



University of Pennsylvania
ScholarlyCommons

Publicly Accessible Penn Dissertations

2020

Informal Institutions And Development

Raj Patel
University of Pennsylvania

Follow this and additional works at: <https://repository.upenn.edu/edissertations>



Part of the [Philosophy of Science Commons](#)

Recommended Citation

Patel, Raj, "Informal Institutions And Development" (2020). *Publicly Accessible Penn Dissertations*. 3892.
<https://repository.upenn.edu/edissertations/3892>

This paper is posted at ScholarlyCommons. <https://repository.upenn.edu/edissertations/3892>
For more information, please contact repository@pobox.upenn.edu.

Informal Institutions And Development

Abstract

Why are some countries wealthier than others? Recent scholarship on this question has begun to converge on the answer to this question: institutions. Some countries are wealthier than others, on this view, because they have the right institutions. These are a set of well-functioning institutions that allow individuals in society to cooperate and coordinate with each other, and, in the end, secure the welfare gains from that cooperation and coordination.

Institutions matter, in other words. But which institutions? In this dissertation, I argue that it is the presence of particular informal institutions (such as social norms) in some societies, and their absence in others, that ultimately decides whether cooperation and coordination is achieved, and welfare gains are secured. I argue that the historical focus on formal institutions (such as formal laws and regulations) is somewhat misplaced, since whether or not people abide by formal institutions is determined by the presence of particular informal rules and norms that govern everyday interaction. In a nutshell: formal institutions matter, but informal institutions rule.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Philosophy

First Advisor

Cristina Bicchieri

Keywords

Corruption, Economic development, Informal institutions, Philosophy of social science, Political philosophy

Subject Categories

Philosophy | Philosophy of Science

INFORMAL INSTITUTIONS AND DEVELOPMENT

Raj Navanit Patel

A DISSERTATION

in

Philosophy

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2020

Supervisor of Dissertation

Cristina Bicchieri

S. J. Patterson Harvie Professor of Social Thought and Comparative Ethics, Professor of
Philosophy and Psychology

Graduate Group Chairperson

Errol Lord, Associate Professor of Philosophy

Dissertation Committee

Samuel Freeman, Avalon Professor of the Humanities, Professor of Philosophy and of Law
Kok-Chor Tan, Professor of Philosophy

For Gita, Navanit, Falgoon, and Aruna

તમારા હાથના કામથી

તમારા હૃદય ના પ્રેમ થી

અહીંના શબ્દો તમારા છે

Acknowledgements

My greatest intellectual and professional debt is to Cristina Bicchieri, my advisor and mentor, who has supported me academically, professionally, and personally from the very beginning. Cristina told me years ago that I needed to look at the world to answer the questions I was interested in. And so I went looking—all over the world.

I owe a great debt to my other committee members, Samuel Freeman and Kok-Chor Tan. Samuel taught me the value of liberalism, lessons that I hope shine through this work and my future philosophical endeavors. KC taught me how to think “globally” about philosophical questions. Taking his class on global justice in my first year in graduate school was indispensable to my intellectual development.

I also owe a debt to teachers and mentors outside of my committee. At Penn, Michael Weisberg, Quayshawn Spencer, Karen Detlefsen, and Lisa Miracchi offered me precious support and advice. At Cambridge, Anna Alexandrova and Marion Godman provided invaluable feedback on my work—they taught me how to think and write “like a philosopher.”

I also show gratitude to my professional colleagues outside of academia. At Chatham House, I owe a special thanks to Leena Koni Hoffman, Elizabeth Donnelly, and Tighisti Amare. At the Nigerian Bureau of Statistics, I owe a special thanks to Shamsudeen Lawal. They have always been supportive of me, both intellectually and professionally, and some of the central ideas of this dissertation were developed during our travels and our work, from London to Abuja, to Lagos, to Addis, and everywhere in between.

