✓	✓	✓	✓
Soil quality	✓	✓	✓	✓	✓	✓	✓
Access to water	✓	✓	✓	✓	✓	✓	✓
Observations	230	230	230	230	229	229	230
R-squared	0.4320	0.2891	0.5499	0.3506	0.6491	0.6580	0.6264

Robust standard errors in parentheses

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

Table A.7: LGA - Quadratic specification

VARIABLES	(1) Gini coefficient	(2) P80/P20	(3) P80/P50	(4) P20/P50	(5) Top 1%	(6) Top 5%	(7) Top 10%
Number of squatters	0.0014**	0.0215	0.0001	-0.0013**	0.0100	0.0348	0.0503
	(0.001)	(0.014)	(0.001)	(0.001)	(0.021)	(0.034)	(0.038)
Number of squatters squared	-0.0000**	-0.0003*	-0.0000	0.0000**	0.0001	-0.0001	-0.0003
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Controlling for:							
Mean Income	✓	✓	✓	✓	✓	✓	✓
Area	✓	✓	✓	✓	✓	✓	✓
Remoteness dummies	✓	✓	✓	✓	✓	✓	✓
State dummy	✓	✓	✓	✓	✓	✓	✓
Terrain ruggedness	✓	✓	✓	✓	✓	✓	✓
Terrain grade	✓	✓	✓	✓	✓	✓	✓
Mineral endowment	✓	✓	✓	✓	✓	✓	✓
Soil quality	✓	✓	✓	✓	✓	✓	✓
Access to water	✓	✓	✓	✓	✓	✓	✓
Observations	230	230	230	230	229	229	230
R-squared	0.4336	0.2970	0.5483	0.3582	0.6498	0.6563	0.6245

Robust standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.8: LGA - Truncated regression

VARIABLES	(1) Gini coefficient	(2) P80/P20	(3) P80/P50	(4) P20/P50	(5) Top 1%	(6) Top 5%	(7) Top 10%
Number of squatters	0.0004*	0.0019	0.0001	-0.0002	0.0154	0.0255	0.0273*
	(0.000)	(0.005)	(0.001)	(0.000)	(0.010)	(0.016)	(0.017)
Controlling for:							
Mean Income	✓	✓	✓	✓	✓	✓	✓
Area	✓	✓	✓	✓	✓	✓	✓
Remoteness dummies	✓	✓	✓	✓	✓	✓	✓
State dummy	✓	✓	✓	✓	✓	✓	✓
Terrain ruggedness	✓	✓	✓	✓	✓	✓	✓
Terrain grade	✓	✓	✓	✓	✓	✓	✓
Mineral endowment	✓	✓	✓	✓	✓	✓	✓
Soil quality	✓	✓	✓	✓	✓	✓	✓
Access to water	✓	✓	✓	✓	✓	✓	✓
Observations	230	230	230	230	229	229	230

Robust standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.9: LGA - Truncated regression (inverse hyperbolic sine transformation)

VARIABLES	(1) Gini coefficient	(2) P80/P20	(3) P80/P50	(4) P20/P50	(5) Top 1%	(6) Top 5%	(7) Top 10%
Log number of squatters	0.0070**	0.0594	-0.0066	-0.0053	0.1560	0.3371*	0.3892*
	(0.003)	(0.076)	(0.008)	(0.003)	(0.126)	(0.201)	(0.228)
Controlling for:							
Mean Income	✓	✓	✓	✓	✓	✓	✓
Area	✓	✓	✓	✓	✓	✓	✓
Remoteness dummies	✓	✓	✓	✓	✓	✓	✓
State dummy	✓	✓	✓	✓	✓	✓	✓
Terrain ruggedness	✓	✓	✓	✓	✓	✓	✓
Terrain grade	✓	✓	✓	✓	✓	✓	✓
Mineral endowment	✓	✓	✓	✓	✓	✓	✓
Soil quality	✓	✓	✓	✓	✓	✓	✓
Access to water	✓	✓	✓	✓	✓	✓	✓
Observations	230	230	230	230	229	229	230

