

https://dspace.org/

Essays on the political economy of state formation and of laboratory federalism

Keeton, Lyndal

2016

http://hdl.handle.net/10539/22115

Keeton, Lyndal (2016) Essays on the political economy of state formation and of laboratory federalism, University of the Witwatersrand, Johannesburg, <http://wiredspace.wits.ac.za/handle/10539/22115>

Downloaded from a DSpace repository. https://dspace.org/

ESSAYS ON THE POLITICAL ECONOMY OF STATE FORMATION AND OF LABORATORY FEDERALISM

A thesis presented for the degree of Doctor of Philosophy
in the School of Economic and Business Sciences,
University of the Witwatersrand

by

Lyndal Keeton
Student No: 9901459W
School of Economic and Business Sciences
University of the Witwatersrand
Private Bag X3, WITS 2050
Johannesburg, Republic of South Africa
+27.11.717.8065 (Tel.)
Lyndal.Keeton@Wits.ac.za

Supervised by Professor Giampaolo Garzarelli

Abstract

This thesis investigates the problem of the economic organisation of the public sector. It begins by establishing context by considering the two related issues central to it: the boundary of the state and the internal organisation of government (Chapter 1). There is a growing literature that explores the boundary of the state in political economy terms. Moreover, the boundary of the state can be viewed in a similar light to the boundary of the firm. The Second Generation Theory of Fiscal Federalism explores the internal organisation of government through the lens of the theory of the firm. Second Generation Theory assumes that governments are subject to the same problems that firms face: for example, just like firms require institutions to align the incentives of managers and shareholders (e.g., better defined contracts), governments require institutions to align the incentives of politicians and citizens (e.g., better defined constitutions).

In order to improve our understanding of economic performance over time, the state should be considered as a complex organisation held together by a series of public choice compromises. Chapter 2 considers one aspect of the state as an organisation: when a boundary change of an existing state generates a new state. It tries to economically capture the birth of a new state through boundary change by taking a cue from the theory of internal exit: the secession of a group of people from an existing state who will then go on to form a new state. Internal exit predicts an internal exit-proof tax rate, i.e., a state will set the tax rate so that internal exit will not occur (e.g., Quebec in Canada). However, in precolonial southern Africa (ca. 1600-1910), internal exit occurred. A well-known example of this is that of Mzilikazi who in the 19th century left the Zulu with his followers and formed his own, new state: the Ndebele. Why is it that in Africa internal exit as a threat failed and internal exit still took place? With the aid of a simple, historically informed model, this chapter offers a political economy explanation of why internal exit took place in precolonial southern Africa. The model shows how internal exit results from the payoff calculation of an elite member's (e.g., Mzilikazi) desire to maximise his share of public revenue surplus.

Chapter 3 considers the internal organisation of government through the role of intergovernmental grants in the context of laboratory federalism. The Public Economics literature on intergovernmental grants is extensive. In this extensive literature, grants are usually analysed according to consumer behaviour theory where income and substitution effects determine community spending (and ultimately community welfare). However, these effects shed little light on how local governments can use grants to experiment with policy (laboratory federalism) in order to develop new, successful policies. In fact, even casual empiricism shows that local governments routinely experiment with policy and achieve varying degrees of success. One recent example is Mayor Bloomberg's range of anti-poverty experiments in New York City. Very little theory has been produced that ties policy experimentation with the role of grants, however. Chapter 3 takes an organisational view of grants, namely it likens them to incomplete contracts to show how certain grants can be policy instruments for the creation and discovery of new knowledge in the public sector. More precisely, the chapter develops an evolutionary learning model that captures the knowledge gains that different types of grants (e.g., lump-sum grants compared to matching grants) can engender. It shows that a lump-sum grant can bring about greater learning at the local government level than a closed matching grant. Chapter 4 concludes by summarizing and suggesting areas for future research.

Acknowledgements

To begin, I would like to thank my supervisor, Professor Giampaolo Garzarelli, for his patience, encouragement, and mentorship.

I am grateful for financial assistance in the form of the Carnegie Time-off Grant and the RW Bethlehem Fellowship.

I would like to thank the participants of the ERSA Workshops on Economic History, on Public Economics, and on Political Economy, as well as the first and second annual WINIR conferences for helpful feedback.

Finally, I would like to thank Eugene for invaluable support, always.

Contents

Abstract i

Acknowledgements ii

Contents iii

List of Figures v

Chapter 1 Introduction 1

1.1 Introduction 1

1.2 The Boundary of the State 3

1.3 Internal Organisation of Government 9

1.4 Conclusion 13

1.5 References 14

Chapter 2 Internal Exit and Natural State Formation in Precolonial Southern Africa (ca. 1500-1910) 16

2.1 Introduction 16

2.2 Land, People, and the Right to Rule 21
2.2.1 African Land 21
2.2.2 African People: Value and Birthright 22
2.2.3 Broadcast of Power and the African Elite 23
2.2.4 A Paradigmatic Illustration 26

2.3 Model 28
2.3.1 The African State as a Public Organisation 29
2.3.2 Internal Exits: Compromise or Conflict 35
2.3.3 The Potential Exiter's Decision: To Exit or Not? 39
2.3.4 Can the Ruler Prevent Exit? 41

2.4 Related Literature 43

2.5 Conclusion 47

2.6 References 49

Chapter 3 Intergovernmental Grants and Local Governmental Learning 55

3.1 Introduction 55

3.2 Intergovernmental Grants as Policy Instruments for Experimentation 59

3.3 Local Government Learning 63

3.4. Melioration as Learning 66

3.5 Intergovernmental Grants: Leeway to Experiment 71
3.5.1 Melioration with a Lump-Sum Grant 72
3.5.2 Melioration with a Closed Matching Grant 73
3.5.3 Graphical Comparison of Melioration under Different Grants 75

3.6 Summary and Areas for Future Research 78

3.7 References 82

Appendix 3A Two Extreme Cases of Grant Experimentation 89

Appendix 3B The Process of Policy Adoption 92

Chapter 4 Conclusion 95

4.1 Summary and Conclusion 95

4.2 Areas for Future Research 97

4.3 References 98

List of Figures

2.1 The Ruler's Broadcast of Power 25

2.2 Map of Mzilikazi's Internal Exit 27

2.3 Population Size and Surplus 33

2.4 Exit Decision Tree 41

3.1 Melioration between Two Policy Choices 70

3.2 Learning under a Lump-sum Grant compared to a Closed Matching Grant 77

3A.1 Closed Matching Grant Conditions Crowd Out Experimentation 89

3A.2 No Crowding Out of Experimentation 90

Chapter 1

Introduction

1.1 Introduction

This thesis regards the problem of the economic organisation of the public sector along the two related dimensions of state boundaries and of internal organisation of government. The difference between state and government will be outlined presently, but let me immediately flag that the uniting thread of both is the complex nature of public organisation populated by self-interested individuals.

As a result, just like firms, state and government are subject to problems such as corruption, opportunism, and shirking. Hence, it is sensible to apply insights from the theory of the firm (or from industrial and economic organisation) to the economic study of the public sector. This is what this thesis does, inspired mainly by the works of Geoffrey Brennan, James M. Buchanan, Douglass C. North, Wallace E. Oates, John Joseph Wallis, and Barry R. Weingast.

The classic Weberian (1946: 78) definition of a state is a community that holds the "monopoly of the legitimate use of physical force within a given territory." While this definition may be useful for "Western" or "modern" states, it restricts our understanding only to those states with fixed territorial bounds or borders. As discussed in Chapter 2, this thesis takes a different view of the state given its African setting. The state is an organisation that holds a relative advantage on the legitimate use of physical force within a particular *social* space: common features such as culture, history, and language provide the principal boundary for the state. Thus, as we will see, the boundaries of a state are principally social; and

physical boundaries are incidental, mattering especially for broadcast of power reasons and not for land value increases intrinsic to territorial expansion.

The problem of the boundary of the state is viewed in parallel to the problem of the boundary of the firm (e.g., Lopez-de-Silanes, Shleifer, and Vishny 1997; Holian 2009; Levin and Tadelis 2010). Coase (1937) provides a definition of the boundary of the firm in terms of transaction costs: if the cost of internalising a transaction is less than that of leaving it on the market, then, all else equal, that transaction should take place within firm boundaries. The political economy literature attempts to make similar analogies for state boundaries (e.g., Alesina and Spolaore 1997, 2005).

The problem of the boundary of the state is related to our second concern – the problem about the internal organisation of government. In this thesis, government is the steering apparatus of the state: the group of people who both make decisions for and control the state (e.g., a president, state governors, mayors), often by means of complex organisational structures, such as bureaucracies and legislatures (e.g., Breton and Wintrobe 1982; Weingast and Marshall 1988).

The organisational form of the state within which our internal organisation of government is studied is the decentralised or fiscally federal one (Oates 1972), though, as we will see, the considerations are equally applicable to centralised forms. The Second Generation Theory of Fiscal Federalism (SGT) explores the internal organisation of government through the lens of the theory of the firm (e.g., Oates 2005; Weingast 2009). The SGT assumes that governments are subject to the same problems that firms face. For example, just like firms require institutions to align the incentives of managers and shareholders (e.g., better defined contracts), governments require institutions to align the incentives of politicians and citizens (e.g., better defined constitutions).

1.2 The Boundary of the State

In exploring the definition of state boundaries in political economy terms, Wittman (1991) and Bolton, Roland and Spolaore (1996) take an organisational view. Wittman (1991) suggests that a state's boundaries are determined by economies of scale: if public goods can be provided at a lower per capita cost within the state than elsewhere (e.g., within another state), then, all else equal, the public goods should be provided within state boundaries. With economies of scale in the production of public goods (e.g., national defense), large states can pay lower per capita costs of providing these goods vis-à-vis smaller states (e.g., when smaller states merge to form one larger state, they do not duplicate the public goods offered). However, larger states typically also have larger policy coordination and administration costs that can offset these economies of scale.

One of Wittman's assumptions is that markets and states overlap perfectly, meaning that one way to reduce transaction costs is through synergy gains from merging nations: synergy "is maximised when the set of merged states creates greater wealth than any other combination of states ... Negative synergy arises when the combination creates disfunction [*sic*]" (p. 127). Wittman continues that two states "would join together (separate) if the economies of scale and scope and the synergy produced by their union created greater (smaller) benefits than costs" (p. 129).

Bolton, Roland and Spolaore (1996) consider the parallel between firms and states as well, but suggest two differences. Firstly, when two firms merge, the owners negotiate the contract for the merger. But when two states merge, if democratic, the median voters decide for the rest of the citizens. Secondly, courts can legally enforce a merger between firms, but a merger between states needs to be self-enforcing (Bolton, Roland and Spolaore 1996: 699).

There are numerous other economic arguments for state boundaries, which do not rely on the economics of organisation. One is Findlay (1996), which models the territorial expansion

of states. The population of the state is separated into soldiers and farmers. A soldier's role is to protect the land and a farmer's role is to farm the land. The soldiers face defending the state's "sphere of influence" or geographic territory from the enemies outside the state. Because of the trade-off between the army required to defend the territory and the farmers to make productive use of the land, the boundary of the state is defined when the marginal productivity of a soldier equals the marginal productivity of a farmer. In short, Findlay posits that the boundary of the state is defined when the cost of defending the territory equals the benefits in terms of output [cf. also North's (1981:127) "'efficient'" size of the manor" in 10th century feudal Europe].

Another is Friedman (1977), which considers how different types of taxes affect the boundaries of the state. Friedman finds that trade taxes (e.g., customs duties) lead to larger state boundaries to internalise trade costs. Taxing labour (e.g., income tax) leads to boundaries that prevent labourers from leaving, i.e. closed boundaries. Finally, rents or land taxes are neutral as a "state, whatever its size, can tax up to the full economic rent of land and no more" (p. 61).

More recent political economy models by Alesina and Spolaore (1997; 2005), Bolton and Roland (1997), and Alesina (2003) are explicit in showing the trade-offs between economic and political benefits. Alesina and Spolaore (1997; 2005) and Alesina (2003) suggest that the optimal boundary of a state is one that reaches the highest level of average welfare for citizens. This level occurs when the costs of heterogeneity in large populations (e.g., diverse public good preferences) equal the benefits of size (e.g., economies of scale in public good provision). For example, as in Wittman (1991), the benefits of a large country include economies of scale in producing public goods (especially pure public goods such as national security), a larger market without trade restrictions, and the ability to provide social insurance among regions. However, once again, these benefits can be mitigated by the diseconomies of

the internal costs of running a large state (e.g., administrative and congestion costs). More importantly, a larger country is likely to have greater political costs in the form of heterogeneity of its citizens' preferences. Heterogeneity can affect the stability of the boundaries of a state because it reduces the proportion of citizens who are satisfied with central government policies. Assuming the median voter determines the public goods, the larger the state the more different the public goods will be from those preferred locally.

Bolton and Roland (1997) extend the Alesina and Spolaore (1997, 2005) framework by considering the specific effects of income distribution within the boundary of the state – for example, when a wealthier region wants to separate from a poorer region to avoid redistribution. The Bolton and Roland model shows that when regions have different preferences over public good policy, state boundaries can become smaller even if it is less economically efficient.

At this stage it is useful to discuss some of the common themes in this political economy literature on state boundaries. Firstly, the models are based on state boundary experience in modern states or in European history, viz. for the most part, open access orders in the sense of North, Wallis and Weingast (2009). Secondly, the common perspective in the literature (with the exception of Findlay 1996) is that there are economies of scale to be gained by providing public goods that should drive the boundaries of a state outwards. There is less agreement to the downside of expanded boundaries. Thirdly, the boundary of a state is discussed as a territorial border. Lastly, the models, especially in the more recent literature (e.g., Alesina 2003; Alesina and Spolaore 1997, 2005; Bolton and Roland 1997), consider democratic countries in which the tax paying, voting citizens are able to express their preferences for state boundaries through, for example, a referendum (think of Scotland in 2014). As will be shown in Chapter 2, the ability to express exit preferences through voting is not always an available option.

The political economy literature thus provides efficiency conditions for the boundary of a state. But why might political boundaries actually contract? To consider this question, a starting point is Hirschman's (1970) seminal work, on *Exit, Voice and Loyalty*. Hirschman provides two "behaviours" or "options" in the case of institutional dissatisfaction. "Exit" is simply leaving when the people believe that a better mix of goods can be obtained elsewhere. "Voice" is staying, but expressing dissatisfaction by complaining or protesting with the objective of improving the provision of goods. "Loyalty" has two possible interpretations: either as a behaviour, like exit and voice, or as an attitude that prevents exit and encourages voice. Hirschman's theory is based on the European and American experience regarding political dissatisfaction.

Buchanan and Faith's (1987) theory of *internal exit* builds on the economic theory of exit. Internal exit is defined as the "secession by a coalition of people from an existing political unit along with the establishment of a new political unit that will then provide public goods to those who defect from the original unit" (p. 1023). Thus, internal exit is a means through which people can form states that supply them with their most preferred institutional arrangements.

Internal exit goes a step beyond Hirschman's (1970) notion of exit to consider the case when people have the option of exit from dissatisfaction and will also go on to form a new state. However, the model of internal exit predicts that the threat of exit will act as a sufficient incentive for a rational government to set an internal exit-proof tax rate, i.e., a tax rate low enough to remove the incentive to exit. Essentially, the threat of exit is sufficient to prompt government to set the tax rate at a level sufficient to prevent state exit from actually happening. Thus, the model suggests that a state boundary is likely to be resistant to change, i.e., by preventing exit the state is able to at least maintain or perhaps ultimately increase its boundaries.

Friedman (1977) and Berkowitz (1997) reach conclusions similar to Buchanan and Faith. Friedman suggests that the threat of exit provides states with an incentive to expand boundaries or to build physical barriers, such as fences or walls, in order to increase the cost of exit (e.g., the minefields that surrounded former East Germany and Mozambique).

Based on the experience of the Soviet Union, Berkowitz (1997) extends Buchanan and Faith's internal exit model. Berkowitz uses a three-stage game in which the state sets the tax rate centrally and the peripheral regions can choose to exit. The model shows how the state will use redistribution and tax policy to prevent exit whenever possible. The peripheral regions take advantage of the state's reluctance to allow exit to influence centrally controlled policies by both threatening to withhold taxes and threatening to exit.

Thus, the literature presented here presents different efficiency arguments for state boundaries. In the case of Wittman (1991), the boundary will be reached when the costs of public good provision (which increases at a decreasing rate) equals the administrative costs of running the state. In Alesina and Spolaore (1997, 2005), the boundary occurs where the costs of public good provision equal the political costs of a heterogeneous population. In models with diseconomies of scale in public good provision, such Findlay's (1996), the boundary exists when it is no longer rational to defend territory: this boundary occurs when the members of the state are no longer able to make full productive use of the territory in terms of farming because all labour is directed at defending the state's territory from enemies outside the boundary.

Moreover, the literature is concerned with carrot and stick approaches to preventing boundary contraction by preventing exit. Buchanan and Faith (1987) provide carrots in the form of lower taxes to prevent exit. More commonly, though, the literature takes a stick approach to preventing exit. Bolton and Roland (1997: 1070-1071) highlight instances in which threat of exit leads states to actually increase taxes to use the higher revenue to prevent

exit. Friedman (1977) suggests expanding state boundaries by building fences and by other methods of increasing the costs of exit. Findlay (1996) and Wittman (1991) suggest military involvement to maintain state boundaries.

Chapter 2 shows in what way political boundary formation in Africa, in particular precolonial southern Africa, differs from most commonly known cases. Firstly, as mentioned, it considers documented historical cases of *natural states* (North, Wallis and Weingast 2009) in precolonial southern Africa instead of the more well-known modern or historical European experiences. Secondly, in contrast to Wittman (1991), Alesina and Spolaore (1997, 2005) and Bolton and Roland (1997), there are no economies of scale in terms of public good provision to be gained from increasing boundaries. Instead, the model presented in Chapter 2 is closer to Findlay (1996), in which there are diseconomies of scale. Thirdly, unlike the modern or historical European experience, in Chapter 2 the state boundaries under consideration are not defined by physical space or territory. Instead, an African state boundary is defined socially by people acknowledging the leadership of a ruler. Lastly, in contrast to the more recent literature (e.g., Alesina and Spolaore 1997, 2005; Bolton and Roland 1997; Alesina 2003) in which tax paying, voting citizens are able to express their preferences for state boundaries, in an African state more emphasis is given to the ruler and the ruling elite. Compared to democratic states, people in precolonial southern Africa have a more limited choice set. Instead of being able to elect suitable rulers (or to vote them out of positions of power), the only choice that African people have is that of choosing which ruler to follow. As Chapter 2 shows, the implication of this lack of choice is that, instead of considering what the incentives are to exit, we look at the incentives for potential rulers to leave with those people who will follow them.

The result of these differences is that many times in precolonial Africa internal exit could not be prevented. As the model presented in Chapter 2 shows, African internal exit results

from the payoff calculation of a member of the ruling elite's desire to maximise his share of public revenue surplus. Moreover, the model shows that an African state boundary is determined by the ability of the ruler to provide each of the other ruling elite members with a sufficient public revenue surplus share to remove the incentive to exit.

1.3 Internal Organisation of Government.

Experience suggests that, similarly to firms, governments are subject to problems of corruption, rent-seeking and shirking. At the heart of these problems is the problem of aligning incentives, e.g., between owners and managers in firms and between voters and elected officials in government. The theory of the firm suggests the use of contracts to address incentive problems (e.g., Williamson 1979; Grossman and Hart 1986; Hart 1995).

Within contract theory, there are two main approaches. The first relies on complete contracts; while the second considers incomplete contracts. Complete contracts "specify all parties' obligations in all future states of the world, to the fullest extent possible. As a result, there will never be a need for the parties to revise or renegotiate the contract as the future unfolds. The reason is that, if the parties ever changed or added a contract clause, this change or addition could have been anticipated and built into the original contract" (Hart 1995: 22). Essentially, a complete contract stipulates every obligation in every eventuality.

Hart (1995: 23) suggests that there are three problems with complete contracts. Firstly, it is difficult to plan for every eventuality that will take place in the future (e.g., very few people foresaw the Great Recession in 2007). Secondly, contracting parties will struggle to negotiate about events that may have never occurred or even existed before. Thirdly, it is very difficult to put every eventuality into writing in the contract in such a way as a third party (e.g., a judge) can understand and enforce it. Thus, complete contracts are problematic due to bounded rationality (Williamson, e.g., 1979).

These problems mean that all contracts are essentially incomplete, namely that it is impossible to include all eventualities in the contract. Hence, "the contract will contain gaps and missing provisions. In particular, it will be silent about the parties' obligations in some states of the world and will specify these obligations only coarsely or ambiguously in other states of the world" (Hart 1995: 23). For example, a contract between a firm and its supplier will likely not specify obligations in the case of an uncommon natural disaster, such as a tsunami, or a man-made, never before seen event, such as 9/11. Since an incomplete contract is unable to specify all eventualities, it must specify which party has the rights to decide how to act in unforeseen circumstances. These rights are known as *control rights* (e.g., Grossman and Hart 1986).

Tirole is one of the first to suggest applying theory of the firm insights, and in particular, the theory of contracts, to the internal organisation of government (Tirole 1994). As he writes, "there is little conceptual difference between governments and firms" (p. 1). Since there are both complete contract and incomplete contract theories, consider the application of each to the organisation of government.

If we apply the theory of complete contracts to government, then the government is a group of people who are "motivated by formal and complete incentive schemes" (Tirole 1994: 16). Taking this view means that a constitution as a contract between the government and citizens needs to specify every eventuality, e.g., every natural disaster and every conflict situation. However, for the same reasons that firm contracts are incomplete (bounded rationality), it is impossible for, e.g., any constitution to do that.

A more realistic approach is to apply the theory of incomplete contracts to government. For example, a constitution may specify welfare rights, such as those right for social security, but will not specify welfare amount for each individual citizen. As in the case of firms, when applying incomplete contracts to government we can consider the importance of control

rights, e.g., which government agency responds in the case of a natural disaster; or which government ministry makes the decisions regarding conflict with another country. As Seabright (1996) proposes, the "allocation of power matters, in short, precisely when it is not possible to specify in advance how the power should be exercised" (p. 65). With control rights, we can view the organisation of government as the "distribution of control rights over various kinds of decisions. This division is determined by constitutions, laws and tradition" (Tirole 1994: 16).

Tirole (1994) emphasises the need to distribute control rights over multiple government branches or levels to ensure accountability. Distributing rights over a single government branch can lead to "self-serving actions and capture" (p. 16). Thus, control rights should be divided over several branches or levels of government, e.g., the well-known separation of judicial, legislative and executive branches of government. Citizens "cannot take for granted that politicians and bureaucrats will act in their interests, and because they cannot write constitutions constraining precisely the ways in which they will do so, they need to allocate power carefully to those parties who will, in pursuing their own perceived interests, be most likely to further those of the citizens themselves" (Seabright 1996: 65).

The SGT of fiscal federalism is specifically concerned with the study of the organisation of government through the lens of the theory of the firm (e.g., Oates 2005; Weingast 2009). However, so far, the application of contract theory is limited. Seabright (1996) provides an incomplete contracts model of the appropriate degree of decentralisation in a federation. In the model, the incomplete contract is an election, which is considered incomplete because citizens and governments cannot verify the welfare of citizens due to both the actions of government and "various random shocks" (p. 65). Seabright's model shows that the choice to centralise or decentralise depends on the trade-off between policy coordination (reducing

interjurisdictional externalities) under centralisation and accountability to local jurisdictions under decentralisation.

Nonetheless, Oates (2005: 359) criticises the Seabright model for failing to take into account the primary fiscal instrument of intergovernmental fiscal relations, namely intergovernmental grants. Oates points out that matching grants are effective tools to solve the problem of interjurisdictional externalities. Thus, with the use of grants, in particular matching grants, Seabright's trade-off , says Oates, "effectively disappears" (p. 359).

Unlike Seabright, Garzarelli (2006) submits that an intergovernmental grant can be viewed as a SGT equivalent of an incomplete contract. This is different from the traditional approach to intergovernmental grants, which mirrors consumer behaviour theory (e.g., King 1984). Under the traditional approach, the income and substitution effects determine community spending, e.g., a lump-sum grant will only have an income effect for the grant recipient.

Instead, it is possible to consider intergovernmental grants as incomplete contracts in which the grant donor and the grant recipient are the contracting parties. Rather than a grant being simply a transfer of funds from one level of government to another, an intergovernmental grant conveys "a preferred spending pattern to the recipient" (Garzarelli 2006: 249).