I also owe a special debt to friends at Penn. Dylan Manson saved my life. (And this dissertation, after a particularly significant conversation in southern New Mexico.) Neither I nor this dissertation would be here were it not for him. Pierce Randall suffered through hundreds of emails, phone calls, and conversations over the years, and will probably do so for many years to come. And Justin Bernstein was always there for me, from my first year as a graduate student until now. They have my profound gratitude for their intellectual, professional, and personal support.

Other friends, colleagues, and teachers have been supportive of me through the years as well: Andrea Alvaradou-Urbina, Elizabeth Argeris, Ben Baker, Jenny Bulstrode, Chetan Cetty, Jarvis Fisher, All Lauren Forbes, William Griffith, Ashley Halkett, Seiko Iyama, Karen Kovaka, Max Lewis, Akilah Martinez, Peter McNally, Tenzin Metoh, Juan Perez, Aliya Rao, Justin Reed, Brian Reese, Carlos Santana, Flavia Vehbiu, and Mingjun Zhang. They all have my profound gratitude.

Finally, I thank my family. My brother, Falgoon Patel, has worked tirelessly to support me over the years, and has always brought me joy, hope, and inspiration whenever I was doubting myself or feeling down. I would not have gotten through this without him.

I owe my mother, Gita Patel, and my father, Navanit Patel, everything. It is hard for me to imagine how difficult it must have been for them to leave India and Uganda, to leave their home and their world, and travel across three continents in search of a better life. They suffered through hardship so I would never have to.

મારું કાર્ય તમારા માટે છે – કાયમ અને હંમેશાં.

ABSTRACT

INFORMAL INSTITUTIONS AND DEVELOPMENT

Raj Navanit Patel

Cristina Bicchieri

Why are some countries wealthier than others? Recent scholarship on this question has begun to converge on the answer to this question: *institutions*. Some countries are wealthier than others, on this view, because they have the right institutions. These are a set of well-functioning institutions that allow individuals in society to cooperate and coordinate with each other, and, in the end, secure the welfare gains from that cooperation and coordination.

Institutions matter, in other words. But *which* institutions? In this dissertation, I argue that it is the presence of particular informal institutions (such as social norms) in some societies, and their absence in others, that ultimately decides whether cooperation and coordination is achieved, and welfare gains are secured. I argue that the historical focus on formal institutions (such as formal laws and regulations) is somewhat misplaced, since whether or not people abide by formal institutions is determined by the presence of particular informal rules and norms that govern everyday interaction. In a nutshell: formal institutions matter, but informal institutions rule.

In Chapter 1, *Institutions, development, and corruption: the Clientelist Trap*, I provide a “macro-level” view of the role that informal institutions play in generating welfare-relevant outcomes in development contexts. Specifically, I argue, following Douglas North, that institutions play a critical role lowering transaction costs and facilitate cooperation. Lower transaction costs allow for advanced, specialized economic markets that ultimately generate wealth. I then develop a neo-institutionalist account of corruption and apply it to corrupt political markets, specifically, to clientelist politics. I sketch a model that shows the conditions under which clientelism is a stable political strategy for both corrupt patrons (elite politicians) and their clients (non-elite voters).

In Chapter 2, *Institutions out of thin air? A classical liberal case against charter cities*, I turn to a “meso-level” view of institutions in a development context. If institutions are so central to development, why not create discrete political units—charter cities—to compete with places that have “bad institutions”? Proponents of charter cities make this argument and claim that building a place like Singapore on the coast of a place like Somalia is the solution to a number of development problems (such as corruption). In this Chapter, I argue against charter cities from a broadly classical liberal view. Specifically, I claim that there are reasons to think that charter cities are not a morally justifiable form of government by the classical liberal’s lights, even if some form of the charter city (i.e., a private city) is an ideal form of libertarian government. I also claim that much of the pro-charter city arguments rest on a set of empirical claims that fuel pragmatic considerations, but these empirical claims are either false or they actually give us reason to question the very workability of charter cities.