Robust standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.10: LGA - Truncated regression (Quadratic specification)

VARIABLES	(1) Gini coefficient	(2) P80/P20	(3) P80/P50	(4) P20/P50	(5) Top 1%	(6) Top 5%	(7) Top 10%
Number of squatters	0.0014**	0.0215	0.0001	-0.0013**	0.0100	0.0348	0.0503
	(0.001)	(0.013)	(0.001)	(0.001)	(0.020)	(0.032)	(0.037)
Number of squatters squared	-0.0000**	-0.0003**	-0.0000	0.0000***	0.0001	-0.0001	-0.0003
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Controlling for:							
Mean Income	✓	✓	✓	✓	✓	✓	✓
Area	✓	✓	✓	✓	✓	✓	✓
Remoteness dummies	✓	✓	✓	✓	✓	✓	✓
State dummy	✓	✓	✓	✓	✓	✓	✓
Terrain ruggedness	✓	✓	✓	✓	✓	✓	✓
Terrain grade	✓	✓	✓	✓	✓	✓	✓
Mineral endowment	✓	✓	✓	✓	✓	✓	✓
Soil quality	✓	✓	✓	✓	✓	✓	✓
Access to water	✓	✓	✓	✓	✓	✓	✓
Observations	230	230	230	230	229	229	230

Robust standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Chapter B

Measuring Inequality

Litchfield (1999) notes that there are five key axioms that are generally deemed desirable for a measure of inequality:

The Pigou-Dalton Transfer Principle (Dalton, 1920; Pigou, 1912). This axiom requires the inequality measure to rise (or at least not fall) in response to a mean-preserving spread: an income transfer from a poorer person to a richer person should result in a rise (or at least not a fall) in inequality. A transfer from a richer person to a poorer person should result in a fall (or at least not a rise) in inequality. Consider a vector $y\prime$ which is a transformation of the vector y obtained by a transfer ϵ from y_j to y_i, where $y_i > y_j$, and $y_i + \epsilon > y_j - \epsilon$. The transfer principle is satisfied if and only if $I(y\prime) \geq I(y)$.

Income Scale Independence. This axiom requires the inequality measure to be invariant to uniform proportional changes to each individual's income. So for any scalar $\lambda > 0$ we should have, $I(y) = I(\lambda y)$. For example if we change currency units, there should be no change in inequality.

Principle of Population (Dalton, 1920). The population principle requires inequality measures to be invariant to replications of the population. So if we merge two identical distributions, there should be no change in inequality.

Anonymity. This axiom requires the inequality measure to be independent of every other characteristic of the individuals in the population other than their income (or other indicator). Hence for any permutation $y\prime$ of y, we have $I(y) = I(y\prime)$.

Decomposability. This axiom implies that there should be a relationship between inequality in a population as a whole and inequality in its constituent parts. The idea is that we would like to be able to express total inequality as a function of inequality within the constituent subgroups.

Bibliography

Acemoglu, D., S. Johnson, and J. A. Robinson (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review 91*(5), 1369–1401.

Allen, D. W. (1991). Homesteading and Property Rights; or, "How the West Was Really Won". *Journal of Law and Economics 34*(1), 1–23.

Alston, L., G. Libecap, and B. Mueller (1999). *Titles, Conflict and Land Use: The Development of Property Rights and Land Reform on the Brazilian Amazon Frontier.* Economics, Cognition, and Society. University of Michigan Press.

Alston, L. J., E. Harris, and B. Mueller (2012). The Development of Property Rights on Frontiers: Endowments, Norms, and Politics. *The Journal of Economic History 72*(03), 741–770.

Alston, L. J., G. D. Libecap, and R. Schneider (1996). The Determinants and Impact of Property Rights: Land Titles on the Brazilian Frontier. *Journal of Law, Economics and Organization 12*(1), 25–61.