The interpretation of a grant as an incomplete contract has two implications. Firstly, that all grants, irrespective of type, are conditional; for example, a lump-sum grant will only have a pure income effect if its payments are not dependent on the recipient's actions. Secondly, the extent of incompleteness varies with the type of grant; for example, a closed matching grant that specifies the good the grant should be spent on, the matching rate, and the maximum grant amount is more complete than an unconditional lump-sum grant which has none of these conditions (Garzarelli 2006: 249).

To view the incompleteness of the grant as leeway for the rational grant recipient to take advantage of is tantamount to viewing the leeway as a policy tool that provides a grant recipient with the freedom to experiment with policy and, in so doing, try to generate new knowledge to solve complex policy problems. Hence, Garzarelli suggests that viewing grants as incomplete contracts provides a role for grants in encouraging grant recipients to experiment with policies.

This thesis builds on Garzarelli (2006) by viewing intergovernmental grants as incomplete contracts. Chapter 3 presents a simple model to show how different types of intergovernmental grants can engender different levels of learning at the local government level. In particular, the model shows how a less complete grant, i.e., a lump-sum grant, can stimulate more recipient government experimentation than a more complete grant, i.e., a closed matching grant.

1.4 Conclusion

The economics of organisation can provide many insights for the political economy study of the boundary of the state and of the internal organisation of government. The view of the public sector as an *organization of organizations* by North, Wallis and Weingast (2009) provides a broader framework for research into state formation. Additionally, viewing the government through the lens of the theory of the firm provides a rich framework for analysing incentive problems in the internal organisation of government.

1.5 References

Alesina, Alberto 2003. "The Size of Countries: Does it Matter?," *Journal of the European Economic Association* 1(2-3): 301-316(April-May).

Alesina, Alberto and Enrico Spolaore 1997. "On the Number and Size of Nations," *Quarterly Journal of Economics* 112(4): 1027-1056 (November).

Alesina, Alberto and Enrico Spolaore 2005. *The Size of Nations* Cambridge, Massachusetts: MIT Press.

Berkowitz, Daniel 1997. "Regional Income and Secession: Center-periphery Relations in Emerging Market Economies," *Regional Science and Urban Economics* 27(1):17-45(February).

Bolton, Patrick and Gerard Roland 1997. "The Breakup of Nations: A Political Economy Analysis," *Quarterly Journal of Economics* 112(4):1057-1090 (November).

Bolton, Patrick, Gérard Roland, and Enrico Spolaore 1996. "Economic Theories of the Break-up and Integration of Nations," *European Economic Review* 40(3-5):697-705(April).

Breton, Albert and Ronald Wintrobe 1982. *The Logic of Bureaucratic Conduct: An Economic Analysis of Competition, Exchange, and Efficiency in Private and Public Organizations* Cambridge: Cambridge University Press.

Buchanan, James M. and Roger L. Faith 1987. "Secession and the Limits of Taxation: Toward a Theory of Internal Exit," *American Economic Review* 77(5): 1023-1031 (December).

Coase, R.H. 1937. "The Nature of the Firm," *Economica* 4(16):386-405(November).

Findlay, Ronald 1996. "Towards a Model of Territorial Expansion and the Limits of Empire," Chapter 3 in Michelle R. Garfinkel and Stergios Skaperkas (Eds.) *The Political Economy of Conflict and Appropriation* New York: Cambridge University Press.

Friedman, David 1977. "A Theory of the Size and Shape of Nations," *Journal of Political Economy* 85(1): 59-77(February).

Garzarelli, Giampaolo 2006. "Cognition, Incentives, and Public Governance: Laboratory Federalism from the Organizational Viewpoint," *Public Finance Review* 34(3): 235-257(May).

Grossman, Sanford J., and Oliver D. Hart. 1986. "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy* 94(4): 691-719 (August).

Hart, Oliver 1995. *Firms, Contracts and Financial Structure Oxford* Great Britain: Oxford University Press.

Hirschman, Albert O. 1970. *Exit, Voice and Loyalty: Responses to Decline in Firms, Organizations, and States* Cambridge, Massachusetts: Harvard University Press.

Holian, Matthew J. 2009. "Outsourcing in US Cities, Ambulances and Elderly Voters," *Public Choice* 141(3/4): 421-445(December).

King, David Neden. 1984. *Fiscal Tiers: The Economics of Multilevel Government* London: George Allen & Unwin.

Levin, Jonathan and Steven Tadelis 2010. "Contracting for Government Services: Theory and Evidence from U.S. Cities," *Journal of Industrial Economics* 58(3): 507-541(September).

López-de-Silanes, Florencio, Andrei Shleifer and Robert W. Vishny 1997. "Privatization in the United States," *RAND Journal of Economics* 28(3): 447–471(Autumn).

Oates, Wallace E. 1972. Fiscal Federalism. New York: Harcourt Brace Jovanovich.

Oates, Wallace E. 2005. "Towards a Second Generation Theory of Fiscal Federalism," *International Tax and Public Finance* 12(4):349-373(August).

North, Douglass C. 1981. *Structure and Change in Economic History* New York: W.W. Norton

North, Douglass C., John Joseph Wallis and Barry R. Weingast 2009. *Violence and Social Orders* New York: Cambridge University Press.

Seabright, Paul 1996. "Accountability and Decentralision in Government: An Incomplete Contracts Model," *European Economic Review* 40(1): 61-89(January).

Tirole, Jean 1994. "The Internal Organization of Government," *Oxford Economic Papers* 46(1): 1-29(January).

Weber, Max 1946. "Politics as a Vocation" in H.H. Gerth and C. Wright Mills (Eds.) *From Max Weber: Chapters in Sociology* New York: Oxford University Press. Available online: https://openlibrary.org/books/OL6498314M/From_Max_Weber_Chapters_in_sociology Last Accessed March 19, 2014.

Weingast, Barry R. 2009. "Second Generation Fiscal Federalism: The Implications of Fiscal Incentives," *Journal of Urban Economics* 65(3): 279-293(May).

Weingast, Barry R. and William J. Marshall 1988. The Industrial Organization of Congress; or, Why Legislatures, Like Firms, Are Not Organized as Markets. *Journal of Political Economy* 96(1): 132–163 (February).

Williamson, Oliver E. 1979. "Transaction-Cost Economics: The Governance of Contractual Relations," *Journal of Law and Economics* 22(2): 233-261(October).

Wittman, Donald 1991. "Nations and States: Mergers and Acquisitions; Dissolutions and Divorce," *American Economic Review* 81(2):126-129(May).

Chapter 2

Internal Exit and Natural State Formation in Precolonial Southern Africa (ca. 1500-1910)

2.1 Introduction

Political institutions in Africa are known to be different from political institutions elsewhere (e.g., Iliffe 2013[1995,2007]; Osafo-Kwaako and Robinson 2013). The study of African institutions is an important focus of recent economic research (e.g., Acemoglu, Johnson and Robinson 2002; Leeson 2005; Nunn 2008). However, one area that has not yet been fully explored is precolonial African state formation. This essay is about the formation of southern African states (ca. 1500-1910) with specific focus on the Ndebele, the Tawana, and the Zulu.[1]

One possible reason for the gap in economic research into African state formation is the classic Weberian (1946:78) definition of a state: a community that holds the "monopoly of the legitimate use of physical force within a given territory." While this definition may be useful for "Western" or "modern" states, it restricts our understanding of states to only those with well-defined territorial bounds.

Precolonial Africa is considered a frontier (Kopytoff 1989) in the sense that it "consists of politically open areas nestling between organized societies but 'internal' to the larger regions in which they are found – what might be called an *internal* or *interstitial frontier*" (p. 9). The frontier is "an institutional vacuum" in the sense that it is possible to start a different type of "social order" (p. 10) at relatively low cost. In this frontier, a precolonial southern African

[1] Southern African states typically settled in one of three regions in Southern Africa: along the coastline of what is now the KwaZulu-Natal province in South Africa and the southern region of Mozambique (e.g., the Zulu people); in the interior grasslands which extend from the Drakensberg, through the Highveld – which now lie in the provinces of the Free State, Gauteng, Mpumalanga, North West and KwaZulu Natal - into what is now Botswana and Zimbabwe (e.g., Sotho, Tswana and Venda people); and in the river valleys of what is now the Eastern Cape province in South Africa (e.g., Xhosa) (Peires (1983[1981]:139 and Iliffe 2013[1995,2007]: Chapter 6).

state was defined along social lines rather than by the territory it occupied (Schapera 1963[1956]: 28; Kopytoff 1989: 22-23). As such, an African state identified itself in terms of its ruler (e.g., a king or a chief). For the same reason a state would take the name of its first ruler (e.g., Zulu).[2]

Another reason for the gap in research is the anthropological guideline suggesting that a state materialises only once the population of a community exceeds 100,000 people (North, Wallis and Weingast 2009:53). Population density was much lower in Africa than in other parts of the world. For example, in 1750 the population density of sub-Saharan Africa was one tenth that of Europe (Herbst 2000: Table 1.1). Hence, we cannot apply either Weber's definition or the anthropological guideline to many precolonial southern African states.

In *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History,* North, Wallis and Weingast (2009) provide a useful framework of states from a (pre-state) "foraging order" to a "natural state" and then to an "open access order" (the most advanced form of state or what we consider a modern state). Since each of these types of states is defined along social rather than territorial lines, we can apply this typology to precolonial southern Africa. The closest type of state to the ones considered in this essay is a natural state or limited access order, which is characterized by a governing self-interested elite. The members of the elite limit access to privileges and, by limiting access, create incentives to respect each other's privileges. The elite also keeps violence in check as any violence will reduce privileges, especially rents extracted.

In addition to the presence of a self-interested elite and the reduced violence compared to a foraging order, precolonial southern African states can be classified as natural states for two other factual reasons. Firstly, a natural state can form from a population of one thousand people (North, Wallis and Weingast 2009: 53). This population size is also a more realistic

[2] See,e.g., Bryant (1905).

lower bound for precolonial southern African states. The size of the Zulu population at its origin (mid-17th century) is estimated at 2,000 (Barfield 1993:47); while, in 1795, Tswana states (Barolong, Kwena and Ngwato) had populations of up to 5,000 (Thompson 1995[1978]: 248). Secondly, a southern African state was an "organization of organizations" (North, Wallis and Weingast 2009: 17): it was governed by a primary ruler in conjunction with the self-interested elite. Members of the elite could be appointed by the ruler to lead wards, namely smaller administrative units of the state that helped collect the state's public revenue. For example, the 19th century Tawana state was divided into wards of approximately 10,000 people (Tlou 1977: 374).

North, Wallis and Weingast (2009: 21) suggest that, in a natural state, if there is a change in the balance of power among members of the elite, then the change will result in civil wars and revolutions. They do not consider another possibility: internal exit. Internal exit is the "secession by a coalition of people from an existing political unit along with the establishment of a new political unit that will then provide public goods to those who defect from the original unit" (Buchanan and Faith 1987: 1023).

This essay draws on Buchanan and Faith's (1987) theory of internal exit.[3] However, Buchanan and Faith predict that internal exit will not happen.[4] Yet I report cases of internal exit that took place in precolonial southern Africa. One well-known example is that of the leader of the Ndebele, Mzilikazi. In 1822, Mzilikazi (and people loyal to him) exited from the Zulu state and went on to eventually form the Ndebele (e.g., Brown 1969, Lye 1969b, Omer-Cooper 1969, Thompson 1969, Cobbing 1974, and Ndlovu-Gatsheni 2008).[5]

[3] This essay is also informed by Eggertsson (1991), McGuire and Olson (1996), and Bates, Greif and Singh (2002).

[4] State formation through organisational offshoot is examined in the modern literature on federalism and secession (Berkowitz 1997; Bolton and Roland 1997; Alesina and Spolaore 2005). The internal exit literature is the converse to "voting with the feet" (Tiebout 1956) and "exit" (Hirschman 1970) where there is an existing organisational alternative, i.e., an external exit, to defect to (e.g., another city, country).

[5] The process of internal exit is often referred to as "fission" in the African History literature. See, e.g., Hammond-Tooke (1964), Schapera (1968), Harinck (1969), Chanaiwa (1980), and Gulbrandsen (1993). The southern African history literature (e.g., Kopytoff 1989; Herbst 2000) postulates that internal exit is both costless and common. However, in precolonial southern Africa, it is not necessarily either of these two things. Firstly, as documented in this essay, internal

Besides historical cases, there is a legacy from internal exit in the form of path dependence (e.g., Nunn 2008). The Royal Bafokeng Nation that now owns land in the North West Province of South Africa originated from internal exit (Mbenga and Manson 2010: 2-7). The King of the Ndebele is still recognized by the South African government as a traditional leader, while the Ndebele are the second largest ethnic group in Zimbabwe. The Ndebele language – *isiNdebele* – is an official language in both South Africa and Zimbabwe (CIA 2015a,b). Essentially, internal exit occurred in precolonial southern Africa and led to the formation of new natural states, some of which still exist today.

Why is it that in Africa internal exit as a threat fails and internal exit still takes place? With the aid of a simple, historically informed model, this essay offers a political economy explanation of why internal exit took place in precolonial southern Africa. The explanation can be briefly summarised.

Disputes among the elites, i.e., members of the royal family, cause one or more members to exit the original state. An exiting member of the elite ("the exiter") is often accompanied by some relatives, e.g., his brothers, and people not necessarily related to him (immigrants, loyal commoners, servants, and slaves).[6] After exiting, the exiter and his people go on to form a new state in a new territory. As the new natural state ruler, the exiter is responsible for enforcing the rule of law as he sees fit. The exiter's people will form a new privileged

exit could have a high cost. Those who attempted internal exit could be hunted down and, if caught, killed by members of the original state. Furthermore, even those state members viewed as potential exiters could be executed (e.g., Lestrade 1930; Schapera 1963[1956]; Ndlovu-Gatsheni 2008). Secondly, internal exit is more common from approximately 1500 to 1800. During this time period, there were many internal exits that led to African states creating smaller natural states. The smaller states would spread out over large geographical areas. From 1800 until approximately 1910, it is less common for internal exit to take place. Instead it is more common for African states to amalgamate and to form larger natural states. This phase of amalgamation, typically, is attributed to the Mfecane, the increase of trade through Delogoa Bay and the encroachment of the European colonialists (Legassick 1969, Omer-Cooper 1969, Wilson 1969). In spite of Zulu efforts at state centralisation in this period the traditional political system in which elite members controlled wards remained. Thus, internal exits from the state still occurred (Omer-Cooper 1969: 212). The phenomenon of exit is not confined to southern Africa alone. Exit also occurred in other parts of Africa (e.g., Asiwaju 1976, Herbst 1990).

[6] I use male pronouns when referring to the ruler and other members of the elite. The reason for this is that most southern African states (e.g., the Zulu) were patriarchal societies. With a few exceptions, women were subordinate to men. Thus, like other patriarchal societies in Medieval Europe, women did not participate in governing the state directly and did not take on the roles of ruler or ward leader. However, it was possible for women to influence politics indirectly through their husbands, brothers and sons, as has been the case in Europe and elsewhere (Thomas (1970[1873]; Ndlovu-Gatsheni 2008: 85).

elite while other people would typically be commoners who have few rights or privileges. Often the new state will take the name of the exiter. For example, many Tswana states are named after their first rulers: the Bakgatla after Mokgatla, the Bakwena after Kwena, the Barolong after Morolong, Baphiring after Phiri, among others (TNAD 1905; Schapera 1963[1956], 1980). The same is true for the Venda states, e.g., a state ruled by Tshivhase is known as *vhahaTshivhase*, which means "Tshivhase's people" (Schapera 1963[1956]: 28). Hence, a new natural state can be formed through internal exit. Notwithstanding the empirical setting being preliterate states, we are still dealing with rational economic actors who respond to their surrounding incentives (Posner 1979).

This essay contends that there are four factors that contributed to the ease of internal exit in precolonial southern Africa. Firstly, the abundance of land makes it easier to form a new state (e.g., Herbst 2000; Iliffe 2013[1995,2007]). The second is the inverse of the first. The relative scarcity of people increases their value. The third is institutional in that the social institution of birthright grants privileges to the elite members of the polity only, facilitating exit. Finally, the fourth is that together land abundance and scarcity of people means that there are decreasing returns to the "broadcast of power" over the people by existing rulers (Herbst 2000), again easing exit. This essay points out that some exits can be from compromise (e.g., Coasian bargain) while others can result in conflict and even death. Whether exit actually takes place or not is a result of the payoff calculation from the exiter's desire to maximise his share of the public revenue surplus: his expected net benefit from exiting being greater than his expected net costs.[7]

[7] This is not to say that all internal exit decisions were purely economic. There were cases where other reasons led to exit. For example, the Mogôpa ruler, Tsoku (ca.1750-1760), is described as a cruel man. Mogôpa tradition suggests that when Tsoku had a live pregnant woman cut open for his own interest, many of his followers exited from his state. When Motshabi became ruler of the Birwa (ca.1830) he was "too cowardly" to participate in his own "installation ceremonies" (similar to a coronation for a European king), Masâke left the Birwa to start his own state. In the early 1900's, Chief Matume of the Phalaborwa, angered his people by "seducing their wives and having any men who complained killed" (Schapera 1963[1956]: 155). In response, his half-brother Selwana took most of the citizens of the state away to start a new one (Krige 1937: 339; Schapera 1963[1956]: 155).

2.2 Land, People, and the Right to Rule

2.2.1 African Land

African exit was unique because people could move from the original state to form a new state on unoccupied land. The African environment was conductive to this movement of people given the abundance of land and the methods of output used. As the land abundance thesis suggests, large parts of Africa were unoccupied because land of comparable quality (i.e., not particularly fertile) was a relatively abundant resource (Austin 2008; Hopkins 2009; Illife 2013[1995, 2007]). Hence land was not considered valuable. Consequently, property rights over land were either ill-defined or non-existent.[8] Therefore, the abundance of land means that internal exit can be a fairly costless option.

Africans relied primarily on the outputs from labour intensive activities, such as cattle farming, crops, and hunting. To see why these activities did not raise the cost of exit, consider each in turn. Cattle farming enabled internal exit as cattle were mobile and exiters could simply take the cattle with them upon exit (Omer-Cooper 1969; Sansom 1974[1937]: 274-275; Hall 1986: 85; Austin 2008: 601). Crops and hunting were seasonal outputs. Crops were planted and harvested in the wet season and the crop yields were not able to be stored. Hunting took place in the dry season when wild animals were easier to find as they remained close to water (Austin 2008 and Hopkins 2009). Hence, internal exit most likely took place during the dry season when the cost of exit was low: an exiter would not lose crop outputs and he could still engage in hunting while exiting.

[8] Cf. Demsetz 1967. Similarly, Anderson and Hill (2004) show in reference to another frontier, namely the *Wild West* (California and Nevada) in the 19th century, that when land has no value there is no incentive to develop property rights.

2.2.2 African People: Value and Birthright

Contrary to the abundance of land, "people were scarcer and, therefore, more important than land" (Chanaiwa 1980: 8). The scarcity of African people meant that people were valued, not just for labour, but also for fulfilling "social and political needs" (Kopytoff and Miers 1977: 69).[9] For a ruler like Shaka of the Zulu, the size of the state in terms of people was a reflection of his power (Schapera 1937: 180; Omer-Cooper 1969: 209; Kuper 1975a: 70; Chanaiwa 1980: 8; Ndlovu-Gatsheni 2008: 77-82).[10]

Typically, precolonial southern Africa states consisted of two groups of people, commoners (those unrelated to the ruler, slaves, servants, and immigrants) and elites (relatives of the ruler). Commoners and elites had different rights, determining the type of internal exit that occurred. The institution of birthright created options for the elite, but not for commoners. First, a member of the elite was of royal descent and eligible to rule a state in his own right; a commoner was not. Second, a member of the elite was able to attract people to follow his lead; again, a commoner was not. Thus, the initiator of internal exit was most likely a member of the elite rather than a commoner.

In point of fact, exit was not a viable option for a commoner as he would never be recognised as a ruler by other African people, i.e., he would not be able to acquire people. In other contexts (e.g., Europe), a commoner could become a member of the elite through other means, such as accumulating wealth (e.g., Buchanan and Faith 1987) or through marriage (e.g., marrying the daughter of the state ruler).[11] In the southern African case, despite the

[9] The scarcity and corresponding value attached to people has thus far been considered in the context of slavery (Kopytoff and Miers 1977; Fenske 2012). It has not yet been considered in the context of African state formation.

[10] Consequently, because people were valued highly, Africans developed property rights over people, especially relatives, commoners, women, and children. Property rights over people are neither unique to Africa nor are they limited to the slave trade. For example, for the greater part of "Western" history, a wife was considered her husband's property. Also, in the Roman Empire, a father had the right to kill or sell his children. Bride price (a payment from a groom to the bride's family) is still practiced in parts of Asia and in sub-Saharan Africa. Cf. Leeson, Boettke and Lemke (2014); Ashraf, Bau, Nunn and Voena (2015).

[11] In African states commoners could acquire status through cattle ownership and by marrying elite women (Chanaiwa 1980: 7; Ndlovu-Gatsheni 2008: 75). However, for the purpose of this essay, I do not consider these state members to be elite as they still will not have the right to rule a state (Richards 1961: 137).

availability of land, the only option available to a commoner was which elite member to follow.[12] Consider now how the ruler (and other members of the elite) exerts power over people in an African state.

2.2.3 Broadcast of Power and the African Elite

Economic theories of the state are typically based on regions in which land is a scarce resource (Friedman 1977; Wittman 1991; Findlay 1996).[13] For example, in considering the Romans, Han Chinese, and Ottomans, Findlay (1996) suggests that a ruler's ability to broadcast power is a land maximisation problem.[14] When land has value, it "is not simply given but has to be acquired and held against the tribe's enemies who surround it" (p. 42). As indicated, though, in Africa land had low value. A rational African ruler would not maximise land simply because the marginal cost of controlling the land would have far exceeded the (negligible) marginal benefit of doing so. Instead, as mentioned, people were the scarce, valuable resource. Thus, an African ruler faced a different problem: a population maximisation problem.

In parallel to Findlay (1996), one can think of a ruler as being in the centre of a "sphere of influence" from which emanates the ability to broadcast power (Herbst 1990) over people. Land still had a role to play, however. The abundance of land together with low population density meant that the ruler's people were spread out over long distances. It was easier for a ruler to maintain his power over the people closest to him. As the distance between the ruler and his people increased, it became increasingly difficult for the ruler to maintain his power

[12] Kopytoff (1989: 17; 36) suggests that there was a duality to the relationship between the commoners and the ruler as African commoners valued rulers in a similar way that rulers valued them. As the ruler's power lay in acquiring people to follow him, the people's power lay in choosing who to follow. Commoners without a ruler would be without value; while a commoner could get a higher status by following a stronger ruler (Kopytoff 1989: 17; 36).

[13] Even Tiebout's (1956) model of external exit assumes land is a fixed factor.