In Chapter 3, *The natural and the social in the metrics of justice*, I turn to a “micro-level” view of the role that institutions play in mediating individual differences in a way that is relevant for liberal egalitarian theories of justice. Specifically, I turn my attention to the debate about the status of so-called “natural inequalities” between Rawlsian resourcists and capability theorists. I argue that capability theorists are able to accept both that distributions of natural endowments are neither just nor unjust while also maintaining that a theory of justice’s metric ought to be sensitive to individual variation (even when this variation has, at bottom, a “natural” (innate, biological) component). This is because capability theorists are able to adopt a more sophisticated understanding of the natural/social distinction, one that stresses the *interaction* between the two through the mediating role of both formal and informal institutions. What lies at the heart of this understanding is a particular conceptualization of the nature of disability, namely, the interaction model.

Table of Contents

Chapter 1. Institutions, development, and corruption: the Clientelist

Trap	1
Introduction.....	1
Collective action theory and corruption: a theoretical overview.....	3
Institutions, development, and transaction costs.....	3
Transaction costs and trust	6
Clientelism and vote markets.....	10
Clientelism: a definition	10
Vote markets beyond the principal-agent problem	13
The Clientelist Trap: contracting for votes in Nigeria	16
Why not programmatic politics?.....	16
Why non programmatic politics?	19
Conclusion	25

Chapter 2. Institutions out of thin air? A classical liberal case against charter cities

Introduction.....	27
What are charter cities?.....	29
Libertarianism and classical liberalism	32
Hobbes and the alienation contract	32
Liberalism, rights, and the separateness of persons.....	34
Nozick and the ultraminimal state.....	37
Classical liberalism, libertarianism, and the charter city	39
The colonialism objection	46
The wrongs of colonialism?	47
Consent	48
Domination.....	50
Pragmatic arguments	51
Informal institutions matter	51
Conclusion	54

Chapter 3. The natural and the social in the metrics of justice.....

Introduction.....	55
Resourcism and the capabilities approach	57
Purely personal heterogeneities.....	62
Compensation claims for natural inequalities.....	63
The status of natural inequalities.....	66
The natural and the social: the interaction model.....	66
Interaction, institutions, and gender	72
Measuring global justice.....	73
Conclusion	74

Appendix 1: Methodology (measuring informal institutions)

Bibliography.....

CHAPTER 1. Institutions, development, and corruption: the Clientelist Trap

1. Introduction
2. Collective action theory and corruption: a theoretical overview
 - a. Institutions and development
 - b. Transaction costs and trust
3. Clientelism and vote markets
 - a. Clientelism: a definition
 - b. Vote markets beyond the principal-agent problem
4. The Clientelist Trap: contracting for votes
 - a. Why not programmatic politics?
 - b. Why non-programmatic politics?
5. Conclusion

1. Introduction

How does Nigeria “get” to Denmark?¹ That is to ask, how does a society characterized by low honesty and high corruption transform itself into a society characterized by high honesty and low corruption? This is the central question of anti-corruption policy, and indeed, the central question of development policy generally. Remarkably, after at least two decades of research and billions of dollars disbursed, we still do not have a satisfactory answer. Anti-corruption success stories are few and far between, and in many cases, countries that have received billions of dollars from the international community for anti-corruption political reform have higher levels of corruption than before they pursued those reforms. It is not much of an exaggeration, then, to say that two decades of anti-corruption policy has been a failure.

Why have these efforts failed? One increasingly influential answer to this question places the blame squarely on the theoretical underpinnings of these efforts, namely, the conceptualization of corruption as a principal-agent problem.² The idea here is that an honest principal (e.g., a public body) contracts with an agent (e.g., a politician or a bureaucrat) who then abuses their public power for personal gain. The key assumptions are that the agent’s interests (in a utility-maximizing sense) diverge from the principal’s interests, and that there is an information asymmetry between the principal and the agent. What determines whether the agent will engage

¹ To use Fukuyama’s turn of phrase: “The international community would like to turn Afghanistan, Somalia, Libya, and Haiti into idealized places like ‘Denmark,’ but it doesn’t have the slightest idea of how to bring this about.” The problem of development, according to Fukuyama, is the problem of “getting to Denmark,” where Denmark is a generic stand-in for an “imagined society that is prosperous, democratic, secure, and well governed, and experiences low levels of corruption” (Fukuyama 2014, p. 26).