Anderson, T. L. and P. J. Hill (1975). The Evolution of Property Rights: A Study of the American West. *Journal of Law and Economics 18*(1), 163–79.

Athanasopoulos, G. and F. Vahid (2003). Statistical Inference and Changes in Income Inequality in Australia. *Economic Record 79*(247), 412–424.

Atkinson, A. B. (2015). *Inequality : What can be done?* Harvard University Press Cambridge, Massachusetts.

Banerjee, A. and L. Iyer (2005). History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India. *The American Economic Review 95*(4), 1190–1213.

Brown, J. C. (1978). A Nineteenth-Century Argentine Cattle Empire. *Agricultural History 52*(1), 160–178.

Burroughs, P. (1967). *Britain and Australia 1831-1855: A Study in Imperial Relations and Crown Lands Administration.* Clarendon P Oxford.

Butlin, N. G. (1983). *Our Original Aggression : Aboriginal Populations of Southeastern Australia, 1788-1850.* G. Allen & Unwin Sydney ; Boston.

Cingano, F. (2014). Trends in Income Inequality and its Impact on Economic Growth. OECD Social, Employment and Migration Working Papers 163, OECD Publishing.

Cinnirella, F. and E. Hornung (2016). Landownership concentration and the expansion of education. *Journal of Development Economics 121*(C), 135–152.

Coase, R. H. (1960). The Problem of Social Cost. *The Journal of Law & Economics 3*, 1–44.

Cowell, F. (2011). *Measuring Inequality.* Oxford University Press.

Dalton, H. (1920). The Measurement of the Inequality of Incomes. *The Economic Journal 30*(119), 348–361.

Dell, M. (2010). The Persistent Effects of Peru's Mining Mita. *Econometrica 78*(6), 1863–1903.

Demsetz, H. (1967). Toward a Theory of Property Rights. *The American Economic Review 57*(2), 347–359.

Dye, A. and S. La Croix (2013). The Political Economy of Land Privatization in Argentina and Australia, 18101850: A Puzzle. *The Journal of Economic History 73*(04), 901–936.

Easterly, W. (2007). Inequality does cause underdevelopment: Insights from a new instrument. *Journal of Development Economics 84*(2), 755–776.

Elbers, C., J. O. Lanjouw, and P. Lanjouw (2003). Micro–level Estimation of Poverty and Inequality. *Econometrica 71*(1), 355–364.

Engerman, S. L. and K. L. Sokoloff (2002). Factor Endowments, Inequality, and Paths of Development among New World Economies. *Economa 3*(1), 41–109.

Fei, J. C. H., G. Ranis, and S. W. Y. Kuo (1978). Growth and the Family Distribution of Income by Factor Components. *The Quarterly Journal of Economics 92*(1), 17–53.

Galor, O., O. Moav, and D. Vollrath (2009). Inequality in Landownership, the Emergence of Human-Capital Promoting Institutions, and the Great Divergence. *Review of Economic Studies 76*(1), 143–179.

Grosjean, P. and R. Khattar (2014). It's Raining Men! Hallelujah? Discussion Papers 2014-29, School of Economics, The University of New South Wales.

Ham, T. (1854). The squatting map of Victoria (Port Phillip District, New South Wales) [cartographic material].

Hansen, Z. K. and G. Libecap (2004). The allocation of property rights to land: US land policy and farm failure in the northern great plains. *Explorations in Economic History 41*(2), 103–129.

Harris, A. (1847). *Settlers and Convicts or, Recollections of Sixteen Years' Labour in the Australian Backwoods* .

Hodgson, G. M. (2006). What Are Institutions?. *Journal of Economic Issues (Association for Evolutionary Economics) 40*(1), 1 – 25.

Kaplow, L. (2005). Why Measure Inequality? *The Journal of Economic Inequality 3*(1), 65–79.