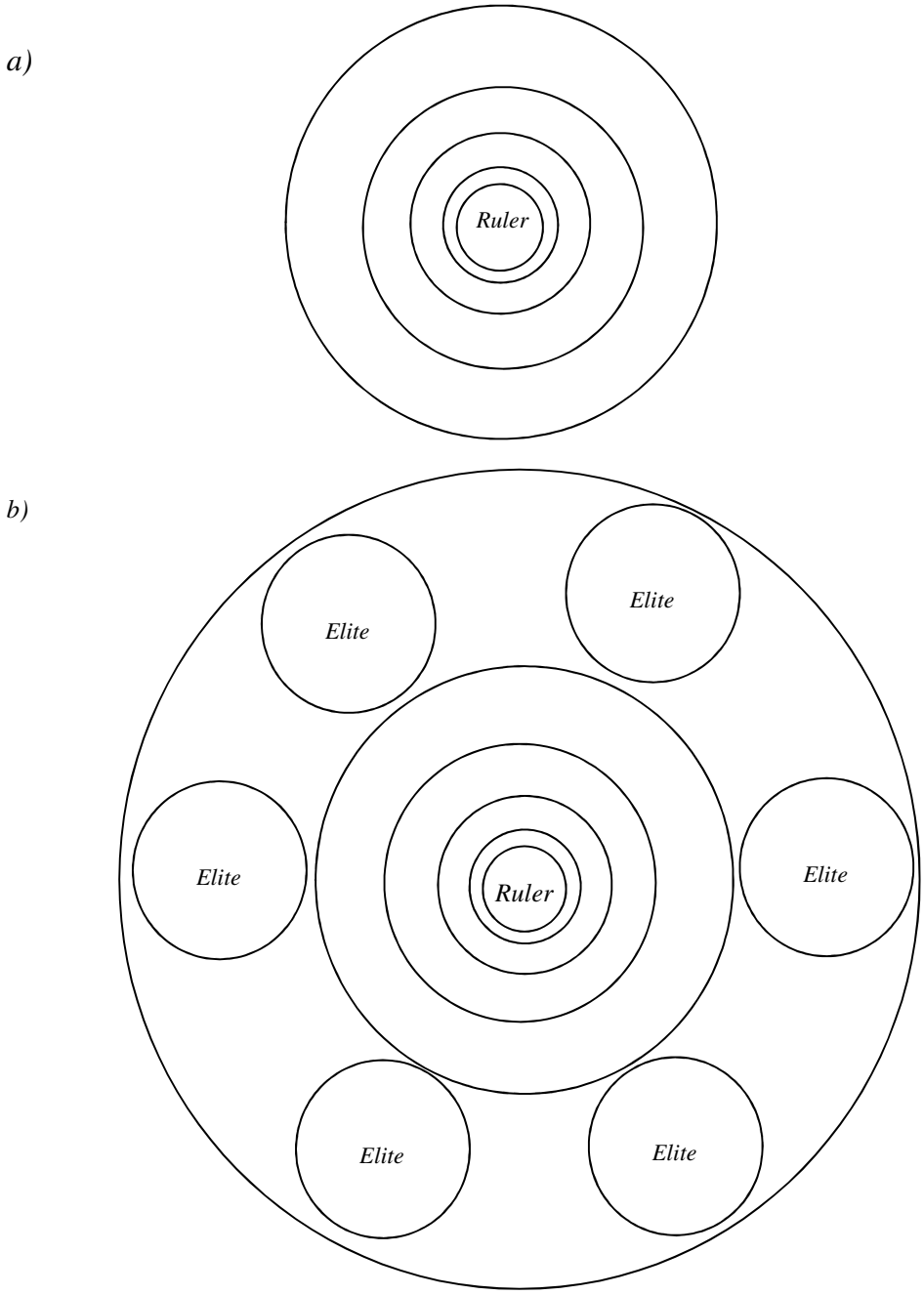
[14] This is a public organisation problem parallel to the Coasean (1937) limit to firm size which depends on increasing transactions costs.

over them (Herbst 2000; Iliffe (2013[1995, 2007]: 3).[15] The physical boundary of this broadcast was incidental though, mattering for broadcast of power reasons only and not for land value increases intrinsic to territorial expansion. Figure 2.1*a* illustrates the ability of the ruler to broadcast power. The ruler's power is higher over the people closer to him. As the distance from the power centre increases, the ruler's power starts to decrease (represented by the increasing distances between the circles) until an upper limit beyond which the people are no longer under his power.

One way for the ruler to increase his power was through the loyalty of other members of the elite. The elite included the ruler and other members of the elite who were usually the ruler's male relatives (brothers and uncles). Different members of the elite had different rights and privileges. The ruler had more rights and privileges than the rest of the elite as he was responsible for protecting and governing the state. The other members of the elite were expected to assist in these responsibilities and, in doing so, would increase the ruler's broadcast of power (Schapera 1937:180; Schapera 1963: 160-161; Uzoigwe 1977:47).

[15] One can think of the ruler's power as the Doppler effect. The closer the people are to the ruler the greater his power is felt but as they move further away the ruler's power diminishes.

Figure 2.1 The Ruler's Broadcast of Power



This type of broadcast of power is documented by Tlou (1977) in reference to the Tawana. The Tawana state (in what is now Botswana) was formed in approximately 1800 when a member of the Ngwato elite, Tawana, exited from the Ngwato state with his ward. Tawana and his people settled in the Kgwebe Hills where they expanded their population by asserting

power over other people in the area. From 1800 to at least 1847 (when Letsholathebe became the ruler of the Tawana) the Tawana state was not defined by land but rather by the power of the Tawana ruler over his people. The ruler was at the core of the state. The ruler enlisted other elite members to help assert his power (Figure 2.1*b*). However, the assistance of the other elite members decreased the extent to which the ruler's power was felt (Tlou 1977: 371-378). Thus, the other members of the elite would increase the ruler's power in absolute terms. However, at the margin, the ruler's power was likely to be diluted as he would have to share his power with the elite members who assist him. Moreover, using elite members to increase power could be a double-edged sword. The more power delegated to elite members, the more likely the same elite members could also weaken, or even destroy, the state by exiting with their own people and forming a new state.

2.2.4 A Paradigmatic Illustration

Mzilikazi was an elite member who initially helped increase Shaka's broadcast of power among the Zulu people by commanding his own people (the Khumalo), but ultimately chose to exit with his people to form his own state.[16] In 1822, after completing a raid, Mzilikazi signalled his intention to exit by refusing to pay taxes on the raid (Becker 1966: 36; Thompson 1969: 347). Mzilikazi then exited from the Zulu with between two and three hundred warriors plus an undetermined number of women and children. After Mzilikazi's exit, another member of the elite was sent after him and his people to force return to the Zulu. As a result, Mzilikazi and his people were unable to settle immediately as they had to first evade the Zulu elite members pursuing them. Having successfully evaded pursuit, by 1829

[16] Historical details of internal exits from states other than Mzilikazi's Ndebele in southern Africa are limited. African history tends to emphasize the successes of particular leaders, like Mzilikazi, while lacking detailed information on other people or events, especially unsuccessful exiters. Nevertheless, we can still find historical evidence that internal exit from many southern African states occurred. Rasmussen (1977) expresses a similar sentiment. The primary sources of internal exit are from missionaries who observed the African states (e.g., Thomas 1970[1873]); while secondary sources rely on oral tradition (e.g., Schapera 1963[1956]; Sansom 1974[1937]).

the Ndebele population had grown to an estimated 60,000 to 80,000 people of which between 1,000 and 5,000 were warriors (1936 estimate) (Lye 1969b: 96).

Figure 2.2 Map of Mzilikazi's Internal Exit



Figure 2.2 shows the route Mzilikazi took after exiting from the Zulu in the context of current African borders. The Zulu state was located near the east coast of what is now South Africa. The route shows how Mzilikazi and his people headed inland from the east coast, eventually crossing the Limpopo River (which is now the border between South Africa and its north eastern neighbours, Zimbabwe and Mozambique).[17] Ultimately, it took approximately twenty-five years for Mzilikazi and his people to form a new state, i.e., the

[17] For more details on the route that Mzilikazi took and his efforts to evade the Zulu, see Lye (1969b), Cobbing (1974) and Chanaiwa (1976). The Ndebele population grew rapidly after exit as the state incorporated Nguni and Sotho people (Lye 1969b: 96).

Ndebele, in what is now Zimbabwe (Thomas 1970[1873]: 162; Cobbing 1974: 615). By 1880, the population of the Ndebele was estimated to be between 150,000 and 200,000 people (Chanaiwa 1976: 60).

2.3 Model

Based on the historical evidence, we can isolate three exogenous factors of African uniqueness that contributed to the ease of internal exit: land abundance; birthright; and decreasing returns to the broadcast of power. The abundant land allowed exit to be a physical movement of people from one area to another. Birthright gave the ruler the right to rule but also entitled other elite members to rule over people. The decreasing returns to the broadcast of power necessitated that a ruler delegate power to the same elite members with the right to rule.

An additional factor unique to Africa was the relative scarcity of people which increased their value. In strictly economic terms, people were valuable to a ruler because they could generate a surplus. Hence, the ruler would try to maximise the number of people which in turn would maximise the surplus. However, to increase his broadcast of power over people and, consequently, the surplus extracted, the ruler needed assistance from other elite members. The other elite members would help to extract a surplus from the state in exchange for a share of the surplus.

The other elite members were rational in that, similarly to the ruler, they were surplus share maximisers. Since these other elite members were entitled to rule their own state, when the surplus they could expect to extract from their own newly created state exceeded the surplus share allocation from the ruler (and any possible costs of exit, e.g., conflict), the rational elite member would exercise the option of internal exit. Whether or not the ruler of the original state decided to fight the exit (and thereby increase the cost of exit) depended on

the difference between the loss of surplus due to the loss of people and the gain in the share of the surplus allocated to the elite exiter.

2.3.1 The African State as a Public Organisation

Consider a state of N people, $N \geq 1000$. The N people consist of two groups, the elite, E, and commoners, M, such that $E + M = N$. In economic terms, the difference between the elite and the commoners is that the members of the elite share in the public revenue surplus of the state while the commoners do not. One of the members of the elite is the ruler of the state. The ruler is unique within the elite because he sets the tax rate and establishes the share of the surplus allocated to himself, and to each of the other elite members.

The history literature suggests that in the southern African states before 1790 the ruler supplied public goods, such as law and spirituality. Subsequently, a ruler's responsibilities also included protecting his people from enemies, rain-making, managing grain production, allocating cattle, and regulating the ivory trade (Chanaiwa 1980: 8; Ndlovu-Gatsheni 2008: 75). Similarly to others (e.g., Bates, Greif, and Singh 2002), for simplicity I assume that these public goods can be captured by a simple public good: order. The cost of order depends on the size of the population,

$$G = G(N). \qquad (1)$$

As mentioned, the scarce populations spread over vast distances lead to the ruler facing decreasing returns to the broadcast of power. These same conditions mean that the cost of providing order increases at an increasing rate as the size of the population increases, i.e., $\frac{dG}{dN} > 0$ and $\frac{d^2G}{dN^2} > 0$ (Figure 2.3*a*).

Public revenue is obtained by taxing output which, as indicated, historically took the form of crops, livestock, and hunting. I assume that, since the ruler provides order, the people

have an incentive to produce output beyond subsistence levels. As output was strictly labour-intensive, the state's output is a function of the population:[18]

$$Y = Y(N). \quad (2)$$

A Venda proverb says, "(y)ou cannot appear before the lion without a piece of firewood in your hand," meaning that a person may not appear before the ruler of a state without paying a tax (Schapera 1963[1956]:101). Taxes in states such as the Zulu took the form of percentage shares of output.[19] Taxes could also be paid in nature, e.g., directly in the form of labour by tending the ruler's cattle or tending his crops (Guy 1983[1981]:43).

Furthermore, southern African history indicates that the amount of tax due was proportional to the output of the state. For example, in the Luapula area (now Northern Zimbabwe), a member of an established state could have a higher tax rate than the member of a newly established state (Vansina 1962:327). Thus, tax is proportional at a rate of τ, where $\tau \in (0,1)$.[20] It follows that public revenue T is obtained through taxing income $Y(N)$ at a rate of τ,

$$T = \tau Y(N), \quad (3)$$

Analogously to providing order, public revenue also exhibits decreasing returns given the difficulties of collecting taxes from sparsely settled people over large distances, i.e., $\frac{dT}{dN} > 0$ and $\frac{d^2T}{dN^2} < 0$ (Figure 2.3*a*).

[18] Every person of the same gender provided essentially one labour unit: the marginal benefit of each additional unit of labour was constant (Kopytoff 1989: 42).

[19] A well-known example of a tax is the Sotho-Tswana's first-fruits ceremony in which the ruler "bit into the first fruits of the harvest," (Legassick 1969: 25), i.e., the Sotho-Tswana ruler would accept tax in the form of a percentage of the total harvest. Beer made from grain was another way to pay taxes on crops. Members of the Lobedu state and the Basotho could pay taxes in the form of pots of beer brewed from maize and sorghum. An example of a hunting tax was a percentage share of the hunting returns, e.g., "one tusk of every elephant killed, and the breast of every ox or large wild animal slaughtered" (Legassick 1969: 25). For other examples of taxes see Hunt (1931), Schapera (1963[1956]: 99-100), Ballard (1983[1981]), and Gulbrandson (1993). More specific taxes could also apply: the Zulu had another unique form of tax in that each large household was required to give an adult daughter to the ruler (Gulbrandson 1993:556; Schapera 1963[1956]: 100-101).

[20] I use the term tax instead of tribute or rent. The relationship between the ruler and the citizens of African states is one of exchange in which the citizen tax payers expected something in return for taxes, usually protection and public order. Hence this is not unlike tax systems that exist today (Schapera 1973[1953]: 26).

Each member of the state, whether elite or not, thus contributes to public revenue. This assumption is in line with Buchanan and Faith (1987) and, more importantly, is historically accurate. The ruler of the state was expected to use his own income for the good of the state, e.g., redistribution (e.g., Vansina 1962: 327; Schapera 1963[1956]: 99-100; Legassick 1969: 51-52 regarding Sotho-Tswana redistribution; Hammond-Tooke 1985; and Ndlovu-Gatsheni 2008: 80-81 regarding Ndebele redistribution). The historical evidence is also in line with the McGuire and Olson (1996) model of autocracy: the rational ruler implements redistribution because it allows him to increase the number of people in his polity, and hence ultimately increase the public revenue he can extract.

Historically, the rational southern African ruler will choose to extract more public revenue than was required to simply cover the cost of providing the public good, i.e., the ruler would extract a surplus. The difference between public revenue (3) and the cost of providing order (1) is the surplus (i.e., the net benefit of providing order):

$$S = T - G, \tag{4}$$

or,

$$S = \tau Y(N) - G(N). \tag{5}$$

The ruler's problem is therefore:

Max: $S = \tau Y(N) - G(N)$.

s.t.: $\tau Y(N) > G(N)$

as shown in Figure 2.3*b*. An African state has a minimum population size (e.g., $N = 1000$) and a maximum population size N_{max}.[21] For very large populations ($N > N_{max}$) the ruler will not provide order because the cost of providing order will outweigh the benefit (i.e.,

[21] Populations below 1000 people are considered foraging orders rather than natural states (North, Wallis, and Weingast 2009). We can assume that when the population is smaller than 1000 that the ruler will not be able to extract sufficient surplus to justify his incentive to rule.

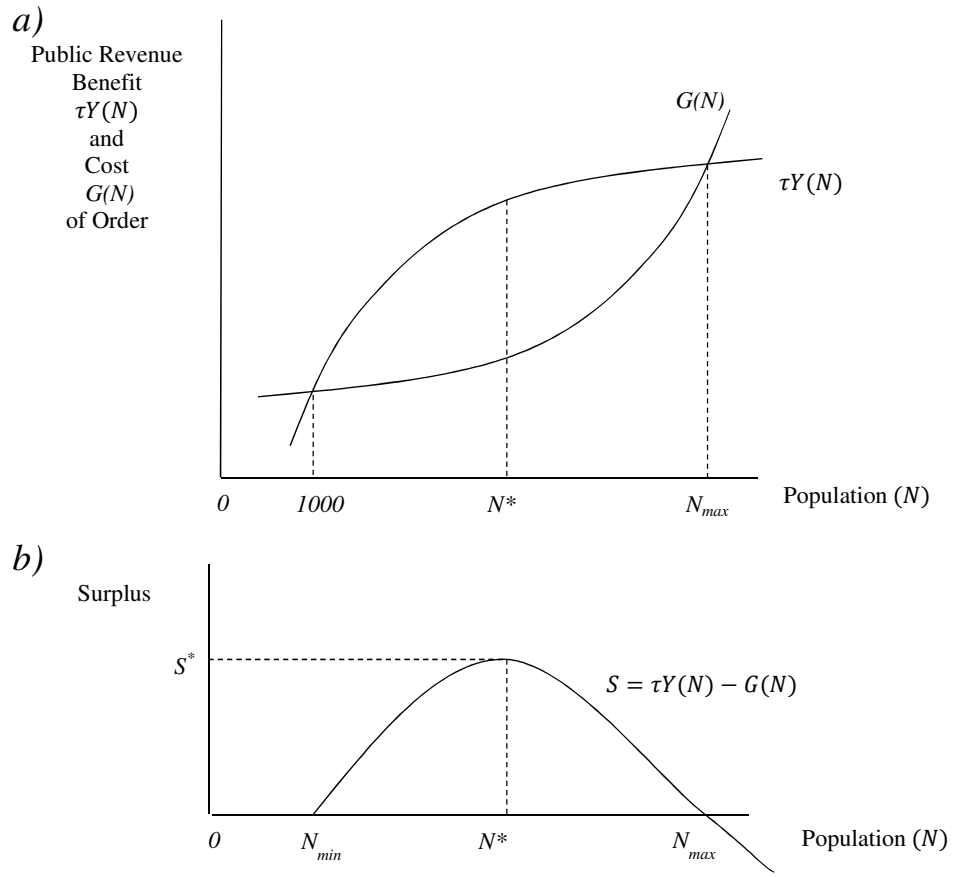
$\tau Y(N) < G(N)$) and the ruler will be unable to extract a surplus. The maximum population size corresponds to the boundary of the broadcast of power in Figure 2.1.

Within these bounds, $1000 \leq N \leq N_{max}$, the ruler will form a state over people as he will be able to extract a surplus.[22] However, because of the difficulties of broadcasting power in terms of both extracting revenue and providing order, the surplus extracted depends on the population size. An optimal point exists at N^* at which point the population size is such that the ruler can maximise the difference between the public revenue collected and the cost of providing order. For a population $1000 < N < N^*$, the surplus increases at a decreasing rate as the population increases, $\frac{\partial S}{\partial N} > 0$, $\frac{\partial^2 S}{\partial N^2} < 0$. After reaching a maximum at N^*, for a population $N^* < N < N_{\max}$, the surplus decreases at an increasing rate as the population increases, $\frac{\partial S}{\partial N} < 0$, $\frac{\partial^2 S}{\partial N^2} > 0$ (Figure 2.3*b*).

Accordingly, if the population size is N^*, then any change in the population size will be detrimental to the total surplus extracted. However, if $N \neq N^*$, then a change in population size can allow the ruler to extract more or less surplus: if $1000 \leq N < N^*$ then a decrease in the population can result in a smaller surplus; if $N^* < N \leq N_{\max}$, then a decrease in the population can result in a larger surplus. To increase the broadcast of power, the ruler enlists the help of the other members of the elite. It is at this juncture that the African state begins to move away from a monolithic structure as in Olson's (1993) bandit theory, to an organisation of organisations as in North, Wallis and Weingast's (2009) theory.

[22] For a given *N*, the rational ruler will set the amount of order and the corresponding tax rate to maximise the surplus extracted.

Figure 2.3 Population Size and Surplus



The other elite members help the ruler to extract a surplus from the rest of the state. In return, the surplus will be shared among the elite members of the state. The surplus allows both the ruler and the other elite members to have higher levels of wealth than commoners. In turn, the elite will use this wealth to marry more women than commoners and to acquire power by attracting more people (Guy 1983[1981]:43; Omer-Cooper 1994: 13; Ndlovu-Gatsheni 2008: 75).

Once the elite members are enlisted to help extract a surplus, the elite members will start to form their own groups of people. Many southern African states (e.g., Ndebele, Tawana, and Zulu) were organised into smaller groups, namely, an African state was an organisation

of organisations (e.g., Kopytoff 1989). The smaller groups were known as wards.[23] A ward was a smaller administrative unit for, e.g., the allocation of pastures and for the adjudication of minor crimes (theft, domestic disputes) populated by artisans, clans, conquered people, families, immigrants, warriors, etc. It was governed by its own elite head, often with the aid of other elite members, and contributed to the central fisc (Schapera 1937:180; Omer-Cooper 1969:209; Cobbing 1974; Kuper 1975a,b; Chanaiwa 1980; Ndlovu-Gatsheni 2008: 75). For example, the Zulu state elite included the *izikhulu* who were the elite ward leaders, like Mzilikazi (Guy 1983[1981]:43).

A ward could be run as a "mini-state" within the state. As noted, a member of the elite would lead the ward, assisted by other members of the elite; for example, a Sotho ward leader was assisted by relatives by marriage and relatives on his mother's side. The rest of the ward would consist of citizens who were commoners. A ward leader could acquire power by increasing the number of people in his ward, i.e., by increasing the broadcast of his own power (Schapera 1937:180; Omer-Cooper 1969:209; Kuper 1975a: 70; Chanaiwa 1980:8; Ndlovu-Gatsheni 2008: 77-82).

As mentioned, in exchange for helping to extract a surplus from the wards, the other elite members would share in the surplus. The proportion of the surplus that elite member i, where $i = 1 \ldots E$ and $i \in \mathbb{N}$, receives is $\sigma_i \in (0,1)$, where $\sum_{i=1}^{E} \sigma_i = 1$, Since the ruler is the primary member of the elite, let his share be σ_1. The ruler's share is thus $S_1 = \sigma_1 S^*$. The ruler would typically take the largest surplus share and would typically be the wealthiest man in the state as a result. In 1932, the ruler of the Kgatla owned one-seventh of his state's income: he received one thousand two hundred bags of corn annually, thirty cattle from any stock that had strayed, and between twenty and forty cattle in fines. Similarly, in 1936, the Swazi ruler was estimated to own ten times as many cattle as the next wealthiest Swazi

[23] African states could have complex political divisions, e.g., a state divided into villages, in turn divided into wards, and then into homesteads. This essay uses the term ward for two reasons: for simplicity and because Mzilikazi was the leader of the Khumalo ward in the Zulu state (Hoernlé 1962[1937]: 86-90).

citizen (Schapera 1963[1956]: 101-102). Even today there is a legacy in terms of surplus shares allocated to traditional rulers (and other elite members). King Goodwill Zwelithini and six other traditional rulers in South Africa still receive annual surplus shares from public revenue ("How Much You Pay for South Africa's Royal Families" 2015)

The other members of the elite each receive a fraction, σ_i, of the remaining surplus $(1-\sigma_1)S^*$, apportioned at the ruler's discretion. That is to say that, the ruler can divide the surplus shares equally among members of the elite, but this is not necessarily the case. For example, an elite member's surplus share might be proportional to his ward size. Thus, each of the other elite members receives a surplus share $S_i = \sigma_i S^*$, where $\sum_{i=2}^{E} S_i = (1-\sigma_1)S^*$. Thus, surplus shares are constrained by the surplus share taken by the ruler, σ_1, the size of the surplus, S^*, the size of the elite, E, and the discretion of the ruler. The Ndebele elite consisted of approximately between 22,500 and 30,000 members (15% of the total population) in 1880 (Chanaiwa 1976: 60).

2.3.2 Internal Exits: Compromise or Conflict

Suppose that one elite member (e.g., a ward leader) is considering exiting the state with his own people, namely, he is a *potential exiter* like Mzilikazi. The potential exiter will take people who are elite, F, and people who are commoners, X.

In deciding whether to exit, the potential exiter will weigh up the expected costs and benefits of exit. In precolonial southern Africa the ruler's response to the exit influenced the expected cost of exit. When a ruler objected to exit, an exiter would be punished for exiting or even for considering exiting.

The circumstances of the documented cases of exit are very different: some occurred through compromise (e.g., Xhosa, Sotho, Venda,) while others ended with violent conflict (e.g., Zulu, Tswana). Exit through compromise is documented among the Xhosa, Sotho and

Venda. With the support of the other members of the elite, a ruler's sons and their people would often exit the original state in order to form their own state (Lye 1969a; Peires 1983[1981]; Iliffe 2013[1995,2007]). For example, sons of the ruler of the Xhosa would each set up their own polity, which was known as their own "Great Place" (Peires 1983[1981]). Venda rulers were considered restrained because the ruler would first try to negotiate a compromise before resorting to force (Lestrade 1930:321).[24] Whether the compromise worked or not would depend on the number of people exiting compared to the number of people remaining in the original state.

Zulu and Tswana exits could be far more costly in terms of the punishment inflicted by the existing rulers. Exiters could still be punished for internal exit if caught by the original polity. Exiters and potential exiters were well aware of the possible punishments. For instance, if caught, exiters from the Zulu state were executed. When Mzilikazi exited from the Zulu, he and his people had to "run for their lives to escape Shaka's reach" (Chanaiwa 1976: 53). In fact, many of Mzilikazi's people were executed for exiting at enTubeni Hills (see Figure 2.2).

Subsequently, despite being an exiter himself, once he had established the Ndebele, Mzilikazi would execute any exiters, and even those he considered potential exiters (Ndlovu-Gatsheni 2008). For example, oral history reports that Gundwane was an exiter from the Ndebele. He travelled with his people over the Limpopo River. However, when Mzilikazi arrived at Gundwane's state, he had Gundwane executed and incorporated Gundwane's people into the Ndebele (Cobbing 1974; Rasmussen 1977).