² Persson et al. (2013, 2019); Rothstein (2005, 2018); Mungiu-Pippidi 2006, 2013, 2019).

in corruption is the cost and benefit of the action: if the benefits of engaging in corruption outweigh the costs, then the agent will engage in corruption. On a particular theoretical understanding of this view, preferences are non-social and exogenous (i.e., the model assumes agents will maximize self-interest independently of what other agents do) and the analytic focus is on formal institutions (the legal rules, penal sanctions, and so on) as opposed to informal institutions (norms, expectations, and so on); the kind of anti-corruption reforms that flow from this framework aim at increasing the costs of engaging in corruption through formal rules and mechanisms, such as increased monitoring and punishment.

But in cases of systemic corruption, corrupt behaviour is not the result of an independent and individual preference to maximize self-interest in the face of formal rules. Instead, in many cases, the best way to understand corruption is as a behaviour supported by a set of shared beliefs and expectations, where social preferences to engage in behaviour are dependent on the expected and actual actions of others. The underlying thought here is that the costs of engaging in corruption are high if one expects nobody else to engage in corruption, but the costs of *not* engaging in corruption are high if one expects everybody else to engage in corruption. From this view, preferences are endogenous, and the key analytic focus is on both the formal institutions as well as the informal institutions that drive the practice. The bottom line is that anti-corruption policy must be sensitive to both the formal and informal sets of rules that structure the incentives to engage or not engage in corruption, as well as the interaction of these formal and informal rules. The broader theoretical current behind this approach is best described as rational choice neo-institutionalism.³

A point that is often lost in the more polemic arguments against the principal-agent approach to corruption is the idea that principal-agent models are, of course, naturally compatible with a broader collective action approach. Principal-agent relationships are almost definitionally integral to the analysis of corruption since the most basic understanding of corruption involves the misuse of public power for private gain—a relationship that implicates a principal-agent relationship at its very core. The arguments in this dissertation align with this general sentiment. Instead of aiming to displace principal-agent approaches to corruption, I show that a sophisticated understanding of corrupt behaviour implicates both principal-agent theory and collective action theory, and indeed, the two approaches are complimentary.

In this chapter, I develop and defend a rational choice neo-institutional account that specifies the kind of informal institutions that drive corrupt practices.⁴ In §2, I clarify the

³ I take the rational choice neo-institutionalist intellectual lineage to stretch back at least to Adam Smith, and to be firmly a part of what Boettke, Haefelle-Balch, and Storr (2016) refer to as ‘mainline economics,’ which accepts the idea that ‘formal and informal institutions guide and direct human activity’ as a key premise (p. 4). For more contemporary exemplars of neo-institutionalist thought, especially with respect to political and economic development, see North (1981, 1987, 1990, 1991, 2005); North, Wallis and Weingast (2010); Ostrom (1990, 2000, 2005, 2009, 2010, 2014); Acemoglu, Johnson & Robinson (2001); Acemoglu and Robinson (2012, 2019).

⁴ The account in this chapter also informs and underpins the theoretical and empirical understanding of institutions in the rest of the dissertation.

empirical and theoretical connections between the account and economic and political development, with a special emphasis on the mediating role institutions play through transaction costs and trust. In §3, I apply my general theoretical account to political clientelism. Here, I provide my own definition of clientelism, one that requires us to look at the formal and informal institutions that structure the political relationships that characterize clientelist politics (which are, of course, ubiquitous in developing societies). In §4, I provide a theoretical model of clientelism that I call the “Clientelist Trap.” The Clientelist Trap draws attention to a set of circumstances whereby a clientelist equilibrium stabilizes as an efficient political strategy because of voter preferences for clientelism. I argue that voters prefer clientelism because it is often the most attractive option in their option set, given credibility and commitment problems that plague other, more seemingly better options. I briefly present fieldwork evidence from Nigeria to provide empirical content to my theoretical framework (see Appendix 1).⁵ As I argue in §3 and §4, the Clientelist Trap is a theoretical and empirical application of the general neo-institutionalist account developed in the chapter, and it is one that helps us make sense of some of the central empirical puzzles that animate much of the rational choice literature on political clientelism. In §5, I conclude.