Knott, M. and F. Hunter (2016, October). More than 150 private schools over-funded by hundreds of millions of dollars each year. *Sydney Morning Herald*. `http://www.smh.com.au/federal-politics/political-news/full-list-1-billion-flowing-to-wealthy-private-schools-officially-classed-as-overfunded-20160930-grs6nz.html`. Accessed: 2016-11-16.

Kuznets, S. (1955). Economic Growth and Income Inequality. *The American Economic Review 45*(1), 1–28.

Lagerlöf, N. P. (2005). Geography, Institutions and Growth: The United States as a Microcosm.

Leigh, A. (2005). Deriving Long-Run Inequality Series from Tax Data. *The Economic Record 81*(s1), S58–S70.

Libecap, G. D. (1989). Distributional Issues in Contracting for Property Rights. *Journal of Institutional and Theoretical Economics (JITE) / Zeitschrift fr die gesamte Staatswissenschaft 145*(1), 6–24.

Litchfield, J. A. (1999). Inequality: Methods and Tools.

McLean, I. W. (2016). *Why Australia Prospered*. Princeton University Press.

Mendolia, S. and P. Siminski (2016). New Estimates of Intergenerational Mobility in Australia. *The Economic Record 92*(298), 361–373.

Miranti, R., R. Cassells, Y. Vidyattama, and J. McNamara (2015). Measuring Small Area Inequality Using Spatial Microsimulation: Lessons Learned from Australia. *International Journal of Microsimulation 8*(2), 152–175.

Mitchener, K. J. and I. W. McLean (2003). The Productivity of US States Since 1880. *Journal of Economic Growth 8*(1), 73–114.

Nugent, J. B. and J. A. Robinson (2010). Are factor endowments fate? *Revista de Historia Econmica 28*(01), 45–82.

Nunn, N. (2008). The Long-term Effects of Africa's Slave Trades. *The Quarterly Journal of Economics 123*(1), 139–176.

Nunn, N. (2009). The Importance of History for Economic Development. *Annual Review of Economics 1*(1), 65–92.

Nunn, N. and D. Puga (2012). Ruggedness: The Blessing of Bad Geography in Africa. *Review of Economics and Statistics 94*(1), 20–36.

Pigou, A. C. (1912). *Wealth and Welfare.* Macmillan and Company, Limited.

Piketty, T. (2014). *Capital in the Twenty-First Century.* Harvard University Press.

Piketty, T. (2016). Thomas piketty: Is increasing inequality inevitable? reflections on capital in the 21st century.

Rajan, R. G. and R. Ramcharan (2011). Land and Credit: A Study of the Political Economy of Banking in the United States in the Early 20th Century. *Journal of Finance 66*(6), 1895–1931.

Ramcharan, R. (2010). Inequality and Redistribution: Evidence from U.S. Counties and States, 1890-1930. *The Review of Economics and Statistics 92*(4), 729–744.

Reuss, F. H., J. L. Browne, and J. N. Degotardi (1860). Reuss & Browne's map of New South Wales and part of Queensland showing the relative positions of the pastoral runs, squattages, districts, counties, towns, reserves.

Roberts, S. H. (1964). *The squatting age in Australia 1835-1847 / Stephen H. Roberts.* Melbourne University Press Melbourne.

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations.* History of Economic Thought Books. McMaster University Archive for the History of Economic Thought.

Stapledon, N. (2013). Australia's Major Terms of Trade and Commodity Shocks, 1800-2013: Sources and Impacts. Ceh discussion papers, Centre for Economic History, Research School of Economics, Australian National University.

Stiglitz, J. E. (2015). Inequality in America A Policy Agenda for a Stronger Future. *The ANNALS of the American Academy of Political and Social Science 657*(1), 8–20.

Tarozzi, A. and A. Deaton (2009). Using Census and Survey Data to Estimate Poverty and Inequality for Small Areas. *The Review of Economics and Statistics 91*(4), 773–792.