If a Tswana ruler learnt of an elite member's intention to exit, he would confiscate that elite member's cattle to try to discourage exit. If a Tswana member of the elite did lead his people to exit, the Tswana ruler would try to bring the exiters back by force. If the exiters

[24] Cf. Coase (1960).

won, they could set up their own state, but would bear a cost for exiting (e.g., loss of people during a conflict). If the exiters lost, they would be forced to return to the original state and be punished (Schapera 1963[1956]: 154-165; 1980). For example, when the Ngwaketse broke away from the Kwena state (c. 1750) and when the Tawana exited from the Ngwato state (c. 1795), the state rulers sent men after them but these men were defeated. Both the Ngwaketse and the Tawana, thus, succeeded in exiting and setting up their own states. Later, around 1840, Phethu and Bathoen attempted to exit from their half-brother, the Ngwato state ruler, Sekgomo I. Sekgomo I reacted to the exit by pursuing and killing them both before bringing their people back to the original state (Schapera 1963[1956]: 154-165, 1980; Mönnig 1988).

Since the potential exiter takes both members of the elite and commoners, the exit generates an externality in the form of a change in the surplus in the original polity. The loss of the members of the elite leads to fewer elite members remaining to share the surplus with. But the loss of both elite members and commoners lowers the possible tax revenue pool.

Historically, the ruler will thus support, be indifferent about, or oppose the internal exit. If the ruler supports exit or, at least, is indifferent, then the potential exiter will be able to exit at negligible cost. However, if the ruler opposes exit, then internal exit will come at a cost, usually in terms of a loss of people in a conflict.

Suppose the ruler's response to the exit depends on the effect of internal exit on his surplus share, S_1. Further suppose that, at least initially, the tax rate stays the same.

The change in the ruler's surplus share due to the loss of people (both commoner and elite) that then exit is:

$$\frac{dS_1}{dN} = \sigma_1\left(\tau\frac{dY}{dN} - \frac{dG}{dN}\right) + \frac{d\sigma_1}{dN}(\tau Y(N) - G(N)). \quad (6)$$

The change in the ruler's surplus share depends on two changes shown in Equation (6). The first, $\sigma_1\left(\tau\frac{dY}{dN} - \frac{dG}{dN}\right)$, is the proportion of the surplus change due to exit that is allocated to the ruler. Note that, as discussed, depending on the size of N at the time of exit, the term

$\left(\tau\frac{dY}{dN}-\frac{dG}{dN}\right)$ can be positive or negative. The second, $\frac{d\sigma_1}{dN}(\tau Y(N)-G(N))$, is the change in the proportion of the surplus that the ruler is allocated. Since some of the exiters are elite members, the ruler can take a greater proportion of the surplus after exit. Consider the implications of exit on $\frac{dS_1}{dN}$.

If $\frac{dS_1}{dN}\geq 0$, then the change in the surplus share allocated to the ruler counterbalances the change in the surplus. For example, if the exit leads to a smaller state surplus, but the ruler gets a larger share of that surplus, then his surplus share will be unchanged or, in fact, could increase. Thus, the ruler will be unaffected, or even better off, after the internal exit. Hence the ruler will be either indifferent or unlikely to oppose the exit. In this case a potential exiter will be able to leave costlessly (similar to a Coasean bargain), e.g., Sotho and Xhosa internal exits.

If $\frac{dS_1}{dN}<0$, then change in the surplus share allocated to the ruler does not counterbalance the change in the surplus. In this case, if the size of the surplus decreases due to the decrease in population size, but the ruler does not get a larger share of that surplus, and he will be worse off than before the exit. Hence, the ruler has an incentive to prevent exit and will impose costs on any potential exits (e.g., Mzilikazi's exit from the Zulu). Historically, the cost of exit was conflict which could have two possible outcomes. The first is that the exiter defeats the ruler and was able to exit successfully, though not costlessly. The second is that the ruler defeats the exiter and his supporters, in which case the exiter is unable to exit and, instead, would bear the cost of punishment for attempting to exit. The ruler's response and the potential costs of exit will influence the potential exiter's decision.

2.3.3 The Potential Exiter's Decision: To Exit or Not?

The potential exiter (e.g, Mzilikazi) faces a decision between the known outcome of staying and the unknown outcome of exiting.[25] See the decision tree in Figure 2.4. If the potential exiter decides not to exit (Outcome 1), then he will continue to receive his surplus share, i.e., $S_{i,i\neq 1}$, i.e., a new natural state will not be formed. If he decides to exit, he will have an expected payoff from the internal exit $E(exit)$ which, as mentioned, depends on the response of the ruler and the outcome of any conflict that may occur (Outcomes 2, 3 and 4).

Suppose that with probability π_1, the ruler (e.g., Shaka) either supports or is indifferent about exit (i.e., when the ruler's surplus share increases or remains unchanged after exit, $\frac{dS_1}{dN} \geq 0$). In this outcome, the potential exiter then can expect, upon internal exiting, to receive a share of the surplus in the new state that he establishes. Thus, with π_1, the potential exiter can expect a payoff of Ψ_X, which is the surplus share the potential exiter expects to receive in his new state (outcome 2). Hence, outcome 2 leads to natural state formation.[26]

Now, with probability $1-\pi_1$, the ruler will oppose internal exit (i.e., when the ruler's surplus share decreases due to the exit, $\frac{dS_1}{dN} < 0$). There are two possible outcomes. Firstly, with probability π_2, that the potential exiter defeats the ruler and is able to exit. Consequently, with π_2, the potential exiter can expect a payoff of $\Psi_X - C$, which is the share of the surplus in the new state that he establishes less the cost of the conflict. Hence, outcome 3 also leads to the formation of a new natural state (e.g., the Ndebele).

Secondly, with probability π_3, the potential exiter's attempt at internal exit fails as he and his people are unable to defeat the ruler. Hence, the potential exiter can expect to be

[25] In parallel to the state as an organisation of organisations, we can think of the exiter as a political entrepreneur (Kopytoff 1989: 17). As a political entrepreneur, an exiter like Mzilikazi had two goals: achieving independence from the original state and acquiring people. Like the entrepreneurs conceptualised by Knight (2006 [1921]), these political entrepreneurs faced a trade-off: the uncertainty of creating new natural state on the African frontier in exchange for expected future rewards in the form of surpluses. Cf. Klein, Mahoney, McGahan, and Pitelis (2010).

[26] The potential exiter will be the ruler in the new state he establishes. Thus, the exiter will be able to determine the amount of order, the tax rate and the surplus shares allocated, in the same way that the ruler does in the original state.

punished by the ruler. Thus, with π_3, the potential exiter can expect a negative payoff of $-P$, where P is the punishment inflicted upon the potential exiter by ruler. Outcome 4, therefore, does not lead to natural state formation.

The expected payoff from internal exit for the potential exiter, i.e., the sum of outcomes 2,3, and 4, is:

$$\begin{aligned} E(internal\ exit) &= \pi_1(supported\ exit) \\ &+ \pi_2(opposed, successful\ exit) \\ &+ \pi_3(opposed, failed\ exit), \end{aligned}$$

or,

$$E(internal\ exit) = \pi_1(\Psi_X) + \pi_2(\Psi_X - C) + \pi_3(-P), \quad (7)$$

where $\sum_{i=1}^{3} \pi_i = 1$.

To decide whether or not to exit from the original state, the potential exiter will compare his surplus share from staying with the state with his expected payoff from exiting, i.e., the potential exiter will compare S_i with equation (6).

The potential exiter will be an exiter if

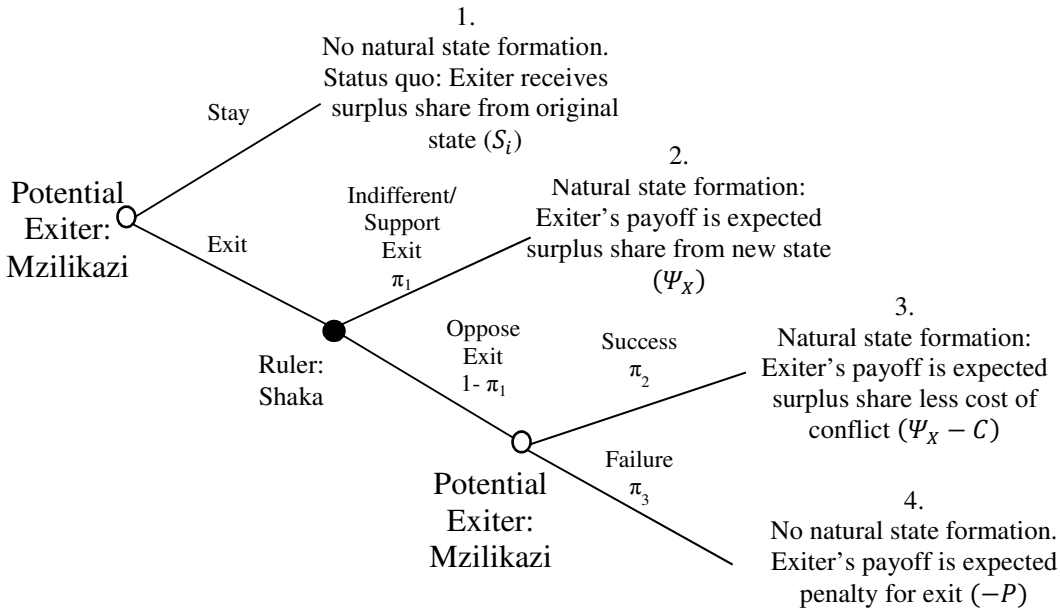
$$S_i < \pi_1(\Psi_X) + \pi_2(\Psi_X - C) + \pi_3(-P)\ . \quad (8)$$

Since $\pi_1 + \pi_2 + \pi_3 = 1$, equation (8) can be restated as

$$(\pi_1 + \pi_2)\Psi_X - S_i > \pi_2 C + \pi_3 P. \quad (9)$$

Equation (9) shows that if expected net increase in surplus share from internal exit, $(\pi_1 + \pi_2)\Psi_X - S_i$, is greater than the expected cost of exit, $\pi_2 C + \pi_3 P$, then the potential exiter will choose to exit and form new natural state.

Figure 2.4 Internal Exit Decision Tree



2.3.4 Can the Ruler Prevent Exit?

Exiting and setting up a new state is a risky undertaking. The state ruler knows that a potential exiter may choose to avoid the risk of internal exit if given a sufficient surplus share. The surplus share sufficient to prevent exit is the certainty equivalent of internal exit. To determine the certainty equivalent, first consider the expected utility of internal exit that corresponds to the expected utility of the payoff of exit for the potential exiter, i.e.,

$$EU(internal\ exit) = EU(\pi_1(\Psi_X) + \pi_2(\Psi_X - C) + \pi_3(-P)). \quad (10)$$

The certainty equivalent, ξ_i, is, thus, the surplus share that provides the same utility as the expected utility of internal exit, i.e.,

$U(\xi_i) = EU(internal\ exit)$, or,

$$U(\xi_i) = EU(\pi_1(\Psi_X) + \pi_2(\Psi_X - C) + \pi_3(-P)). \quad (11)$$

If the ruler provides a surplus share greater than or equal to the certainty equivalent, i.e., if $S_i \geq \xi_i$, then a potential exiter will choose not to exit. Similarly, if the ruler provides a

surplus share less than the certainty equivalent, i.e., if $S_i < \xi_i$, then a potential exiter will choose internal exit.[27]

Consequently, to prevent exit from his state, the ruler needs to ensure that each of the other elite members receives a share of the surplus that is at least equal to his certainty equivalent of internal exit, i.e., $S_i \geq \xi_i$. Accordingly, the sum of the surplus shares allocated to the other members of the elite must at least be equal to the sum of the other elite members' certainty equivalent, i.e.,

$$\sum_{i=2}^{E} S_i = (1 - \sigma_1)\mathrm{S}^* \geq \sum_{i=2}^{E} \xi_i. \quad (12)$$

Whether the ruler is able to provide every member of the elite with his certainty equivalent depends on three factors: the ruler's share of the surplus, σ_1; the size of the elite, E; and the surplus extracted, S. Assume that the ruler sets his own surplus share at the minimum that he is willing to accept in return for ruling the state, σ_1^*.

Two factors remain: the size of the elite and the surplus. Consider the size of the elite first. Since, historically, elite status was a birthright, the ruler would find it difficult to constrain the size of the elite. Consequently, as the size of the elite grows, the ruler will reach a threshold after which he can no longer ensure that each elite member receives a surplus share that corresponds to the certainty equivalent of exit. Thus, when,

$$(1 - \sigma_1{}^*)\mathrm{S}^* < \sum_{i=2}^{E} \xi_i, \quad (13)$$

at least one member of the elite will no longer receive a surplus share equal to his certainty equivalent. The rational response of the elite member who does not receives a surplus share below his certainty equivalent is to exit and form a new natural state. Thus, in terms of surplus shares, equation (13) provides a boundary for the African state.

[27] A potential exiter's risk preferences will inform the certainty equivalent. For example, a risk-averse potential exiter, the certainty equivalent can be less than the expected payoff from exiting, i.e., $\xi_i < E(internal\ exit)$. Thus, a risk-averse potential exiter can be given surplus shares less than his expected payoff of internal exit. Similarly a risk-neutral exiter requires a surplus share equal to their expected payoff of internal exit, while a risk-seeking exiter requires a surplus share greater than his expected payoff of internal exit.

A precolonial southern African ruler needed to find ways to constrain the size of the elite if he intended to prevent internal exit. One way to constrain the size of the southern African elite that is documented in the history literature is for the ruler to have his male relatives, including his sons and brothers, killed. Mzilikazi is thought to have had two of his sons killed (Thomas 1970[1873]: 227; Cobbing 1974:626; Ndlovu-Gatsheni 2008: 77). Zulu rulers, including Shaka and Dingane, had their sons killed at birth. Dingane also killed his brothers (Schapera 1963[1956]:173). However, the ruler needed to have some restraint as too much violence would cause him to lose his own followers.

The last factor is the surplus. To grow the surplus, the ruler would need to increase the number people over whom he can broadcast his power. To do so he will have to delegate power to elite members who can help him reach more people. But this strategy can backfire as it empowers the other elite members to lead people in their own right. Thus, the delegation of power ultimately can lead to exit.

2.4 Related Literature

It is useful to consider how the model and results of this essay relate to Buchanan and Faith's (1987) model of internal exit, Kopytoff's (1989) frontier social order, and North, Wallis and Weingast's (2009) framework of a state as an organisation of organisations. Consider each in turn.

The original Buchanan and Faith (1987) internal exit model presents three main differences with the model presented here. The first regards the issue of land. Buchanan and Faith envision a scenario in which a community declares independence and decides to provide its own public goods. The community remains on their existing land, simply changing political institutions. A recent example of an attempt at this type of internal exit is Scotland's bid for independence from Great Britain in 2014. African internal exit instead

shows how groups of people will physically exit from an existing state to move to a new land where they will go on to establish a new state.

The second difference is the initiator of internal exit. The Buchanan and Faith model suggests that the initiator will be a tax paying citizen who is essentially a commoner. The commoner's motive for exit is to avoid an exploitative tax rate. The African internal exit initiator is a surplus-sharing elite member who takes his own people with him to start a new state. This elite member is motivated to start his own state in which he will be the ruler and will be able to extract a surplus greater than the surplus share he received in the original state.

Thirdly, the models differ in predictions. The Buchanan and Faith model predicts that internal exit will not happen. Because the exiter is a commoner, the state can prevent exit by setting a non-exploitative tax rate.[28] However, in Africa the situation is reversed. A ruler facing conditions of land abundance, decreasing returns to the broadcast of power, a scarcity of people, and the elite institution of birthright was not always capable to prevent an elite member from exiting. As this model shows, there are three factors that determine whether or not a ruler is able to set an internal exit-proof surplus shares: the surplus extracted, the ruler's share, and the size of the elite.

This model offers an economic explanation for Kopytoff's (1989) "frontier social order." Kopytoff (p.16-17) identifies eleven features of the creation of new social orders on the frontier: (1) "a social dynamic" that produced exiters, (2) group exits, (3) "the institutional vacuum," (4) replication of existing states, (5) relatives as part of the group, (6) non-relative, non-elite subjects, (7) a distinction between "firstcomers" and "latecomers," (8) "patrimonialism," (9) inter-dependence between the ruler and his people, (10) "regional context," and (11) "the frontier as a historical process."

[28] Buchanan and Faith (1987) also extend the model to allow commoners who are likely to exit to become surplus-sharing members of the elite in order to remove the incentive to exit. As shown, in the African context this is not possible due to institutional constraints.

Consider each of these features in comparison to this model. (1) Rather than exiters being a social construction, this model suggests that the decision to exit was an economic one determined by the expected net increase in the surplus share due to exiting. (2) In this model, the reason for exit in groups is that the scarcity of valuable people meant that any prospective ruler would take as many people as possible with them. (3), (4), and (10) While the institutional vacuum of the frontier might have allowed for the construction of any desirable natural state, the examples of exit reported in this essay are, for the most part, more likely to be examples of the replication of existing states.[29] Kopytoff attributes replication to exiters requiring social acceptance among other states. Another likely economic reason for replication is that new state faced the same unique African conditions (i.e., abundant land, scarcity of people, decreasing returns to the broadcast of power and the institution of birthright) as the original state.

Features (5), (6) and (7) provide a more comprehensive social structure of African states than in this model. This model simply distinguishes between the elite and non-elite to show that the elite receive surplus shares and are entitled to rule in their own right. (8) and (9) provide the social standing of the ruler. Again this is simplified in economic terms. The model presented in this chapter shows that the ruler sets tax rates and surplus shares (11) And, lastly, while Kopytoff suggests that the exit and creation of new states is a social/historical process, this essay instead emphasizes that the creation of new states through internal exit is a result of elite members making rational economic calculations regarding their surplus shares.

The creation of new natural states on the African frontier represents a missing link in the North, Wallis and Weingast (2009) framework of a state as an organisation of organisations. In this framework, the predicted path for state development is from a natural state to an open

[29] Cf. Winter and Szulanski (2001)

access order. To transition from a natural state to an open-access order, North, Wallis and Weingast suggest that three "doorstop conditions" need to be in place: a rule of law for elite members, perpetual life for the state, and consolidated control of the military (p. 180-208). However, African natural states do not meet any of the doorstop conditions.

Firstly, rather than an impersonal rule of law among elite members, in African states relationships between elite members are personal, based on familial ties. An elite member may receive preferential treatment when favoured by the ruler. But the ruler's preferences can easily change and, when they do, he can have an elite member killed. As Schapera (1963 [1956]: 174-175) notes, relationships were fickle, and there was no rule of law that will provide an "adequate safeguard against bloodshed" among members of the elite.

Second, while some states do meet the condition of perpetual life, many do not. The perpetual life of the state is essential to develop impersonal "perpetually lived organisations" (p. 151). These types of organisations include business corporations as well as "political, municipal, educational, fraternal, and religious corporations" (p. 152). In African states the institution of birthright makes the perpetual life of the state unattainable. As birthright entitles all elite members to rule, when a ruler dies, many African natural states are bogged down in succession arguments between elite members that can destroy the state. For example, while Shaka was a member of the elite, he was not first in line to rule. However he still managed to use force to take over the state from his half-brothers and become the Zulu ruler (Krige 1965; Chanaiwa 1980). As a result of the institution of birthright, many southern African natural states did not survive beyond the third or fourth generation of succession (Richards 1961).

Lastly, consolidated military power requires "the existence of an organization with control over all the military resources" (p. 153) of the state. In the African context, in the same way that the ruler delegates the power rule to ward leaders, he also typically delegates military

power to the same leaders. Hence military power is diffused among the elite members. Therefore many African natural states do not meet the doorstop condition of a consolidated military. The Zulu under Dingane and Zwide (ca. 1970-1818) and later Shaka are an exception though. These rulers consolidated military power. Instead of organising the military along personal or familial lines, the Zulu army was organised into age groups led by experienced generals (Chanaiwa 1980). Since the majority of African states do not meet any of the doorstop conditions, and the Zulu state only meets the third one, African natural states are unable to progress to open access orders.

Yet state transitions did take place. As this essay shows, rather than transitioning between different types of orders, new natural states formed from existing natural states. Hence, the North, Wallis and Weingast framework can be extended, based on the African experience, to show how new natural states from existing natural states, in particular through the mechanism of internal exit.

2.5 Conclusion

This essay is a first attempt to build a historically grounded model in which internal exit is a state formation mechanism. In doing so, it contributes to the economic history of Africa by providing an economic rationale for the specific experiences of internal exit among precolonial southern African States (ca. 1500-1910) such as the Ndebele and the Tawana. More precisely, this essay suggests that the internal exit in precolonial southern African originates from the surplus share maximising behaviour of members of the elite.

Additionally, this essay contributes to economic theory in three ways. Firstly, this essay extends the theory of internal exit by exploring the consequences of internal exit when the initiator is a member of the elite and the exit involved the physical movement of people. Secondly, this essay provides an economic model of Kopytoff's (1989) frontier social order.

Thirdly, this essay extends the North, Wallis and Weingast (2009) framework for a state as an organisation of organisations to include the formation of new natural states from existing ones.

While this chapter sheds new light on the phenomenon of internal exit and state formation in Africa, it is not without its limitations. One such limitation is that this model does not account for an exit forced by a ruler. The economic rationale for a ruler to force exit and actually lose people is beyond the scope of this essay. Thus an area for future research is one that considers other mechanisms of state formation, like forced exit in southern Africa, during both the precolonial and early colonial period.

2.6 References

Acemoglu Daron, Simon Johnson and James A. Robinson 2002. "Reversal of fortune: geography and institutions in the making of the modern world income distribution," *Quarterly Journal of Economics* 117(4): 1231–1294 (November).

Alesina, Alberto and Enrico Spolaore 2005. *The Size of Nations* Cambridge, Massachusetts: MIT Press.

Anderson, Terry L. and Peter J. Hill 2004. *The Not So Wild, Wild West: Property Rights on the Frontier* Stanford, California: Stanford University Press.

Ashraf, Nava, Natalie Bau, Nathan Nunn and Allesandra Voena 2015. *Bride Price and the Returns to Education for Women* Working Paper Available online: http://scholar.harvard.edu/nunn/publications/bride-price-and-returns-education-women Last Accessed August 10, 2015.

Asiwaju, A.I. 1976. "Migrations as Revolt: The Example of the Ivory Coast and the Upper Volta before 1945," *Journal of African History* 17:4: 577-594(October).

Austin, Gareth 2008. "Resources, Techniques, and Strategies South of the Sahara: Revising the Factor Endowments Perspective on African Economic Development, 1500-2000," *Economic History Review* 61(3): 587-624 (August).

Ballard, Charles 1983[1981]. "Trade, Tribute and Migrant Labour: Zulu and Colonial Exploitation of the Delagoa Bay Hinterland 1818-1879," Chapter 4 in J.B. Peires (Ed.) *Before and After Shaka: Papers in Nguni History* Grahamstown, South Africa: Institute for Social and Economic Research, Rhodes University.

Barfield, Thomas J. 1993. *The Nomadic Alternative* New Jersey: Prentice Hall

Bates, Robert, Avner Greif, and Smita Singh 2002. "Organizing Violence," *Journal of Conflict Resolution* 46(5): 599-628(October).

Becker, Peter 1966. *Path Of Blood: A Tragic Era of Bloody Warfare that Culminated in Thousands of Starved Wanderers being reduced to Cannibalism* London, Great Britain: Panther Books.

Berkowitz, Daniel 1997. "Regional Income and Secession: Center-periphery Relations in Emerging Market Economies," *Regional Science and Urban Economics* 27(1):17-45(February).

Bolton, Patrick and Gerard Roland 1997. "The Breakup of Nations: A Political Economy Analysis," *Quarterly Journal of Economics* 112(4):1057-1090 (November).

Brown, Richard 1969. "The External Relations of the Ndebele Kingdom in the Prepartition Era," Chapter 12 in Leonard Thompson (Ed.) *African Societies in southern Africa: Historical Studies* London, Great Britain: Heinemann.