2. Collective action theory and corruption: a theoretical overview

a. Institutions, development, and transaction costs

Neo-institutionalism rests on the idea that the pursuit of individual self-interest, under the appropriate institutional framework, generates net social benefits to society as a whole. It is the existence of these institutions in some societies and not in others that provides the dominant explanation of patterns of comparative development; in other words, the differences in prosperity between societies seen around the world today is best explained, in the final analysis, with reference to the sorts of institutions that one finds in those societies. The diagnosis of developing societies and the problem of development in general is thus framed, on this view, in terms of a lack of an appropriate institutional framework that aligns individual self-interest with the interests of the broader society in which those individuals live. The difficulty in securing the right institutional framework for any society is that a state that has the capacity to deliver even on the most basic institutions (i.e., to provide a minimum of law and order, property rights, enforcements of other contracts, and so on) is also a state strong enough to predate on the efforts of its citizens. “Getting to Denmark,” then, really means getting to a strong, competent state that is sufficiently constrained so as to not interfere with the security of property and other underlying market-relevant institutions.

⁵ I note also that the general theoretical account I develop in this chapter compliments fieldwork I have undertaken and published elsewhere (Patel and Hoffman, 2017).

The “right” institutions align private self-interest and social benefit and thus leave everyone better off; thus we would expect a demand for the right institutions. *Almost all* individuals in the developing world would prefer to live in a society with the right institutions, i.e., those under which honesty was rewarded and dishonesty was punished and where one could be reasonably confident that others would generally act honestly. But a well-functioning society that provides the preconditions of prosperity is, on this view, the ultimate public good and thus suffers from the same problems that afflict the provision of any public good. Individuals stand to benefit from the provision of a well-functioning society without having to contribute to it in the same way they stand to benefit from any other public good without having to contribute to it. And thus the ultimate public good of a successful society that facilitates cooperation is predictably underprovided.⁶ Developing societies thus face a second-order collective action problem with respect to the supply of institutions; it is precisely the inability to overcome this collective action problem that explains the pervasive market failures that cause welfare losses associated with those societies.⁷

At the theoretical core of neo-institutionalism as it relates to corruption is an understanding of *institutions*.⁸ There are two important, related approaches to mention here. First, the *institutions-as-rules* approach, elucidated in a body of work beginning with Douglas North and Elinor Ostrom,⁹ views institutions-as-rules, where institutions are considered to be the “rules of the game in a society.”¹⁰ Formal rules, such as laws and legal regulations, are conscientiously decided upon by the relevant political community, and the extent to which these rules can be designed so as to minimize transaction costs is to the extent to which the “Smithian result” of productive markets will obtain. Both North and Ostrom provide a hierarchy of rules, such as operational rules that govern the day-to-day, collective-choice rules which determine operational rules, constitutional rules which determine collective-choice rules, and meta-constitutional rules which determine constitutional rules, and at the very top, the rules of the biophysical world—rules that are beyond human design.

Much of the institutions-as-rules literature focuses on conscious political rule choice; if individuals are not satisfied with the rules governing interactions on one level of the hierarchy, they “move up” the levels to change rules at the higher level. But there is no reason to think choice of rules will be efficient or beneficial for everyone or even a majority, since the set of rules that “ultimately emerges will depend on the perceived interests of actors involved in setting the rules.”¹¹ In most cases, these interests will be the those of the powerful who are able to set, and crucially, enforce the rules. These interests will often diverge from the general interest, even if

⁶ Keefer (2018, p. 10).

⁷ Bates (1988); Ostrom (1990).

⁸ For a general overview, see Greif & Kingston (2011) and Hindriks & Guala (2015).