Bryant, Alfred T. 1905. *A Zulu-English Dictionary with Notes on Pronunciation, a Revised Orthography and Derivations and Cognate Words from Many Languages; Including also*

a Vocabulary of Hlonipa Words, Tribal-Names, etc., a Synopsis of Zulu Grammar and a Concise History of the Zulu People from the Most Ancient Times Maritzburg and Durban: P. Davis and Sons, Cape Town, Johannesburg and Durban: J.C. Juta and Co.; Pinetown, South Africa: Marianhill Mission Press. Available online: https://archive.org/details/zuluenglishdicti00brya Last Accessed April 18, 2015.

Buchanan, James M. and Roger L. Faith 1987. "Secession and the Limits of Taxation: Toward a Theory of Internal Exit," *American Economic Review* 77(5): 1023-1031 (December).

Central Intelligence Agency (CIA) 2015a. "South Africa" *The World Factbook* Available online: https://www.cia.gov/library/publications/the-world-factbook/geos/sf.html Last Accessed June 10, 2015.

Central Intelligence Agency (CIA) 2015b. "Zimbabwe" *The World Factbook* Available online: https://www.cia.gov/library/publications/the-world-factbook/geos/zi.html Last Accessed June 10, 2015.

Chanaiwa, David 1976. "The Army and Politics in Pre-Industrial Africa: The Ndebele Nation," *African Studies Review* 19(2): 49-67(September).

Chanaiwa, David Shingirai 1980. "The Zulu Revolution: State Formation in a Pastoralist Society," *African Studies Review* 23(3): 1-20 (December).

Coase, R.H. 1937. "The Nature of the Firm," *Economica* 4(16):386-405(November).

Coase, R.H. 1960. "The Problem of Social Cost," *Journal of Law and Economics* 3:1-44(October).

Cobbing, Julian 1974. "The Evolution of the Ndebele Amabatho," *Journal of African History* 15(4): 607-631.

Demsetz, Harold 1967. "Toward a Theory of Property Rights," *American Economic Review* 57(2): 347-359(May).

Eggertsson, Thráinn 1991. *Economic Behavior and Institutions* Cambridge, Great Britain: Cambridge University Press.

Fenske, James 2012. "Land Abundance and Economic Institutions: Egba Land and Slavery 1830-1914," *Economic History Review* 65(2): 527-555(May).

Findlay, Ronald 1996. "Towards a Model of Territorial Expansion and the Limits of Empire," Chapter 3 in Michelle R. Garfinkel and Stergios Skaperkas (Eds.) *The Political Economy of Conflict and Appropriation* New York: Cambridge University Press.

Friedman, David 1977. "A Theory of the Size and Shape of Nations," *Journal of Political Economy* 85(1): 59-77(February).

Gulbrandsen, Ørnulf 1993. "The Rise of the North-Western Tswana Kingdoms: On the Dynamics of Interaction between Internal Relations and External Forces," *Africa: Journal of the International African Institute* 63(4): 550-582.

Guy, J.J. 1983[1981]. "Production and Exchange in the Zulu Kingdom," Chapter 1b in J.B. Peires, (Ed.) *Before and After Shaka: Papers in Nguni History* Grahamstown, South Africa: Institute for Social and Economic Research, Rhodes University.

Hall, Martin 1986. "The Role of Cattle in Southern African Agropastoral Societies: More than Bones Alone Can Tell," *Goodwin Series: Prehistoric Pastoralism in Southern Africa* 5: 83-87(June).

Hammond-Tooke, David 1964. "Chieftainship in Transkeian Political Development," *Journal of Modern African Studies* 2(4): 513-529 (December).

Hammond-Tooke, W.D. 1985. "Descent Groups, Chiefdoms and South African Historiography," *Journal of Southern African Studies* 11(2): 305-319 (April).

Harinck, Gerrit 1969. "Interaction between the Xhosa and Khoi: Emphasis on the Period 1620-1750," Chapter 7 in Leonard Thompson (Ed.) *African Societies in Southern Africa*: 145-170 London, Great Britain: Heinemann.

Herbst, Jeffrey 1990. "Migration, the Politics of Protest, and State Consolidation in Africa," *African Affairs* 89(355): 183-203(April).

Herbst, Jeffrey 2000. *States and Power in Africa: Comparative Lessons in Authority and Control* Princeton, New Jersey: Princeton University Press.

Hirschman, Albert O. 1970. *Exit, Voice and Loyalty: Responses to Decline in Firms, Organizations, and States* Cambridge, Massachusetts: Harvard University Press.

Hopkins, A.G 2009. "The New Economic History of Africa," *Journal of African History* 50(2): 155-177 (July).

Hoernlé, A. Winifred 1962[1937]. "Social Organization," Chapter IV in Schapera, I. (Ed.) *The Bantu-Speaking Tribes of South Africa: An Ethnological Survey* London, Great Britain: Routledge and Kegan Paul.

"How Much You Pay for South Africa's Royal Families" 2015 (June 1) *BusinessTech* Available online: http://businesstech.co.za/news/government/88860/how-much-you-pay-for-south-africas-royal-families/ Last Accessed October 30, 2015.

Hunt, D.R. 1931. "An Account of the Bapedi," *Bantu Studies* V: 275-326.

Iliffe, John 2013[1995, 2007]. *Africans: The History of a Continent* 2nd Edition New York: Cambridge University Press.

Klein, Peter G, Joseph T Mahoney, Anita M McGahan, Christos N Pitelis 2010. "Toward a Theory of Public Entrepreneurship" *European Management Review* (2010) 7(1): 1–15(Spring).

Knight, Frank. H. 2006 [1921]. *Risk, Uncertainty and Profit* Mineola, New York: Dover Publications.

Kopytoff, Igor 1989. "The Internal African Frontier: The Making of African Political Culture," Chapter 1 in Igor Kopytoff (Ed.) *The African Frontier: The Reproduction of*

Traditional African Societies Bloomington and Indianapolis, Indiana: Indiana University Press.

Kopytoff, Igor and Suzanne Miers 1977. "African 'Slavery' as an Institution of Marginality" Chapter 1 in Suzanne Miers and Igor Kopytoff (Eds) *Slavery in Africa* Madison, Wisconsin: University of Wisconsin Press.

Krige, J.D. 1937. "Traditional Origins and Tribal Relationships of the Sotho of the Northern Transvaal," *Bantu Studies* 11: 321-356.

Kuper, Adam 1975a. "The Social Structure of the Sotho-Speaking Peoples of southern Africa. Part I," *Journal of the International African Institute* 45(1): 67-81.

Kuper, Adam 1975b. "The Social Structure of the Sotho-Speaking Peoples of southern Africa. Part II," *Journal of the International African Institute* 45(2): 139-149.

Leeson, Peter 2005. "Endogenizing Fractionalization," *Journal of Institutional Economics* 1(1): 75-98 (June).

Leeson, Peter T., Peter J. Boettke, and Jayme S. Lemke 2014. "Wife Sales" *Review of Behavioral Economics* 1(4): 349-379.

Legassick, Martin 1969. "The Sotho-Tswana Peoples before 1800," Chapter 5 in Leonard Thompson (Ed.) *African Societies in Southern Africa: Historical Studies* London: Heinemann.

Lestrade, G.P 1930. "Some Notes on the Political Organization of the Venda-speaking Tribes," *Africa: Journal of the International African Institute* 3(3): 306-322 (July).

Lye, William 1969a. "The distribution of the Sotho peoples after the Difiqane," Chapter 9 in Leonard Thompson (Ed.) *African Societies in southern Africa: Historical Studies* London, Great Britain: Heinemann (pages 190-206)

Lye, William F. 1969b. "The Ndebele Kingdom South of the Limpopo River," *Journal of African History* 10(1): 87-104.

Mbenga, Bernard and Andrew Manson 2010. *'People of the Dew' A History of the Bafokeng of Phokeng-Rustenburg Region, South Africa, from Early Times to 2000* Johannesburg, South Africa: Jacana.

McGuire, Martin C. and Mancur Olson Jr. 1996. "The Economics of Autocracy and Majority Rule: The Invisible Hand and the Use of Force," *Journal of Economic Literature* 34(1):72-96(March).

Mönnig, H.O. 1988. *The Pedi* Pretoria, South Africa: Van Schaik.

Ndlovu-Gatsheni, Sabelo J. 2008. "Who Ruled by the Spear? Rethinking the Form of Governance in the Ndebele State," *African Studies Quarterly* 10(2-3): 71-94(Fall).

North, Douglass C., John Joseph Wallis and Barry R. Weingast 2009. *Violence and Social Orders* New York: Cambridge University Press.

Nunn, Nathan 2008. "The Long Term Effects of Africa's Slave Trades," *Quarterly Journal of Economics* 123(1): 139-176 (February).

Olson, Mancur 1993. "Dictatorship, Democracy, and Development," *American Political Science Review* 87(3): 567-576(September).

Omer-Cooper, John 1969. "Aspects of Political Change in the Nineteenth-Century Mfecane," Chapter 10 in Leonard Thompson (Ed.) *African Societies in Southern Africa: Historical Studies* London, Great Britain: Heinemann (pages 207-229)

Omer-Cooper, J.D. 1994. *History of Southern Africa* Second Edition London, Great Britain: James Currey.

Osafo-Kwaako, Philip and James Robinson 2013. "Political Centralization in Pre-Colonial Africa," *Journal of Comparative Economics* 41(1):6-21 (February).

Peires, J.B. 1983[1981]. "Chiefs and Commoners in Precolonial Xhosa Society," Chapter 5 in J.B. Peires (Ed.) *Before and After Shaka: Papers in Nguni History* Grahamstown, South Africa: Institute for Social and Economic Research, Rhodes University.

Posner, Richard A. 1979. "The Homeric Version of the Minimal State," *Ethics* 90(1): 27-46(October).

Rasmussen, R. Kent 1977. "A Lost Man in Southern African History: Kaliphi/Gundwane of the Ndebele," *International Journal of African Historical Studies* 10(1): 96-110.

Richards, Audrey I. 1961. "African Kings and Their Royal Relatives," *Journal of the Royal Anthropological Institute of Great Britain and Ireland* 91(2): 135-150 (July-December).

Sansom, Basil 1974[1937]. "Traditional Rulers and their Realms," Chapter 8 in W.D Hammond-Tooke (Ed.) *The Bantu-Speaking Peoples of Southern Africa 2nd Edition*: 135-176 London, Boston and Henley: Routledge & Kegan Paul.

Schapera, I 1937. "Political Institutions," Chapter VIII in I. Schapera (Ed.) *The Bantu-Speaking Tribes of South Africa* London: George Routledge & Sons; Cape Town, South Africa: Maskew Miller.

Schapera, I. 1963[1956]. *Government and Politics in Tribal Society* London, Great Britain: Watts.

Schapera, I 1963. "Kinship and Politics in Tswana History," *Journal of the Royal Anthropological Institute of Great Britain and Ireland* 93(2): 159-173 (July-December).

Schapera, I 1968. "Early Migrations of the Southern Bantu," Chapter 19 in P.J.M. McEwan (Ed.) *Africa from Early Times to 1800*: 205-209 London, Great Britain: Oxford University Press.

Schapera, I 1973[1953]. *The Tswana* London, Great Britain: International African Institute.

Schapera, I 1980. "Notes on the early history of the Kwena (Bakwena-bagaSechele)," *Botswana Notes and Records* 12: 83-87.

Thomas, Thomas Morgan 1970[1873]. *Eleven Years in Central South Africa*: Bulawayo, Zimbabwe: Books of Rhodesia.

Thompson, Leonard 1969. "Co-operation and Conflict: The Zulu Kingdom and Natal," Chapter VIII in Monica Wilson and Leonard Thompson (Eds.) *The Oxford History of South Africa I: South Africa to 1870.* London, Great Britain: Oxford.

Thompson, Leonard 1995[1978] "Southern Africa to 1795," Chapter 9 in Phillip Curtin, Steven Feierman, Leonard Thompson and Jan Vansina (Eds.) *African History* London, Great Britain: Longman.

Tiebout, Charles M. 1956. "A Pure Theory of Local Expenditures," *Journal of Political Economy* 64(5): 416-424 (October).

Tlou, Thomas 1977. "Servility and Political Control: Botlhanka among the BaTawana of Northwestern Botswana, ca. 1750-1906," Chapter 14 in Suzanne Miers and Igor Kopytoff (Eds) *Slavery in Africa* Madison, Wisconsin: University of Wisconsin Press.

Transvaal Native Affairs Department (TNAD) 1905. *Short History of the Native Tribes of the Transvaal*.

Uzoigwe, G.N. 1977. "The Warrior and the State in Precolonial Africa," *Journal of Asian and African Studies* 12(1-4): 20-47(January).

Vansina, J. 1962. "A Comparison of African Kingdoms," *Africa: Journal of the International Africa Institute* 32(4): 324-335(October).

Weber, Max 1946. "Politics as a Vocation" in H.H. Gerth and C. Wright Mills (Eds.) *From Max Weber: Chapters in Sociology* New York: Oxford University Press Available online: https://openlibrary.org/books/OL6498314M/From_Max_Weber_Chapters_in_sociology Last Accessed March 19, 2014.

Wilson, Monica 1969. "The Sotho, Venda, and Tsonga," Chapter IV in Monica Wilson and Leonard Thompson (Eds.) *Oxford History of South Africa I: South Africa to 1870.* London, Great Britain: Oxford.

Winter Sidney G. and Gabriel Szulanski 2001. "Replication as Strategy," *Organization Science* 12(6): 730-743 (November-December).

Wittman, Donald 1991. "Nations and States: Mergers and Acquisitions; Dissolutions and Divorce," *American Economic Review* 81(2):126-129(May).

Chapter 3

Intergovernmental Grants and Local Government Learning

3.1 Introduction

In his 2004 Presidential Address to the International Schumpeter Society Meeting, Malerba (2006) identified four challenges for the research programme of evolutionary economics: demand, knowledge, networks, and coevolution. The purpose of this chapter is to pick up the two challenges of knowledge and networks. However, rather than considering the two challenges in the context of industrial economics, as Malerba mainly suggests doing, this essay considers the challenges in the context of public economics as learning in a fiscal federation.[30] In fact, even though Schumpeter is considered also a pioneer for the study of public economics, especially in its non-constitutional Public Choice manifestation (Buchanan 1984), the premise of this essay is that modern economic evolutionary theory has so far paid little attention to the potential of a dynamic analysis of the public economy.

In public economics, one way to view federalism is as a laboratory in which local governments help the central government by running experiments simultaneously to solve policy problems (e.g., Oates 1999: 1131-4).[31] The 'collective intelligence' of the federated network means that local governments are able to learn from each other's policy mistakes and successes and, ultimately, will generate positive externalities in the form of knowledge

[30] Boschken (1982) is the first to link federations to networks through organisation theory. For more recent contributions see Garzarelli (2004) and Oates (2011).

[31] All states are fiscally decentralised to some degree: even if a state is de jure unitary, it is de facto decentralised (e.g., Breton 2000). Immediate concrete (democratic) examples include France, Italy, South Africa, and the United Kingdom. Cai and Treisman (2009) identify China and the USSR as examples in which dictatorships/centrally planned economies manifested decentralisation. This essay is framed in terms of fiscal federalism as this this is the sphere in which the literature mostly addresses the subject matter I am concerned with. However, the analysis is equally valid for unitary states with de factor decentralisation. For this same reason, some of the decentralisation examples cited are from states that are de jure unitary.

growth for the federation as a whole (e.g., Sabel 2004). Local policy experiments that prove unsuccessful will have a low cost to the federation as a whole, as the experiments are only performed on a local level; whereas those experiments that prove successful will have a high return as they can be adopted by other governments in the federation (e.g., Kollman, Miller and Page 2000).

Casual empiricism suggests that local governments routinely experiment with policy, achieving varying degrees of success. One example is Mayor Bloomberg's range of anti-poverty experiments in New York City (Bosman 2009). Another example is that of Mayor Mockus of Bogotá who introduced many innovative policy experiments. In one experiment, Mayor Mockus employed mimes to direct traffic and to alleviate traffic congestion (Caballero 2004).[32]

This chapter considers laboratory federalism in conjunction with intergovernmental grants, the primary fiscal policy tool of decentralisation. The Public Economics literature on intergovernmental grants is extensive (among others, Courant, Gramlich and Rubinfeld 1979; King 1984; Oulasvirta 1997; Mieszkowski and Musgrave 1999). In this extensive literature, grants are usually analysed according to consumer behaviour theory. These analyses assume that grants are spent either according to the preferences of the median voter (e.g., Romer and Rosenthal 1980) or according to the preferences of the community as a whole (e.g., Oates 1972).[33] Under these analyses, income and substitution effects determine community spending and, ultimately, community welfare.[34]

[32] US states often implement innovative policies. For example, Pennsylvania is known for leading policy innovation regarding organ donation (Stolberg 1999). Oregon has experimented different voting systems (Bradbury 2005). Numerous states experiment with different forms of corporate law (Romano 2006), and Colorado and Washington State are leading experiments in legalising cannabis (Kleiman and Ziskind 2014).

[33] Another view on grants considers the preferences of only of the grant recipient (Wilde 1968). This view does not take into account the preferences of the community or the preferences of the grant donor. Brennan and Pincus (1990:130) provide an example of a university board that failed to take the preferences of the grant donor into account.

[34] There are many shortcomings to analysing grants in this way. For example, the consumer behaviour analysis of intergovernmental grants is unable to distinguish between an unconditional lump-sum grant and a conditional lump-sum grant (Oates 1972: 77).

At the same time, very little literature exists tying policy experimentation with intergovernmental grants. This is somewhat surprising since intergovernmental grants are routinely used to stimulate policy experimentation. In one such example, when the Public Safety Department (which consists of both fire and police) of Fairfield, California was allowed to keep savings from a lump-sum grant, the Department changed its strategy from only fighting fires to fire prevention. This change in policy lowered the running costs as well as improved the insurance rating of the department (Osborne and Gaebler 1993: 225). Grants were used for the State Children's Health Insurance Program (SCHIP) in the United States. SCHIP was introduced in 1997 to improve medical insurance coverage for children. SCHIP allocated lump-sum grants to states, which enabled them to design and apply the policy in different ways (Weissert and Scheller 2008). Another example is the Patient Protection and Affordable Care Act (ACA) from the United States. The ACA uses federal support in the form of intergovernmental grants to create incentives for states to innovate. The results of the policy experiments are evaluated, data is collected and the information is made publicly available (Madison 2014: 788-789).[35] Other countries also use grant funding for policy experimentation. Public universities in Scotland used lump-sum grant funding to successfully experiment with different fees payment policies (Keating 2005a,b).

Intergovernmental grants have also been used successfully to promote policy diffusion (i.e., the transfer of successful policies between governments) in the US. The federal government uses grants either by allocating grants as a reward to those states that implement newly developed policies or by withdrawing grants to punish those states that do not implement the policies (Welch and Thompson 1980; Allen, Pettus and Haider-Markel 2004).

[35] Policy experimentation is not limited to local governments. Kerber and Eckardt (2007) consider policy learning through from countries in the European Union. Cai and Treisman's (2009) model of laboratory federalism assumes that both central and local governments can experiment with policy. Historically, former US President Franklin D. Roosevelt encouraged policy experiments at different levels of government during the Great Depression (Rodrik 2014: 204). Currently Great Britain's Prime Minister David Cameron and US President Barack Obama both encourage policy experimentation in terms of behavioural interventions at the central level (Bennhold 2013; Social Sciences and Behavioral Team 2015).

Hence, it seems worthwhile to consider how intergovernmental grants can help local governments to perform policy experiments and ultimately learn from them. Through a simple model, this essay shows that a local government grant recipient (e.g., a city manager, an Italian regional president, a mayor, a South African provincial premier, a US state governor) is likely to experiment more (and also learn more) from a lump-sum grant than a closed matching grant because an unconditional lump-sum grant allows greater leeway for spending than a closed matching grant.[36] This means that grant donors can use different types of grants as policy instruments to induce different levels of experimentation at the recipient government level.[37] An additional implication of the model is that an unconditional lump-sum grant can produce greater external benefits in the form of knowledge gains to other local governments (and even to the central government) than a closed matching grant. [38]

The model proxies the undertaking of different policy experiments by the local government representative using an intergovernmental grant with *melioration* (Herrnstein and Prelec 1991; Herrnstein et al. 1993; Metcalfe 2001). In its simplest sense, melioration is the process of making something better. In the broad evolutionary economics interpretation, melioration is a process in which an economic agent continually allocates resources (e.g., time and money) to the best feasible alternative. Through experimentation with an intergovernmental grant, the local government representative, behaving as a meliorator, engages in different policy experiments, ultimately allocating more of the grant funding to the

[36] Intergovernmental grants can be given vertically, i.e., from higher level governments to lower ones (or vice versa) or horizontally, i.e., between governments on the same level. Typically (and in this essay) vertical intergovernmental grants are transferred from a higher level of government to a lower one. For a model of the reverse, see Boadway and Keen (1996).

[37] The problem considered here is analogous to the positive version of the traditional soft budget contstraint problem discussed by Kornai (1986), Oates (2005), Vigneault (2007), Besfamille and Lockwood (2008), and Weingast (2009), among others. As Osbourne and Gaebler (1993: 135-137) notes, "permission to fail" is critical to innovation in the public sector.

[38] Bednar (2011: 215) also suggests that a federation needs to have sufficient rules in place to tolerate "selfish experimentation." For example a local government that experiments with legalising cannabis successfully (e.g., in terms of economic growth, higher public revenue and lower crime rates) may actually produce negative externalities to those states in which the median voter finds drug usage repugnant. This type of externality is also known as a "transcendental externality" (Langlois 1982: 295). Clearly, the knowledge gains from "selfish experimentation" will be more limited than experimentation on policies that are more likely to be valued throughout the federation (e.g., welfare policies).

policy experiment that proves most successful. Moreover, as a meliorator, the local government representative will also put a stop to any policy experiments that prove unsuccessful. Melioration accordingly allows us to capture trial-and-error experimentation.

3.2 Intergovernmental Grants as Policy Instruments for Experimentation

The laboratory federalism literature originates from Public Economics and Law and Economics (e.g., Rose-Ackerman 1980; Oates 1999; Kollman, Miller and Page 2000; Strumpf 2002; Sabel 2004; Kotsogiannis and Schwager 2006; Cai and Treisman 2009; Galle and Leahy 2009; Schnyder 2011). It examines two issues: firstly, whether decentralised governments produce more policy experimentation than centralised governments and, secondly, the incentives needed to promote experimentation in a decentralised government. This essay is concerned with the second issue, i.e. with creating the incentives to promote experimentation (see, e.g., Rose-Ackerman 1980, Strumpf 1998, Garzarelli 2006, Bednar 2011 and Madison 2014).

Policy experimentation implies a degree of uncertainty. The local government has to implement policy choices which may or may not work. Strumpf (2002) suggests that one reason for a local government not to engage in experimentation is risk aversion. But when the intention of an intergovernmental grant is to promote experimentation, grants can be used to cushion unsuccessful experimentation. By rewarding experimentation, regardless of outcome, a local government grant recipient will be able to experiment with policy without fear of losing future grants.

But, as always, there is no free lunch: with any type of grant comes the implication that future grants will be related to how effectively the previous grant is used (Brennan and Pincus 1990: 129; Garzarelli 2006: 248). The rational grant recipient, therefore, has an

incentive to use the funds efficiently or else run the risk of not receiving additional grants in the future.[39] Hence, we can view an intergovernmental grant as an incomplete contract between the grant donor (the central government) and the grant recipient (the local government) (cf. Brennan and Pincus 1990; Breton 1998[1996]).[40] Taking this view suggests that an intergovernmental grant is more than simply a transfer payment from the donor to the recipient.[41] Instead, a grant is able to convey information that the price mechanism fails to transmit. Thus, like any contract, a grant is actually an instrument of exchange between the donor and the recipient.

Traditionally, there are two main categories of intergovernmental grants: conditional and unconditional (e.g., King 1984). For example, a closed matching grant is considered a conditional grant while a lump-sum grant is considered an unconditional grant. For conditional grants, the donor places restrictions on the use of the funds; while for unconditional grants, the donor does not place any restrictions on the use of the funds. A closed matching grant is classified as a conditional grant which specifies that for each dollar spent by the grant recipient on a particular activity, the grant donor will match it with a specified sum to a maximum sum of money (e.g., for every one dollar spent by the recipient, the donor will contribute fifty cents but to a maximum of two million dollars).[42] A lump-sum grant is classified as an unconditional grant. Unconditional lump-sum grants are sums of money given with no restriction on the use of funds (e.g., two million dollars).

---

[39] In the Brennan and Pincus (1990:130) example, the grant recipient (a public university) was punished by the grant donor (the government) with a cut in grant funding. Another example of implicit conditions attached to "unconditional" grants comes from South Africa: during an electricity crisis in 2015, the South African National Treasury withheld fiscal equalisation grants (which are typically unconditional lump-sum grants) from municipalities which had failed to pay their electricity bills (Slabbert 2015).

[40] An incomplete contract is a contract between two or more parties that contains gaps or missing stipulations (Hart 1988:123). An intergovernmental grant could not be viewed as a complete contract as "[a] complete contract would specify all relevant contingencies over the life of the contract, foresee a course of action for each of these contingencies, deny the…possibility of renegotiation so as not to deprive the original agreement of its credibility, and provide for each contingency" (Breton 1998[1996]:214).

[41] Other views of intergovernmental grants beyond the traditional consumer theory analysis also include Ostrom and Ostrom (1965) and Bradford and Oates (1971).

[42] Matching grants can be either closed-ended or open-ended. A closed-ended matching grant limits the total amount the donor will contribute; while the size of the open-ended matching grant is dependent only on the contribution made by the recipient.

However, when viewing an intergovernmental grant as an incomplete contract, any grant has conditions attached regardless of whether it is a conditional grant or not (Brennan and Pincus 1990; Breton 1998[1996]; Garzarelli 2006). It is simply the case that some of the grant conditions are explicit (e.g., a matching ratio and a maximum amount in the case of a closed matching grant) while others are implicit (e.g., verbal suggestions) (Brennan and Pincus 1990; Breton 1998[1996]). An unconditional lump-sum grant is only unconditional in so far as there are no explicit conditions attached, i.e., implicit conditions will still be attached.[43] A closed matching grant that is explicitly conditional is also likely to have implicit conditions attached. Hence, a grant donor can use both explicit conditions and implicit conditions to direct the spending of a grant.[44] Thus, all grants, whether explicitly conditional or not, are, in fact, conditional (Brennan and Pincus 1990). Hence, different types of intergovernmental grants can have different degrees of conditionality. For example, as a lump-sum grant only has implicit conditions attached, it is relatively less conditional than a closed matching grant, which has both explicit and implicit conditions attached.

The inverse of grant conditionality is the leeway or discretion left to the grant recipient: a grant with a high degree of conditionality will allow very little leeway, while a grant with a relatively low degree of conditionality will have greater leeway. Thus, an unconditional lump-sum grant has relatively more leeway as compared to a conditional closed matching grant, and vice versa.[45] Garzarelli (2006:249) suggests that this leeway can "serve as a policy tool. It can be used by a central government in a federation to induce different levels of

[43] As Breton (1987[1985]: 315) indicates, "[u]nconditionality…is the product of a reliance on a theory of intergovernmental grants in which the decision makers are neither real governments, nor competitive."

[44] It's unlikely that a grant donor will provide grant funds without any conditions at all because, like the recipient, the donor is also subject to political constraints (Brennan and Pincus 1990: 137).

[45] When experimentation is not required, a grant donor and a grant recipient often have preferences for grant spending. The donor or central government prefers to promote those policies that it views as important, e.g., those that generate positive externalities and merit goods. Thus, a donor will prefer a grant with explicit conditions attached, e.g., a matching grant. In contrast, the recipient or local government prefers to have leeway to spend the funds according to local preferences. Thus, a recipient will prefer a grant with as few explicit conditions as possible, e.g., an unconditional lump-sum grant (Oulasvirta 1997: 401).

experimentation from its local governments."[46] A grant that is less complete (an unconditional lump-sum grant) allows more leeway for the recipient to engage in policy experimentation than a grant that is relatively more complete (a closed matching grant).[47]

Moreover, a local government grant recipient who chooses not to engage in experimentation will be held accountable by *both* the donor and its citizens.[48] The donor can punish the grant recipient by increasing grants to other local governments. The citizens of the local community will recognize that they have lost out on these grants and will hold the local government grant recipient responsible (Goodspeed 2002: 418-419).[49]

For example, when Norway replaced conditional grants with unconditional lump-sum grants in 1986, Carlsen (1995:56) found that local government grant recipients had an incentive to "use spending decisions as strategic instruments to achieve additional grants." The reason for this incentive is that the Norwegian public believed that the behaviour of the local grant recipients could be the reason for future grant reductions. Thus, if fewer grants were received, the public would hold the local grant recipient responsible in the next local election.[50]

Even when there is imperfect information, there are incentives for both the grant donor and the grant recipient to uphold the conditions of a grant. The conditionality attached to a grant creates political rents for both the grant donor and the grant recipient.[51] Provided the

---

[46] Bednar (2011) also suggests grants to create incentives to experiment but does not discuss the relative conditionality issue raised by both Strumpf (1998) and Garzarelli (2006).

[47] The conditionality associated with grants to induce policy experimentation also means that intergovernmental grants effectively 'soften' the budget constraint of the local government. This is because the rational grant recipient understands that future grants can be expected once the current grant has been used for policy experimentation. Thus, the central government can use future intergovernmental grants to both encourage and reward experimentation. The positive incentives created by soft budget constraints have not been examined in the literature. Typically, a soft budget constraint is considered harmful to a federation as it creates perverse incentives for local governments to overspend (e.g., Kornai 1986; Oates 2005; Vigneault 2007; Weingast 2006).

[48] Such accountability will avoid problems of transfer dependence (Rodden 2003; Oates 2008).

[49] Volden (1999: 73) finds that grant recipients do respond to the incentives created by intergovernmental grants but the nature of response depends on institutions of recipient government.

[50] Additionally, local governments compete with each other to get grants from higher levels of government (Salmon 2005:3). Moreover, Callander and Harstad (2013) suggest that local governments compete with each other when it's likely that one of their policy innovations will be adopted by the central government.

[51] One way that local governments generate political rents for central governments is when the local governments "hire" the central governments to collect taxes. The central government has economies of scale in preventing both tax avoidance and

conditions attached to a grant generate sufficient rents for both the grant donor (the central government) and for the grant recipient (the local government), then neither the donor nor the recipient will wish to lose out on these rents. Hence, both the local government and the central government will have incentives to uphold the contractual terms of the grant and, in this case, to use the grant funding effectively to experiment with policy (Breton 1998[1996]: 214-219).

Now consider how a local government grant recipient, behaving as a meliorator, uses an intergovernmental grant to experiment with policies. Subsequently, this chapter analyses how different types of grants (e.g. a closed matching grant compared to a lump-sum grant) can engender different levels of experimentation and, hence, learning.

3.3 Local Government Learning

Creating new, successful policies is difficult (e.g., Lindblom 1959, 1979). Often, policy experimentation is needed because policy makers do not necessarily have the knowledge of which policy will produce a desired objective (Callander 2011b). For example, in the case of legalising cannabis in the US, Kleiman and Ziskind (2014: 78) note, "[g]iven the range of potential gains and losses, and of policy options, the probability of finding the perfect combination right from the start must surely be near zero. Thus, the best initial policy will not be the one that comes closest to some calculated optimum, but instead the one easiest to adjust in light of experience, which among other things means building in evaluation and policy feedback mechanisms."

Evolutionary theory is especially useful to study instances where agents are not explicitly optimising, such as when policy makers are attempting policy innovation (Nelson and Winter

tax evasion. Thus, the local government will be in a position to transfer political rents to the central government when the cost of collecting taxes locally (i.e., without economies of scale) is greater than the sum of the cost of collecting taxes centrally and the cost of creating the political rents (Breton 1998[1996]: 214-219).

1977). Hence, this essay uses tools of evolutionary theory to help to enrich the understanding of how local governments experiment with policy and how the experimentation leads to learning.[52]

Learning is not a perfect process. As Arrow (1962:155) indicates, "[l]earning is the product of experience. Learning can only take place through the attempt to solve a problem and therefore only takes place during activity." Because experimentation and innovation involve a degree of uncertainty, it is difficult for a policy maker to choose the correct policy straight away. This is especially true when the problem itself is not easy to formulate. Hence, learning from experimentation implies some degree of mistakes (see, e.g., Nelson and Winter 1977 and Callander 2011a).[53] In spite of the mistakes, though, the concept of learning does imply that the policy maker will make an improvement, over a period of time, from poorer policy choices to better choices.[54]

As an evolutionary process, learning requires three mechanisms, namely selection, variation and retention. Consider each of these mechanisms in turn. Selection is a "mechanism providing differential rewards and penalties" (Dosi, Marengo and Fagiolo 2005: 318). In the context of learning at the local government level, selection is the act of choosing the most appropriate policy from the available policy options (Forte 1982: 234).[55] Each policy option under consideration for experimentation should have an associated value or reward function attached to it that provides a guide to selecting the most appropriate policy

[52]Callander (2011b) and Callander and Hummel (2014) model experimentation and learning in the context of policies, though not necessarily at the local government level.

[53] Note that a good policy is considered one that achieves the objectives desired by society. However, a policy that is a mistake is not necessarily a policy that does not achieve the objectives desired by society. Instead, a policy that is a mistake means that there is another policy that would be preferred (Lindblom 1959:83).

[54] Cf. Börgers and Sarin (1997).

[55] There are two types of selection in the evolutionary literature, namely natural selection and artificial selection. Natural selection takes place without human intervention, while artificial selection occurs when humans manipulate the "criteria or environment of selection; the selection process is under the control of a human agent" (Hodgson and Knudsen 2010: Section 3.2 Paragraph 2). Following Hodgson and Knudsen (2010), this chapter does not separate out these two forms of selection because, in spite of human intervention, institutions are still exposed to both competition and natural events. Thus, some institutions may persist while others may not. For more on artificial selection see Commons (2009[1934]).

(Dosi and Nelson 1994: 154).[56] Thus, the policy options are competing with each other, and each policy option will be considered in terms of its value relative to the values of the other policies.[57]

Variation takes place when the local government switches between policies. The local government representative experiments by choosing a policy with the highest value, relative to the mean value of all the available policy options, and switches between the policies depending on the relative values for each policy (Dosi, Marengo and Fagiolo 2005). The act of switching between the policy options is known as variation.[58] Selection works together with variation to choose the better policy options, and making the less successful policies redundant (Hodgson and Knudsen 2010: Section 2.1). Thus, learning at the local government level takes place by the combination of selection and variation among the various policy options.[59]

Consequently, through selection and variation, learning allows for multiple equilibria. Hence, there can be multiple stable policy choices (Dosi and Nelson 1994: 154; Dosi, Marengo and Fagiolo 2005: 294). Thus, different local governments, even if homogeneous, may find different policy solutions for the same problem.

The final step in local government learning is retention. Retention takes place when a successful policy is kept or retained when it has a higher value – it generates more benefits or achieves the same results at lower costs, for example – relative to the other available policy options (Vanberg and Kerber 1994). Another form of retention is policy diffusion, i.e., when

[56] This is similar to Lancaster's (1966) notion of a characteristic vector that determines the consumption of a particular good.

[57] It is the competition between experimental policies that leads to new policy development (See Saviotti and Mani (1995) for a parallel in technological innovation).

[58] Variation is essential to learning as policy makers, "must anticipate the possibility of a reconsideration of their situation, a reformulation of their problems, and a change of strategies in light of experience and new information. The condition of human fallibility requires analysis, reason, deliberation, choice, experience, reconsideration, and an opportunity to alter, amend, or change as new information and new understanding give rise to new possibilities" (Ostrom 2008: Chapter 2, Paragraph 35). See also Börgers and Sarin (1997).

[59] Note that pure selection implies no learning as it relies solely on random switching behaviour. Instead of considering the value of each policy option, pure selection means that a local government official will pick a policy arbitrarily without taking into account past experience or whether the policy will be successful (Dosi, Marengo and Fagiolo 2005: 318-319).

the successful policy is reproduced by other governments (Hodgson and Knudsen 2010: Section 1.7).

3.4 Melioration as Learning

To understand how the grant recipient learns by experimenting with grant funding, we consider the recipient's behaviour upon receiving a grant. As Brennan and Pincus (1990: 136) note, "[i]f the grant is in any way contingent on the recipient's behavior (either the making of the grant at all or its size), then the grant tends to induce behavioural responses in (potential) recipients." On receipt of a grant, the rational behaviour of the local government representative will be to take policy actions tied to the grant (e.g., constructing a new road, reviving a local park).

One way to capture experimentation and learning at the local government level is as if it takes place through melioration (e.g., Herrnstein and Prelec 1991; Herrnstein et al. 1993; Metcalfe 2001). Melioration is "the process of choosing that alternative among the set of alternatives which currently has the higher yield in (value)" (Herrnstein et al 1993: 150). The local government representative, behaving as a meliorator, engages in different experiments simultaneously to find a solution to a problem. For example, different experiments to reduce crime can be tried in different boroughs of a city. Through these experiments, the local government official will switch his preference for the best policy solution depending on the value functions attached to each experiment.

Melioration is different but related to the traditional utility maximisation framework. In a maximisation framework, alternative policies are articulated together so that they lose their individuality, that is, only the utility attached to the whole choice bundle matters in a typical community iso-welfare function. Instead, melioration requires that the utility be unbundled and attached to each individual policy so that each policy can be considered separately. This

unbundled utility or disaggregated welfare is referred to as the value of the policy (Herrnstein and Pralec 1991: 154-155). [60]

The value of a policy may encompass attributes of the policy (such as compatibility and observability) and can include knowledge of how successful that policy has been in other local policy experiments. The value can also include any rents captured by implementing the policy.

The local government representative will always choose the policy with the highest value relative to any other policy options. If the policy continues to prove successful, the policy will continue to have the highest value of possible polices and the local government will retain the policy. If the policy does not prove successful, then the value of the policy will fall and when another policy has a higher value relative to that policy, the local government will switch to a different policy. Thus, implementing a chosen policy can have positive or negative effects on the relative value of the available policy options. In other words, the values associated with the policy options change endogenously during melioration. Melioration can continue up to the point where funds are used up or to the point at which the average value of the alternative policies is equal.[61]

Thus, melioration is appropriate to model experimentation and learning at the local government level because it encompasses the three mechanisms of evolutionary learning that we discussed, namely selection, variation, and retention. Under melioration a policy will be selected on the value (i.e., disaggregated welfare) attached to that policy. Variation takes place when a policy with a higher value will replace a policy with a lower value. A policy that maintains a higher value than any other policies will be retained.

[60] For more on the policy attributes see Makse and Volden (2011). Osborne and Gaebler (1993: 141-142) provide examples in the US in which output or results are measured to determine the success or failure of a policy.

[61] The point at which the average value of the alternative policies is equal is the meliorator's equilibrium where the local government representative is indifferent between different policies as they all have the same value (Herrnstein and Pralec 1991).

To illustrate melioration, consider how a local government representative switches in preference between two policies, x_1 and x_2, to solve the problem of potholes on local roads. In this illustration, x_1 may involve simply fixing each pothole individually, while x_2 may involve re-tarring the roads where there are potholes.[62] Suppose that there are two identical areas within the local community and that each policy is tested in a different area.

Each policy, x_i, has a corresponding value function $v_i(\mathrm{x})$, which is the amount of satisfaction perceived to be obtained from x_i given $x = (x_1, x_2)$. $v_1(\mathrm{x})$ is the value of x_1 which is a faster, short-run solution with little inconvenience to drivers on the road; $v_2(\mathrm{x})$ is the value of x_2 from a long-run solution because the potholes will not reappear. The weighted mean of the values is $\bar{v}(\mathrm{x}) = \sum_{i=1}^{n=2} x_i\, v_i(\mathrm{x})$. The value function indicates relative preferences for different policies: $v_1(\mathrm{x}) > v_2(\mathrm{x}) \Leftrightarrow x_1 \succ x_2$, and $v_1(\mathrm{x}) < v_2(\mathrm{x}) \Leftrightarrow x_1 \prec x_2$.

As mentioned, the value function provides feedback to the local government representative as to the relative success or failure of the policy chosen. Thus, the local government representative is aware of both value functions and their mean. Accordingly, the local representative is able to compare the two policies based on their relative values.

As previously indicated, selecting from the alternative policies can have positive or negative effects on the relative value of the policy. If the relative value of that policy starts to decline, and falls below the relative value of another policy, the recipient will "switch" to or prefer the policy with a higher value. The changes in the relative values of the different policies provide the feedback for the recipient. Hence, the change in value function $\left(\frac{\partial v_i(\mathrm{x})}{\partial x_i}\right)$ indicates whether a policy's value is increasing or decreasing. That is to say that if a policy experiment proves unsuccessful it will have a decreasing value function $\left(\frac{\partial v_i(\mathrm{x})}{\partial x_i} < 0\right)$; whereas if a policy experiment proves successful it will have an increasing value function

[62] An example of an innovation to solve the pothole problem in Johannesburg, South Africa is a smartphone application that allows a citizen to report a pothole problem with both the GPS coordinates of the location and a photograph of the pothole ("JRA plans to fix problems in a day" 2015).

$\left(\frac{\partial v_i(\mathrm{x})}{\partial x_i} > 0\right)$. A policy can also have a constant value function $\left(\frac{\partial v_i(\mathrm{x})}{\partial x_i} = k\right)$ if the value of the policy remains unchanged.

Suppose the local government receives a grant with which to experiment. The local government representative is tasked with spending the grant. Acting as a meliorator, the local government representative will always first choose that policy that has the highest value of the two policy choices. The values of the two policies are $v_1(\mathrm{x})$ and $v_2(\mathrm{x})$. As the experiment progresses, with more information about the relative success or failure of the experimental policy, the values of the two policies will change.[63]

Figure 3.1 shows how the value functions of the policies change relative to each other. Initially, policy 2 has a higher value than policy 1. However, the value of policy 2 is declining ($\frac{\partial v_2(\mathrm{x})}{\partial x_2} < 0$) and the value of policy 1 is increasing ($\frac{\partial v_1(\mathrm{x})}{\partial x_1} > 0$). Then at ME_1 the values of policy 1 and policy 2 are equal ($v_1(\mathrm{x}) = v_2(\mathrm{x})$). Since the value of policy 2 is declining and the value of policy 1 is increasing, at ME_1 the local government representative will also experiment with policy 1. The value of policy 1 continues to increase ($\frac{\partial v_1(\mathrm{x})}{\partial x_1} > 0$), but at a decreasing rate ($\frac{\partial^2 v_1(\mathrm{x})}{\partial {x_1}^2} < 0$), then falls ($\frac{\partial v_1(\mathrm{x})}{\partial x_1} < 0$), while the value of policy 2 begins to rise again ($\frac{\partial v_2(\mathrm{x})}{\partial x_2} > 0$). At ME_2, the values of both policies are equal (i.e., $v_1(\mathrm{x}) = v_2(\mathrm{x})$). Thus, ME_2 is another switching point, and, since the value of policy 1 is declining and the value of policy 2 is increasing, the local government representative switches back to policy 2, which has the higher value beyond ME_2.

Melioration continues as long as funds are available, namely until the grant funding runs out. The number of times the local government representative switches between the available

[63] From Herrnstein's (1974) matching law, since x_i represent responses and if we consider that the value functions $v_i(\mathrm{x})$ act as reinforcements, then $\frac{x_1}{x_1+x_2} = \frac{v_1(\mathrm{x})}{v_1(\mathrm{x})+v_2(\mathrm{x})}$, and, hence, $\frac{x_1}{x_2} = \frac{v_1(\mathrm{x})}{v_2(\mathrm{x})}$.

policies is an indication of the number of experiments undertaken, and, therefore, the extent of the learning that has taken place.

Figure 3.1 Melioration between Two Policy Choices

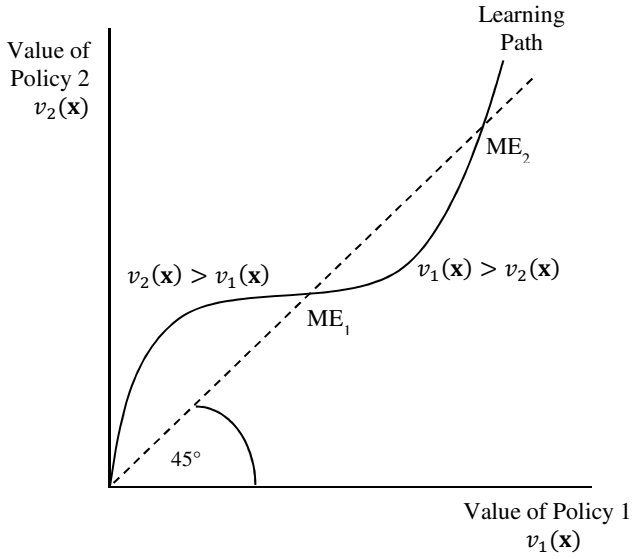


Figure 3.1 also shows a "learning path" along which policies will be implemented based on the value functions. Throughout melioration, the relatively more successful policy (as determined by the value function) will be implemented. Clearly, though, if the initial policy is successful, and maintains a value higher than the other policy choice, then there will be no need to try out a second experiment or to switch between the two policies at all. Conversely, if the initial policy is unsuccessful, the local representative can realise the mistake and correct it before the mistake becomes too costly.

Melioration is thus a trial-and-error process of selecting policy experiments that may prove successful or unsuccessful. [64] Only those policy experiments that prove successful will be retained. Melioration therefore reduces the choice set available to the policy representative to only those policy choices where the relative value is higher than the mean of the relative

[64] While economists (e.g., Dur 2001) may be concerned that government representatives may stick with their mistakes rather than admit failure to retain votes, political scientists (e.g., Volden 2007) finds that representatives actually do abandon unsuccessful policies and adopt more promising policies.

values. Consequently, it is possible to remove some of the possible policy choice mistakes, though not all. For example, if the value of a policy falls to zero, that policy will be abandoned completely. Melioration, thus, encompasses the concept of learning because, through adjustments to the value functions, the local representative ultimately reduces the mistakes and improves on the policy choices.[65]

3.5 Intergovernmental Grants: Leeway to Experiment

Assume there are two levels of government: central and local. The central government is the intergovernmental grant donor and the local government is the intergovernmental grant recipient. More specifically, within the local government, there is a local government representative who is a representative agent of the local community. The local government representative is interested in maximising political rents, and, thus, is not a perfect representative of the median voter. Instead, we can think of the local government representative as a rule follower (e.g., Langlois 1998) who has incentives to uphold the contractual terms of the grant. In this case the local government representative understands that she should use the grant funding effectively to experiment with policy to maximise rents.

Any public spending by the local government representative is limited by the local community's budget constraint. Initially, the local community budget constraint depends only on taxes raised from the local community. Assume there is no local government borrowing and taxes both at the local level and at the central level remain constant.

[65] For simplicity, this model shows melioration as a static process in which time is exogenous. However, like all evolutionary processes, there is a time component. It takes time for the local government representative to implement a policy experiment and to evaluate the experiment's success or failure. To extend this analysis, melioration can be described by a simple dynamic replicator equation, $\dot{z}_i = \gamma z_i(\alpha_i, x_i)(v_i(\mathrm{x}) - \bar{v}(\mathrm{x}))$, which shows the change in the share of income direct to a policy x_i over time, $\dot{z}_i$. $\dot{z}_i \geq 0$ is the share of income over time directed to the policy x_i, where $i = 1,2$. γ is the rate of switching between policies. Melioration will take place provided $\gamma > 0$. Thus, γ determines the extent of learning. $(v_i(\mathrm{x}) - \bar{v}(\mathrm{x}))$ is the difference between the value of the policy chosen $v_i(\mathrm{x})$ and the mean of the policy values $\bar{v}(\mathrm{x})$. Thus, the share of income directed to policy x_i over time will increase provided $\gamma > 0$, $z_i > 0$, and $v_i(\mathrm{x}) > \bar{v}(\mathrm{x})$. Underlying this equation is the principle that as the average reward for a successful policy increases, so too does the share of income assigned to it. The replicator equation shows that as time progresses, learning and experience alter the structure of policy experimentation, leading to local efficiency gains (see, e.g., Loewenstein 2010).

Assume that the local representative has local public budget B to spend. Suppose, also, that the local representative can choose to spend the local budget on two policies, x_1 and x_2. The initial budget constraint is,

$$p_1x_1 + p_2x_2 = B, \quad (1)$$

where p_i is the price paid to implement policy x_i, where $i = 1,2$.

We can rearrange equation (1) to provide the share of the budget that is assigned to each policy as follows (Metcalfe 2001):

$$\frac{p_1}{B}x_1 + \frac{p_2}{B}x_2 = 1. \quad (2)$$

Let

$$\alpha_1 = \frac{p_1}{B}, \text{ and} \quad (3)$$

$$\alpha_2 = \frac{p_2}{B}, \quad (4)$$

where α_i is the share of income devoted to policy x_i. Thus, the local budget constraint can also be expressed as,

$$\alpha_1x_1 + \alpha_2x_2 = 1. \quad (5)$$

3.5.1 Melioration with a Lump-Sum Grant

Suppose that the central government provides the local government with a lump-sum grant, (G_L). The budget constraint becomes,

$$p_1x_1 + p_2x_2 = B + G_L \quad (6)$$

Rearranging,

$$\frac{p_1}{B+G_L}x_1 + \frac{p_2}{B+G_L}x_2 = 1, \quad (7)$$

where $\alpha_i = \frac{p_i}{B+G_L}$

Proceeding along the lines of Herrnstein and Prelec (1991) and Metcalfe (2001), the melioration equilibrium can be specified with matching conditions,

$$v_1(x_1, x_2) = v_2(x_1, x_2), \quad (8)$$

in which the values of both policies are equal, and

$$\alpha_1 x_1 + \alpha_2 x_2 = 1, \quad (9)$$

and where the budget (including the lump-sum grant) is spent.

Taking total differentials of the matching condition, the system of adjustments that result from the lump-sum grant can be represented as,

$$\begin{bmatrix} \frac{\partial v_1}{\partial x_1} & \frac{\partial v_2}{\partial x_2} \\ \alpha_1 & \alpha_2 \end{bmatrix} \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{B + G_L} dG_L \end{bmatrix}.$$

Solving the system of adjustments yields

$$\frac{dx_1}{dG_L} = \frac{\frac{\partial v_2}{\partial x_2}}{\Delta(B+G_L)} > 0, \text{ and} \quad (10)$$

$$\frac{dx_2}{dG_L} = \frac{\frac{\partial v_1}{\partial x_1}}{\Delta(B+G_L)} > 0, \quad (11)$$

where $\Delta = \alpha_2 \frac{\partial v_1}{\partial x_1} + \alpha_1 \frac{\partial v_2}{\partial x_2}$.

Equations (10) and (11) show that an increase in G_L results in an increase in policy experimentation on both policies. Any decision to choose between experimenting on policy 1 or policy 2 depends only on the changes to the relative values of both policies and the size of the budget. Hence, the local government representative can choose freely between either of the two policy options, and can meliorate between the two until the lump-sum grant is used up.

3.5.2 Melioration with a Closed Matching Grant

Suppose that the central government provides the local government with a closed matching grant (G_M) with the explicit condition that it be spent on x_1, i.e. the closed matching grant increases the spending on x_1 by making it relatively cheaper. Let the closed matching grant

rate be m, where $0 < m < 1$. So, for example, if the central government decides to give 50% of what the local government has spent on the public policy as a grant, then $m = 0.5$. Let $G_M = mp_1x_1$. Thus the budget constraint becomes,

$$p_1x_1 + p_2x_2 = B + G_M, \tag{12}$$

which, again, can be rearranged as

$$\frac{p_1}{B+G_M}x_1 + \frac{p_2}{B+G_M}x_2 = 1 \tag{13}$$

where $\alpha_i = \frac{p_i}{B+G_M}$.

Once again, proceeding along the lines of Herrnstein and Prelec (1991) and Metcalfe (2001), the melioration equilibrium can be specified with matching conditions,

$$v_1(x_1, x_2) = v_2(x_1, x_2), \tag{14}$$

in which the values of both policies are equal, and

$$\alpha_1x_1 + \alpha_2x_2 = 1, \tag{15}$$

and where the budget (including the closed matching grant) is spent.

The system of adjustments that result from a closed matching grant can be represented as,

$$\begin{bmatrix} \frac{\partial v_1}{\partial x_1} & \frac{\partial v_2}{\partial x_2} \\ \alpha_1 & \alpha_2 \end{bmatrix} \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix} = \begin{bmatrix} 0 \\ (B + G_M)dG_M \end{bmatrix}.$$

Solving the system of adjustments for a closed matching grant yields

$$\frac{dx_1}{dG_M} = \frac{\frac{\partial v_2}{\partial x_2}}{\Delta(B+G_M)} = \frac{\frac{\partial v_2}{\partial x_2}}{\Delta(B+mp_1x_1)} > 0, \text{ and} \tag{16}$$

$$\frac{dx_2}{dG_M} = \frac{\frac{\partial v_1}{\partial x_1}}{\Delta(B+G_M)} = \frac{\frac{\partial v_1}{\partial x_1}}{\Delta(B+mp_1x_1)} > 0, \tag{17}$$

where $\Delta = \alpha_2\frac{\partial v_1}{\partial x_1} + \alpha_1\frac{\partial v_2}{\partial x_2}$.

Equations (16) and (17) are both positive. This means that a closed matching grant to one policy can increase experimentation on both policies in absolute terms. However, if we compare equation (10) with equation (16) and equation (11) with equation (17), the change in

expenditure on both policies is dependent on the grant-aided policy, x_1. This means that although experimentation can increase, the experimentation is constrained by the commitment of the grant recipient to honour the matching condition of the grant. Hence, the matching condition can, in effect, crowd out experimentation on a policy that it is not specifically intended for. The lump-sum grant has no such restrictions. Hence, this illustration shows that a lump-sum grant provides the grant recipient with greater leeway to experiment through melioration than a closed matching grant of the same size.

3.5.3 Graphical Comparison of Melioration under Different Grants

The learning effects of a lump-sum grant compared to a closed matching grant can also be shown graphically. In Figure 3.2, the closed matching grant-aided policy, x_1, is on the horizontal axis and an alternative policy, x_2, is on the vertical axis. The local community budget constraint in the absence of intergovernmental grants is represented by AB. We can think of the budget constraint as a boundary of spending by the local community determined by local public fund-raising efforts (e.g., local taxes).

Similarly, we can consider the budget line that incorporates an intergovernmental grant as another boundary, the shape of the boundary is dictated by the type of grant. DF represents the shift in the budget constraint due to the lump-sum grant which has no explicit conditions attached (Equation 7). AMF represents the shift in the budget constraint due to the closed matching grant (Equation 13), which specifies the matching rate m for which the local government's spending on closed matching grant-aided policy, x_1.

When a local community receives a grant, the outward shift of the budget constraint together with the initial budget constraint defines a space in which the local representative can experiment with policies: for a lump-sum grant (DF) the space is ABFD and for a closed matching grant (AMF) the space is ABFM. The experimentation with the grant by the local

representative will also be constrained by the initial level of local government spending at I. Assume the prices of the policy experiments remain unchanged by undertaking the experiment.

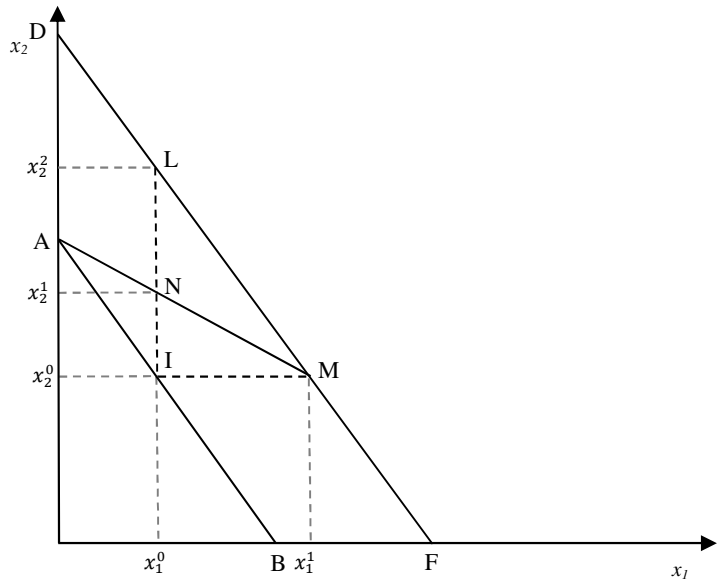
First, consider the lump-sum grant (DF). Upon receipt of DF, the local representative can experiment by meliorating between whichever policy (x_1 or x_2) has the higher value until the lump-sum grant is spent, viz. (along line LM). Thus, the potential space for experimenting between policies for a lump-sum grant is represented by the area ILM.

Now consider the closed matching grant (AMF). Upon receipt of AMF, the local representative can, as for the lump-sum grant, choose how to spend the closed matching grant but now has to keep to the conditions attached to the closed matching grant. The local representative can experiment though meliorating between the two policies (x_1 and x_2) until the closed matching grant is spent (i.e., along the line NM). The conditions attached to the closed matching grant crowd out some of the potential to experiment in comparison to the lump-sum grant. Thus, the potential for switching between policies for a closed matching grant is represented by the area, INM, which is smaller than ILM.

As shown in Figure 3.2, the closed matching grant conditions crowd out potential melioration and, thus, policy experimentation. The extent to which the potential experimentation is crowded out is represented by area NLM which is the difference between ILM (i.e., the area for potential experimentation by melioration for a lump-sum grant) and INM (i.e., the area for potential experimentation by melioration for a closed matching grant). We can thus think of the two areas, ILM and NLM, representing the potential for learning from both grants. Since ILM is the bigger area, a lump-sum grant with no explicit conditions has greater potential for learning through experimentation. The conditions attached to a closed matching grant limits the freedom of the local representation to choose between potential policies and, thus, crowds out potential learning through experimentation.

While the graphical representation of the learning effects is static, it has the benefit of indicating the learning potential of these two types of intergovernmental grants by extending the traditional analysis of intergovernmental grants. See Appendix 3A for two near-corner examples.

Figure 3.2 Learning under a Lump-sum Grant compared to a Closed Matching Grant



Notes: The initial level of local government spending is at I where $(x_1, x_2) = (x_1^0, x_2^0)$. Upon receipt of a lump-sum grant (DF), the local representative can choose whether to spend the grant solely on x_1 (i.e., to point M where $(x_1, x_2) = (x_1^1, x_2^0)$); on x_2 (i.e., to point L where $(x_1, x_2) = (x_1^0, x_2^2)$); or on a combination of x_1 and x_2 (i.e., along the line LM). Starting with whichever policy (x_1 or x_2) has the higher initial value, the local representative can meliorate between the two policies until the lump-sum grant is spent, i.e., along line LM. Thus, the potential for switching between policies for a lump-sum grant is represented by the area, ILM. Upon receipt of the closed matching grant (AMF), the local representative can, as for the lump-sum grant, choose how to spend the closed matching grant but now has to keep within the conditions attached to the closed matching grant. Only if the local representative chooses to spend the closed matching grant solely on x_1 (i.e., to point M where $(x_1, x_2) = (x_1^1, x_2^0)$), then the size of the grant will be the same as the lump-sum grant. If the local representative chooses to spend the closed matching grant solely on x_2, she will only be able to spend to point N (i.e., where $(x_1, x_2) = (x_1^0, x_2^1)$). Hence the matching conditions have the potential to crowd out $x_2^2 - x_2^1$ spending on x_2. As for the lump-sum grant, starting with whichever policy, x_1 or x_2, has the higher initial value the local representative can meliorate between the two policies until the closed matching grant is spent (i.e., along line NM). Thus, the potential for switching between policies for a closed matching grant is represented by the area, INM.

3.6 Summary and Areas for Future Research

This chapter contributes to the laboratory federalism literature by showing how intergovernmental grants can be used to induce policy experimentation and which types of intergovernmental grants will be most appropriate for this purpose. The field of laboratory federalism has developed largely independently of the theory of intergovernmental grants. As we saw, one issue that laboratory federalism is concerned with is how to create incentives for local governments to experiment with policy (see, e.g., Rose-Ackerman 1980, Strumpf 1998, Garzarelli 2006, Bednar 2011 and Madison 2014). However, as noted by Garzarelli (2006), the field has been missing a policy instrument to create those incentives. Thus, as Garzarelli (2006) suggests, and as the model presented in this chapter shows, the field of laboratory federalism can be extended to view intergovernmental grants as the missing policy instrument to incentivise experimentation.

The traditional approach to intergovernmental grants suggests two main recommendations for grants. The first prescribes lump-sum grants to address fiscal imbalances between governments. The second prescribes matching grants (either closed- or open-ended) to internalise interjurisdictional externalities, i.e., a matching grant is equivalent to a Pigouvian subsidy (e.g., Oates 1972; King 1984). This chapter extends the possible applications for intergovernmental grants to policy experimentation. As shown, a grant with relatively fewer conditions attached, such as a lump-sum grant, is the policy instrument of choice to incentivise local government learning. SCHIP is an example where a lump-sum grant gave different states the flexibility to experiment with different policy applications and where the federal government learned from these experiments to later develop federal legislation (Weissert and Scheller 2008). Similarly, in the case of Scottish university funding, Keating (2005a,b) suggests that the lump-sum grants allocated allowed the universities to experiment

and find new policy solutions that were later adopted by English universities.[66] Lump-sum grants and closed matching grants are not the only two types of grants though. The analysis here can be extended to other types of intergovernmental grants that fall in between these two grants in terms of leeway as well, e.g., a conditional lump-sum grant or an open-ended matching grant.

It is against this backdrop that this chapter shows one possible application of evolutionary theory to public economics. Evolutionary theory has thus far been applied in many economic contexts such as technological change in both organisations and industries, as well as in economic growth (see, e.g., Nelson and Winter 1977, 2002; Dosi 1988; Boulding 1991; Saviotti and Mani 1995; Metcalfe 2001; Consoli 2008; Dopfer 2013). More recently, evolutionary theory has been applied to issues of public policy too (e.g., Wilson and Gowdy 2013). However, the application of evolutionary theory to the field of public economics has been lacking.

In this chapter, experimentation with policy is modelled through the evolutionary process of melioration. Melioration was first used to model both animal and human behaviour (Herrnstein 1974). It was later used in consumption models as an alternative to utility maximisation (Herrnstein and Prelec 1991; Herrnstein et al. 1993; Metcalfe 2001; Consoli 2008). The model presented in this chapter uses melioration differently. This model extends the application of melioration by using it as a proxy for the trial-and-error process of undertaking different policy experiments. As this model shows, experimenting with policy by melioration encompasses selection, variation and retention. One way that retention takes place is when the experimenting local government keeps using the most successful policy. Another way that retention takes place is when other governments in the federation (e.g.,

[66] See Appendix 3B for more on how other local governments adopt successful policies.

other local governments or even the central government) learn from the experiments and start to adopt the successful policies.

When governments in a federation learn from each other's policy experiments, we can view the federation as a network for knowledge creation. Thus, within a federation, the central government can use grants strategically either to direct multiple local governments to experiment with different policy solutions to the same problem or to direct multiple local governments to experiment to solve different problems.

The experiments can generate knowledge of both successful and unsuccessful policies. The knowledge of both can be valuable to the federation: successful policies can be replicated, and the repeating of unsuccessful policies can be avoided. Thus, the results of all experiments can generate positive externalities in the form of improved policy knowledge for the rest of the federation. The the model presented in this chapter finds that there is greater potential for experimentation from a lump-sum grant than from a closed matching grant: a result of this finding is that there is potential for a larger knowledge externality to be generated from a lump-sum grant than from a closed matching grant.

An area in which to extend this research would be the flypaper effect, in particular the *micro* flypaper effect.[67] The *micro* effect considers the different effects that different types of intergovernmental grants have on local government spending. For example, a conditional lump-sum grant is typically more stimulating for that public good for which it is intended than a conditional closed matching grant (Brennan and Pincus 1990: 134-136). In a meliorating framework, the flypaper effect is likely explained by money sticking to those policies that prove successful (or at least are attached to higher value functions than other available policies).

[67] For a review of the flypaper literature see Bailey and Connolly (1998). Brennan and Pincus (1990) suggest a distinction between *macro* and *micro* flypaper effects. The *macro* flypaper effect is more well-known and is the phenomenon in which an intergovernmental grant leads to greater public spending than an equivalent increase in community income, e.g. a tax cut.

As this is a start in considering an evolutionary view of public economics, specifically in the field of laboratory federalism, the scope for research is wide open. One possibility is to evaluate the evolutionary imbalance hypothesis in the context of policy experimentation to determine whether competitive local governments produce more successful polices than less competitive ones.

3.7 References

Arrow, Kenneth. J. 1962. "The Economic Implications of Learning by Doing," *Review of Economic Studies* 29(3): 155-173(June).

Allen, Mahalley D., Carrie Pettus, and Donald P. Haider-Markel 2004. "Making the National Local: Specifying the Conditions for National Government Influence on State Policymaking," *State Politics and Policy Quarterly* 4(3): 318-344(Fall).

Bailey, Stephen J. and Stephen Connolly 1998. "The Flypaper Effect: Identifying Areas for Further Research," *Public Choice* 95(3/4): 335-361 (June).

Bednar, Jenna 2011. "Nudging Federalism towards Productive Experimentation," *Regional and Federal Studies* 21(4/5): 503-521(October/December).

Bennhold, Katrin 2013(December 7) "Britain's Ministry of Nudges," *New York Times* Available online: http://www.nytimes.com/2013/12/08/business/international/britains-ministry-of-nudges.html?_r=0 Last Accessed October 19, 2015.

Besfamille, Martin and Ben Lockwood 2008. "Bailouts in Federations: Is a Hard Budget Constraint Always Best?," *International Economic Review* 49(2): 577-593(May).

Boadway, Robin and Michael Keen 1996. "Efficiency and the Optimal Direction of Federal-State Transfers," *International Tax and Public Finance* 3(2): 137-155(May).

Börgers, Tilman and Rajiv Sarin 1997. "Learning through Reinforcement and Replicator Dynamics," *Journal of Economic Theory* 77(1):1-14(November).

Boschken, Herman L. 1982. "Organization Theory and Federalism: Interorganizational Networks and the Political Economy of the Federalist," *Organization Studies* 3/4: 355-373.

Bosman, Julie 2009. (April 22). "City Ends Some Antipoverty Experiments," *New York Times* Available online: http://cityroom.blogs.nytimes.com/2009/04/22/bloomberg-ends-some-antipoverty-experiments Last Accessed March 22, 2013.

Boulding, K.E. 1991. "What is Evolutionary Economics?" *Journal of Evolutionary Economics* 1(1):9-17(March).

Bradford, David F. and Wallace E. Oates 1971. "Towards a Predictive Theory of Intergovernmental Grants," *American Economic Review* 61(2): 440-448(May).

Bradbury, Bill 2005. (January 1). "Vote-by-Mail: The Real Winner is Democracy," *Washington Post* Available online: http://www.washingtonpost.com/wp-dyn/articles/A40032-2004Dec31.html Last Accessed January 26, 2015.

Brennan, Geoffrey and Jonathan Pincus 1990. "An Implicit Contract Theory of Intergovernmental Grants," *Publius: The Journal of Federalism* 29(4): 129-44(Fall).

Breton, Albert 1987[1985]. "Towards a Theory of Competitive Federalism," *European Journal of Political Economy* 3 (1-2): 263-329.

Breton, Albert 1998[1996]. *Competitive Governments: An Economic Theory of Politics and Public Finance*. New York: Cambridge University Press.

Breton, Albert 2000. "Federalism and Decentralization: Ownership Rights and the Superiority of Federalism," *Publius* 30 (2): 1-16 (Spring).

Buchanan, James M. 1984. "Schumpeter as Precursor to Non-Constitutional Public Choice," *Economia delle scelte pubbliche* 1: 51–53.

Caballero, María Cristina 2004 (March 11). "Academic turns city into a social experiment," *Harvard University Gazette* Available online: http://news.harvard.edu/gazette/2004/03.11/01-mockus.html Last Accessed January 28, 2015.

Cai, Hongbin and Daniel Treisman 2009. "Political Decentralization and Policy Experimentation," *Quarterly Journal of Political Science* 4(1): 35-58.

Callander, Steven 2011a. "Searching and Learning by Trial and Error," *American Economic Review* 101(6): 2277-2308(October).

Callander, Steven 2011b. "Searching for Good Policies," *American Political Science Review* 105(4):643-662(November).

Callander, Steven and Bård Harstad 2013. "Experimentation in Federal Systems," NBER Working Paper Series No 19601 (October).

Callander, Steven and Patrick Hummel 2014. "Preemptive Policy Experimentation," *Econometrica* 82(4)" 1509-1528(July).

Carlsen, Fredrik 1995. "Why Is Central Regulation of Local Spending Decisions So Pervasive? Evidence from a Case Study," *Public Budgeting & Finance* 15(1): 43-57 (March).

Commons, John R. 2009[1934]. *Institutional Economics: Its place in Political Economy Volume 1* (Reprint with an introduction by Malcolm Rutherford) New Jersey: Transaction Publishers.

Consoli, Davide 2008. "Co-evolution of Capabilities and Preferences in the Adoption of New Technologies," *Technological Analysis and Strategic Management* 20(4): 409-425(July)

Courant, Paul N., Edward M. Gramlich and Daniel L. Rubinfeld 1979. "The Stimulative Effects of Intergovernmental Grants: Or Why Money Sticks Where it Hits," Chapter 2 in Peter Mieszkowski and William H.Oakland (Eds) *Federalism and Grants-in-Aid* Washington D.C.: The Urban Institute.

Dopfer, Kurt 2013. "Evolutionary Economics," Papers on Economics and Evolution, Working Paper 1308. Max Planck Institute of Economics,

Dosi, Giovanni 1988. "Sources, Procedures and Microeconomic Effects of Innovation," *Journal of Economic Literature* 26(3): 1120-71(September).

Dosi, Giovanni, Luigi Marengo, Giorgio Fagioli 2005. "Learning in Evolutionary Environments," Chapter 9 in Kurt Dopfer (Ed.) *The Evolutionary Foundation of Economics* Cambridge: Cambridge University Press.

Dosi, Giovanni and Richard Nelson 1994. "An Introduction to Evolutionary Theories in Economics," *Journal of Evolutionary Economics* 4(3): 153-172(September).

Dur, Robert A.J. 2001. "Why Do Policy Makers Stick to Inefficient Decisions?" *Public Choice* 107(3/4): 22—234 (June).

Social Sciences and Behavioral Team (2015) *Annual Report* Executive Office of the President, National Science and Technology Council Available online: https://www.whitehouse.gov/sites/default/files/microsites/ostp/sbst_2015_annual_report_final_9_14_15.pdf Last Accessed October 19, 2015.

Forte, Francesco 1982. "The Law of Selection in the Public Economy as compared to the Market Economy," *Public Finance - Finances publiques* 37(2): 224-245.

Galle, Brian and Joseph Leahy 2009. "Laboratories of Democracy? Policy Innovation in Decentralized Governments," *Emory Law Journal* 58(6):1333-1400.

Garzarelli, Giampaolo 2004. "Open Source Software and the Economics of Organization," in J. Birner and P. Garrouste (Eds.), *Markets, Information and Communication* London and New York: Routledge (An earlier version is available online: http://ideas.repec.org/p/wpa/wuwpio/0304003.html Last Accessed October 25, 2015)

Garzarelli, Giampaolo 2006. "Cognition, Incentives, and Public Governance: Laboratory Federalism from the Organizational Viewpoint," *Public Finance Review* 34(3): 235-257(May).

Goodspeed, Timothy J. 2002. "Bailouts in a Federation," *International Tax and Public Finance* 9:409-421(August).

Hart, Oliver 1988. *Firms, Contracts, and Financial Structure* Oxford: Oxford University Press.

Herrnstein, R. J. 1974. "Formal properties of the matching law," *Journal of the Experimental Analysis of Behavior* 21(1): 159–164(January).

Herrnstein, Richard J. and Drazen Prelec 1991. "Melioration: A Theory of Distributed Choice," *Journal of Economic Perspectives* 5(3): 137-156(Summer).

Herrnstein, Richard J., George F. Loewenstein, Drazen Prelec and William Vaughan Jr. 1993. "Utility Maximization and Melioration: Internalities in Individual Choice," *Journal of Behavioral Decision Making* 6(3):149-185(September).

Hodgson, Geoffrey M. and Thorbjørn Knudsen 2010. *Darwin's Conjecture: The Search for General Principles of Social and Economic Evolution* Chicago and London: University of Chicago Press [Kindle version, Retrieved from www.amazon.com].

"JRA plans to fix problems in a day" 2015 (March 17) *Rosebank Killarney Gazette* Available online: http://rosebankkillarneygazette.co.za/154520/jra-plans-to-fix-problems-in-a-day/ Last Accessed March 26, 2015.

Keating, Michael 2005a. "Higher Education in Scotland and England after Devolution," *Regional and Federal Studies* 15(4): 423-435(December).

Keating, Michael 2005b. "Policy Convergence and Divergence in Scotland under Devolution," *Regional Studies* 39(4): 453-463(June).

Kerber, Wolfgang and Martina Eckardt 2007. "Policy Learning in Europe: the Open Method of Co-ordination and Laboratory Federalism," *Journal of European Public Policy* 14(2):227-247(March).

King, David Neden. 1984. *Fiscal Tiers: The Economics of Multilevel Government* London: George Allen & Unwin.

Kleiman, Mark A.R and Jeremy Ziskind 2014. "Lawful Access to Cannabis: Gains, Losses and Design Criteria," in LSE Expert Group on the Economics of Drug Policy Report *Ending the Drug Wars* (May).

Kollman, Ken, John H. Miller and Scott E. Page 2000 "Decentralization and the Search for Policy Solutions," *Journal of Law, Economics, & Organization* 16(1): 102-128(April).

Kornai, János 1986. "The Soft Budget Constraint," *Kyklos* 39(1): 3-30.

Kotsogiannis, Christos and Robert Schwager 2006. "On the Incentives to Experiment in Federations," *Journal of Urban Economics* 60(3): 484–497(November).

Lancaster, Kelvin 1966. "Change and Innovation in the Technology of Consumption," *American Economic Review* 56(1/2): 14-23 (March).

Langlois, Richard N. 1982. "Cost-Benefit Analysis, Environmentalism, and Rights," *Cato Journal* 2(1):279-300(Spring).

Langlois, Richard N. 1998. "Rule-Following, Expertise, and Rationality: A New Behavioral Economics?," Kenneth Dennis (Ed.) *Rationality in Economics: Alternative Perspectives*. Dordrecht: Kluwer Academic Publishers. Reprinted in Massimo Egidi and Salvatore Rizzello (Eds) 2004. *Cognitive Economics*. Cheltenham: Edward Elgar. Available online: http://web.uconn.edu/ciom/RATIONL2.HTML Last Accessed October 25, 2015.

Lindblom, Charles E. 1959. "The Science of 'Muddling Through'," *Public Administration Review* 19(2): 79-88(Spring).

Lindblom, Charles E. 1979. "Still Muddling, Not Yet Through," *Public Administration Review* 39(6): 517-526(November-December).

Loewenstein, Yonatan 2010. "Synaptic Theory of Replicator-like Melioration," *Frontiers in Computational Neuroscience* 4(17):1-12(June).

Madison, Kristin M. 2014. "Building a Better Laboratory: The Federal Role in Promoting Health System Experimentation," *Pepperdine Law Review* 41(4): 765-816 (May).

Makse, Todd and Craig Volden 2011. "The Role of Policy Attributes in the Diffusion of Innovations," *Journal of Politics* 73(1):108-124(January).

Malerba, Franco 2006. "Innovation and the Evolution of Industries," Journal of Evolutionary Economics 16(1-2):3-23(April).

Metcalfe, J.S 2001. "Consumption, Preferences, and the Evolutionary Agenda," Journal of Evolutionary Economics 11(1):37-58(January).

Mieszkowski, Peter and Richard A. Musgrave 1999. "Federalism, Grants, and Fiscal Equalization," *National Tax Journal* 52(2):239-60 (June).

Nelson, Richard R., and Sidney G. Winter 1977. "In Search of Useful Theory of Innovation," *Research Policy* 6(1): 36-76(January).

Nelson, Richard R. and Sidney G. Winter 2002. "Evolutionary Theorizing in Economics," *Journal of Economic Perspectives* 16(2): 23-46(Spring).

Oates, Wallace E. 1972. *Fiscal Federalism*. New York: Harcourt Brace Jovanovich.

Oates, Wallace E. 1999. "An Essay on Fiscal Federalism," *Journal of Economic Literature* 37(3):1120- 49(September).

Oates, Wallace E. 2005. "Toward a Second-generation Theory of Fiscal Federalism," *International Tax and Public Finance* 12(4): 349-373(August).

Oates, Wallace E. 2008. "On The Evolution of Fiscal Federalism: Theory and Institutions," *National Tax Journal* 16(2): 313-334(June).

Oates, Wallace E. 2011. "On the Development of the Theory of Fiscal Federalism: An Essay in the History of (Recent) Economic Thought," *Journal of Public Finance and Public Choice - Economia delle scelte pubbliche*, XXIX(1-3):3-24.

Osborne, David and Ted Gaebler 1993. *Reinventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector* New York: Plume.

Ostrom, Vincent 2008. *The Political Theory of a Compound Republic: Designing the American Experiment* (3rd Edition). Lanham: Lexington Books [Kindle version, Retrieved from www.amazon.com].

Ostrom, Vincent and Elinor Ostrom 1965. "A Behavioral Approach to the Study of Intergovernmental Relations," *Annals of the American Academy of Political and Social Science* 359: 137-146(May).

Oulasvirta, Lasse 1997. "Real and Perceived Effects of Changing the Grant System from Specific to General Grants," *Public Choice* 91: 397-416.

Rodden, Jonathan 2003. "Reviving Leviathan: Fiscal Federalism and the Growth of Government," *International Organization* 57(4): 695-729(November).

Rodrik, Dani 2014. "When Ideas Trump Interests: Preferences, World Views, and Policy Innovations," *Journal of Economic Perspectives* 28(1):189–208(Winter).

Romano, Roberta 2006. "The States as a Laboratory: Legal Innovation and State Competition for Corporate Charters," *Yale Journal on Regulation* 23:209-247.

Romer, Thomas and Howard Rosenthal 1980. "An Institutional Theory of the Effect of Intergovernmental Grants," *National Tax Journal* 33(4): 451-458(December).

Rose-Ackerman, Susan 1980. "Risk Taking and Reelection: Does Federalism Promote Innovation?" *Journal of Legal Studies* 9(3): 593-616(June).

Sabel, Charles F. 2004. "Beyond Principal-Agent Governance: Experimentalist Organizations, Learning and Accountability," draft discussion paper prepared for WRR meeting, Amsterdam (May 10-14, 2004), in Ewald Engelen & Monika Sie Dhian Ho (Eds.), *De Staat van de Democratie. Democratie voorbij de Staat*. WRR Verkenning 3: 173-195 Amsterdam: Amsterdam University Press

Salmon, Pierre 2005. "Horizontal Competition Among Governments," Working Paper No 2005-02, LEG - Document de travail - Economie from LEG, Laboratoire d'Economie et de Gestion, CNRS, Université de Bourgogne Available online: http://leg.u-bourgogne.fr/images/stories/pdf/doc_trav2005/e2005-02.pdf Last Accessed November 12, 2014.

Saviotti, P.P and G.S Mani 1995. "Competition, Variety and Technological Evolution: A Replicator Dynamics Model," *Journal of Evolutionary Economics* 5(4): 369-392 (December).

Schnyder, Simon 2011. "Laboratory Federalism: Policy Diffusion and Yardstick Competition," Paper presented to the 8th Public Economics at the Regional and Local Level (PEARL) Conference Brno, Czech Republic, 24 May.

Seabright, Paul 1996. "Accountability and Decentralision in Government: An Incomplete Contracts Model," *European Economic Review* 40(1): 61-89(January).

Slabbert, Antoinette 2015(April 1). "Treasury Cracks Whip on Municipalities owing Eskom," *Moneyweb Available online:* http://www.moneyweb.co.za/news/south-africa/treasury-cracks-whip-on-municipalities-owing-eskom/ Last Accessed October 19, 2015.

Social Sciences and Behavioral Team (2015) *Annual Report* Executive Office of the President, National Science and Technology Council Available online: https://www.whitehouse.gov/sites/default/files/microsites/ostp/sbst_2015_annual_report_final_9_14_15.pdf Last Accessed October 19, 2015.

Stolberg, Sheryl Gay 1999. (May 6). "Pennsylvania Set to Break Taboo on Reward for Organ Donations," *New York Times* Available online: http://www.nytimes.com/1999/05/06/us/pennsylvania-set-to-break-taboo-on-reward-for-organ-donations.html Last Accessed January 26, 2015.

Strumpf, Koleman S. 2002. "Does Government Decentralization Increase Policy Innovation?" *Journal of Public Economic Theory* 4(2): 207-41(April).

Vanberg, Viktor and Wolfgang Kerber 1994. "Insitutional Competition Among Jurisdictions," *Constitutional Political Economy* 5(2): 193-219 (March).

Vigneault, Marianne 2007. "Grants and Soft Budget Constraints," Chapter 5 in Robin Boadway and Anwar Shah (Eds.) *Intergovernmental Fiscal Transfers: Principles and Practice* Washington D.C.: The World Bank.

Volden, Craig 1999. "Asymmetric Effects of Intergovernmental Grants: Analysis and Implications for U.S. Welfare Policy," *Publius* 29(3): 51-73(Summer).

Volden, Craig 2007. "Failures: Diffusion, Learning, and Policy Abandonment," *SSRN Working Paper* (October). Available online: http://ssrn.com/abstract=1020750 Last Accessed November 13, 2013.

Weingast, Barry R. 2006. "Second Generation Fiscal Federalism: Implications for Decentralized Democratic Governance and Economic Development," IFIR-CESifo Conference on "New Directions in Fiscal Federalism" Available online: http://www.ifigr.org/workshop/IFIR-CESifo/papers/weingast.pdf Last Accessed January 27, 2015

Weingast, Barry R. 2009. "Second Generation Fiscal Federalism: The Implications of Fiscal Incentives," *Journal of Urban Economics* 65(3): 279-293(May).

Weissert, Carol S. and Daniel Scheller 2008. "Learning from the States? Federalism and National Health Policy," *Public Administration Review* 68(Issue Supplement): S162–S174 (December).

Welch, Susan and Kay Thompson 1980. "The Impact of Federal Incentives on State Policy Innovation," *American Journal of Political Science* 24(4): 715-729(November).

Wilde, James A. 1968. "The Expenditure Effects of Grant-in-Aid Programs," *National Tax Journal* 21(3):340-348(September).

Wilson, David Sloan and John M. Gowdy 2013. "Evolution as a General Theoretical Framework for Economics and Public Policy," *Journal of Economic Behavior and Organization* 90(Supplement):S3-S10(June).

Appendix 3A
Two Extreme Cases of Grant Experimentation

The area for potential experimentation under a lump-sum grant, ILM, depends on the size of the grant and the initial point of consumption I. The area for potential experimentation under a closed matching grant, INM, depends on the matching ratio, size of the grant, and the initial point of consumption I. Consider now two extreme cases shown in Figure 3A.1 and Figure 3A.2.

Figure 3A.1 Closed Matching Grant Conditions Crowd-Out Experimentation

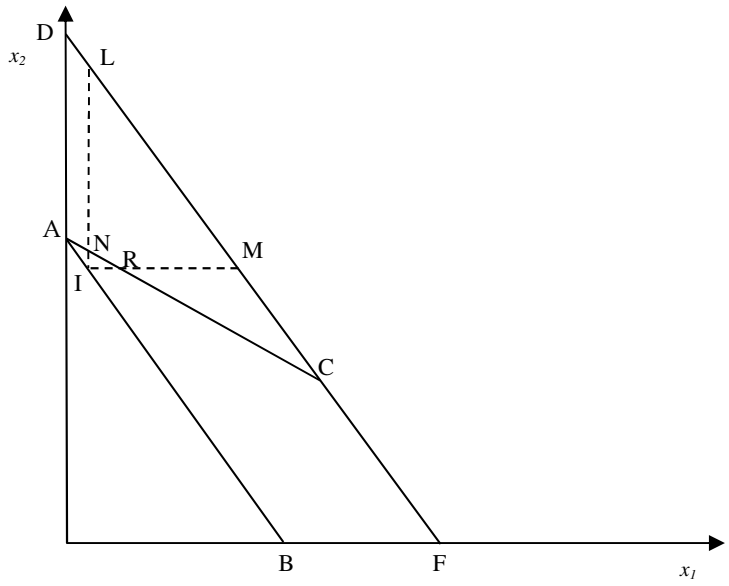


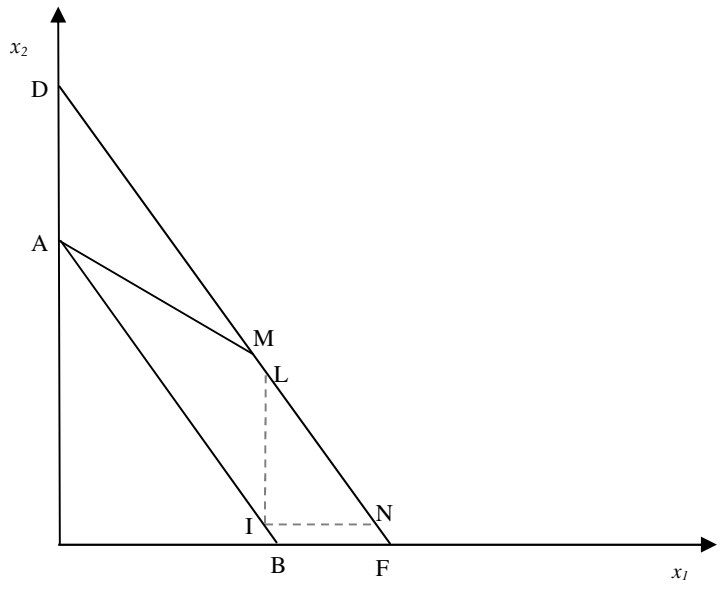
Figure 3A.1 shows a near corner solution[68] at which the local representative has chosen to spend very little on the grant-aided policy initially at point I. In this case, with a closed matching grant the area of possible experimentation by melioration is very small, i.e., INR. A lump-sum grant, however, will allow the same scope for experimentation as in Figure 3.2,

[68] Intergovernmental grant analysis is usually restricted to interior solutions (see, e.g., King 1984).

represented by the area ILM. Since the initial spending on the grant-aided policy is so low, much of the possible additional experimentation is crowded out by the close matching grant conditions.

Figure 3A.2 shows another near corner solution. However, this time the local representative has chosen to spend very little on any policy other than the grant-aided policy. In this case, though, the closed matching grant can provide as much flexibility in the scope for experimenting as the lump-sum grant, i.e., ILN. This extreme case presents the opposite case to that in Figure 3A.1. Here the local representative has chosen to spend sufficiently on the grant-aided policy initially to prevent the conditions of the closed matching grant from crowding out any possible grant-aided policy experimentation.

Figure 3A.2: No Crowding Out of Experimentation



The two extreme cases show that the amount of learning that can be crowded out by the conditions attached to a grant depends on how the local representative chooses to spend portion of the community budget that excludes the grant. Thus, an intergovernmental grant

can effectively impose implicit conditions on the rest of the local community's budget, such as influencing the initial point of local governments spending, I.

Appendix 3B
The Process of Policy Adoption

3B.1 Replication of a Policy

Consider two policy options in a federation, policy P, which is the existing policy, and policy N, which is a new policy innovation developed by an innovator local government. Let $g(t)$ be number of local governments that have implemented P (i.e., the frequency of policy P) and $h(t)$ be the number of local governments that have implemented policy N (i.e., the frequency of policy N).

Every local government has chosen either P or N, such that $g(t) + h(t) = 1$. Policy P is copied by local governments at a rate p and Policy N is copied by local governments at rate n, with $p \neq n$. Thus, the system of equations that represents the growth of the two policies is:

$$\dot{g} = g(p - \varphi), \tag{B.14}$$

$$\dot{h} = h(n - \varphi), \tag{B.2}$$

where $\varphi = pg + nh$ is the average fitness of both policies in the federation, which can be thought of as the average number of local governments in the federation that have chosen to adopt P or N.

Given $g(t) + h(t) = 1$ and $\varphi = pg + nh$ the system of equations (Equations B.1 and B.2) can reduce to a single equation,

$$\dot{g} = g(1 - g)(p - n), \tag{B.15}$$

which has two equilibria. The first equilibrium is when $g = 0$ and the second equilibrium is when $g = 1$. For the first, if $g = 0$, policy P is replaced entirely by policy N in the federation. For the second, if $g = 1$, then policy P replaces any instances of policy N in the federation, meaning that N never gets adopted.

Additionally, if $p > n$, then $\dot{g} > 0$ for all $g \in (0,1)$. Thus, for a federation consisting of some local governments using policy P and others using policy N, the number of local governments using P will increase as the rate at which P is copied by other local governments exceeds the rate at which N is being adopted by local governments. Similarly if $n > p$, $\dot{h} > 0$ for all $h \in (0,1)$. Thus, again, for a federation consisting of some local governments using policy P and others using policy N, the number of local governments using N will increase as the rate at which N is adopted by local governments exceeds the rate at which N is being adopted by local governments.

3B.2 Replication with Mutation

Policy mutation can be defined as "unplanned policy experimentation that arises along the margins of existing policies" (Rodrik 2014:203), e.g., when a policy cannot be imitated perfectly by adopter government. Consider policy P which is the existing policy and policy M which is what policy P mutates into. Let $g(t)$ be number of local governments that have implemented P (i.e., frequency of policy P) and $h(t)$ be the number of local governments that have implemented policy M. Every local government has chosen either P or M, such that $g(t) + h(t) = 1$. The mutation rate from P to M is μ. Thus, the system of equations that represents the growth of the two policies is:

$$\dot{g} = g(1 - \mu) - \varphi g, \tag{B.16}$$

$$\dot{h} = h\mu + h - \varphi h. \tag{B.5}$$

Assume both P and M have the same fitness, $p = m = 1$ and average fitness, φ, is also $\varphi = 1$. Given $g(t) + h(t) = 1$, we can simplify the system of equations to:

$$\dot{g} = -g\mu \tag{B.6}$$

Thus, the frequency of policy N declines over time as:

$$g(t) = g_0 e^{-\mu t} \quad . \tag{B.7}$$

The frequency of B increases as:

$$h(t) = 1 - (1 - h_0)e^{-\mu t}. \tag{B.17}$$

Since this shows mutation only from N to M, and not the other way around, policy N will eventually be phased out and M will become the policy of choice.

Chapter 4

Conclusion

4.1 Summary and Conclusion

This thesis shows that the economics of organisation can add to the political economy study of the boundary of the state and of the internal organisation of government. It begins by investigating the problem of the economic organisation of the public sector, and, in particular, two related issues that are central to the thesis: the boundary of the state and the internal organisation of government.

Chapter 2 considers the boundary of the state as a parallel problem to the boundary of the firm (cf. Wittman 1991 and Bolton, Roland and Spolaore 1996). To do so it takes a broader view of the state in which the state is bounded in terms of social space rather than physical space. By doing so, it builds on the North, Wallis and Weingast (2009) framework in which the state is seen as an organisation of organisations. The framework is extended to include the formation of new natural states from existing ones via internal exit.

Specifically, Chapter 2 uses a model to economically capture the birth of a new state through boundary change. It extends Buchanan and Faith's (1987) theory of internal exit by providing a model in which internal exit is initiated by public revenue surplus sharing elite members. While Buchanan and Faith's theory predicts an internal exit-proof tax rate, as this thesis shows, in precolonial southern Africa (ca. 1600-1910), internal exit is not always preventable. In fact, internal exit has left a legacy in the form of the Ndebele people who originated from the Ndebele state that started when Mzilikazi (together with people loyal to him) left the Zulu 19th century formed his own, new state.

The historically informed model in Chapter 2 provides a political economy explanation of why internal exit took place in precolonial southern Africa. The model shows how internal exit results from the payoff calculation of an elite member's (e.g., Mzilikazi) desire to maximise his share of public revenue surplus. As in the political economy literature (e.g., Friedman 1977; Wittman 1991; Findlay 1996; Alesina and Spolaore 1997, 2005) Chapter 2 provides a boundary condition for the state. However, because of the unique African conditions (e.g., the institution of birthright), the boundary condition concerns the size of the elite and the ability to raise a public revenue surplus rather than other factors more commonly emphasised in the literature (costs of public goods, political costs of heterogenous populations, etc.). The model shows that an African state boundary is determined by the ability of the ruler to provide each elite member with a sufficient public revenue surplus share (i.e., a certainty equivalent of exit) to remove the incentive to exit.

Chapter 3 considers the internal organisation of government in the context of the SGT of fiscal federalism. It submits that an intergovernmental grant can be viewed as the public economics counterpart to an incomplete contract (Brennan and Pincus 1990). More specifically, it provides a role for intergovernmental grants in the context of laboratory federalism. Very little theory has been produced that ties policy experimentation with grants. By taking an organisational view of intergovernmental grants by likening them to incomplete contracts, it shows that grants can be instruments for the creation and discovery of new knowledge in the public sector (cf. Garzarelli 2006).

Accordingly, Chapter 3 introduces a model to show how intergovernmental grants can be used to stimulate local policy experimentation. It is an evolutionary learning model that captures the knowledge gains that different types of grants-as-contracts can engender. The model proxies the undertaking of different policy experiments by the local government representative using an intergovernmental grant with melioration (Herrnstein and Prelec

1991; Herrnstein et al. 1993; Metcalfe 2001). Melioration is able to capture the trial-and-error process of local government learning. Through experimentation with an intergovernmental grant, the local government representative, behaving as a meliorator, engages in different policy experiments, ultimately allocating more of the grant funding to the policy experiment that proves most successful. Moreover, as a meliorator, the local government representative will also put a stop to any policy experiments that prove unsuccessful.

Ultimately, the model shows that a lump-sum grant allows a local government representative more freedom to experiment than a closed matching grant. Moreover, the model shows that the conditions attached to the closed matching grant can actually crowd out potential policy experimentation.

4.2 Areas for Future Research

As this thesis is largely theoretical, even though historically and factually well-grounded, an area for future research is to investigate the empirical implications of the theoretical results. In terms of the boundaries of precolonial southern African states it will be useful to corroborate the results of this model with any available data. This may prove challenging because, as we saw, the states under consideration are preliterate.

As for the implications of Chapter 3, it will be useful to test the efficacy of different types of intergovernmental grants in promoting local government experimentation. Specifically, it would be useful to test empirically whether a lump-sum grant induces greater experimentation by the grant recipient than a closed matching grant.

4.3 References

Alesina, Alberto and Enrico Spolaore 1997. "On the Number and Size of Nations," *Quarterly Journal of Economics* 112(4): 1027-1056 (November).

Alesina, Alberto and Enrico Spolaore 2005. *The Size of Nations* Cambridge, Massachusetts: MIT Press.

Bolton, Patrick, Gérard Roland, and Enrico Spolaore 1996. "Economic Theories of the Break-up and Integration of Nations," *European Economic Review* 40(3-5):697-705(April).

Brennan, Geoffrey and Jonathan Pincus 1990. "An Implicit Contract Theory of Intergovernmental Grants," *Publius: The Journal of Federalism* 29(4): 129-44(Fall).

Buchanan, James M. and Roger L. Faith 1987. "Secession and the Limits of Taxation: Toward a Theory of Internal Exit," *American Economic Review* 77(5): 1023-1031 (December).

Findlay, Ronald 1996. "Towards a Model of Territorial Expansion and the Limits of Empire," Chapter 3 in Michelle R. Garfinkel and Stergios Skaperkas (Eds.) *The Political Economy of Conflict and Appropriation* New York: Cambridge University Press.

Friedman, David 1977. "A Theory of the Size and Shape of Nations," *Journal of Political Economy* 85(1): 59-77(February).

Garzarelli, Giampaolo 2006. "Cognition, Incentives, and Public Governance: Laboratory Federalism from the Organizational Viewpoint," *Public Finance Review* 34(3): 235-257(May).

Herrnstein, Richard J. and Drazen Prelec 1991. "Melioration: A Theory of Distributed Choice," *Journal of Economic Perspectives* 5(3): 137-156(Summer).

Herrnstein, Richard J., George F. Loewenstein, Drazen Prelec and William Vaughan Jr. 1993. "Utility Maximization and Melioration: Internalities in Individual Choice," *Journal of Behavioral Decision Making* 6(3):149-185(September).

Metcalfe, J.S 2001. "Consumption, Preferences, and the Evolutionary Agenda," Journal of Evolutionary Economics 11(1):37-58(January).

North, Douglass C., John Joseph Wallis and Barry R. Weingast 2009. *Violence and Social Orders* New York: Cambridge University Press.

Wittman, Donald 1991. "Nations and States: Mergers and Acquisitions; Dissolutions and Divorce," *American Economic Review* 81(2):126-129(May).