⁹ North (1981, 1987, 1990, 1991, 2005); North, Wallis and Weingast (2010); Ostrom (1990, 2000, 2005, 2009, 2010, 2014).

¹⁰ North (1991).

¹¹ Avner & Greif (2011, p. 16).

more efficient rules would leave everyone better off. Thus, institutional rule-choice is plagued by a dual commitment and credibility problem, where powerful elite interests cannot commit to non-abuse of power for their own private interests, and other, less powerful groups cannot commit to compensating the powerful for the loss of power that would come along with more efficient institutions. This is an insight that stretches back to Adam Smith, who explained the persistence of inefficient institutions such as slavery in exactly these terms.¹² The commitment and credibility problems that plague developing societies is an issue that I return to in §4.

The second, related strand of literature on institutions focuses on *institutions-as-equilibria*.¹³ Where the institutions-as-rules approach emphasizes institutional choice, institutions-as-equilibria emphasizes an underlying theory of motivation which generates equilibrium-stabilization pressures that result in particular behavioral regularities. From this view, what induces behavior is not the overarching formal rules given by the state, but rather social expectations on what others do, are expected to do, and what others expect us to do (empirical and normative expectations in Bicchieri's terminology).¹⁴ Behavioral regularities are often best thought of as stabilized through *informal* institutions, in these terms, for there are often an attendant set of "invisible," informal rules that motivate individuals to comply with a behavioral rule or bear the costs of deviating from that rule. Since deviation is costly relative to compliance, these informal rules generate an equilibrium that is the result of a "structure that motivates each individual to follow a regularity of behaviour in that social situation and to act in a manner contributing to the perpetuation of that structure."¹⁵ The problem of corruption from the institutions-as-equilibria view is not the absence or presence of formal anti-corruption rules, but rather the self-sustaining presence or absence of particular informal rules and the behavioral regularities they give rise to.¹⁶ It is the presence self-sustaining of "bad" informal rules, i.e., rules that make honesty costly relative to dishonesty, that forms the core of the collective action problem at the heart of the problem of development.

I note here that these two approaches are often framed as in tension with each other despite some recent unificationist accounts.¹⁷ I do not aim to settle the debate here. Suffice to say that the two approaches complement each other given that institutions-as-rules focuses on institutional choice, and institution-as-equilibria focuses on the theory of motivation that generates stabilization pressures that reproduce institutions. More importantly, both approaches point to an influential line of critique of the principal-agent approach to corruption which holds that the problem with the principal-agent view is that it focuses on the individual when the problem is often institutional (or systemic). Diagnosing corruption, on this view, requires us to use the theoretical tools of neo-institutionalism and collective action theory rather than principal-

¹² For a good overview on Smith's arguments on the persistence of inefficient institutions, see Weingast (2016).

¹³ Avner & Greif (2011, p. 25).

¹⁴ Bicchieri (2006).

¹⁵ Avner & Greif (2011, p. 25).

¹⁶ Ibid.

¹⁷ Hindriks & Guala (2015).

agent theory alone, since endemic corruption is best viewed a “malicious” equilibrium sustained by interdependency of beliefs about whether others will engage in corruption. Collective action theory, in other words, focuses on the institutions which structure interactions between the principals and the agents—both formal and informal.

b. Transaction costs and trust

While the “post-Washington Consensus”¹⁸ view amongst the development community (and the underlying social sciences) has come to accept the core neo-institutionalist claim that it is the institutions of a society that will determine its overall economic growth, the fundamental emphasis on institutions is not at all new. Indeed, the idea that the relevant institutions are necessary for the functioning of markets, and crucially for the coordination and efficiency gains of markets to obtain (with their attendant welfare benefits), stretches back to at least Adam Smith. In the notes of lectures written some 21 years before the publication of *Wealth of Nations* (1776), Smith wrote that “little else is requisite to carry a state to the highest degree of opulence from the lowest barbarism, but peace, and a tolerable administration of justice.”¹⁹ Obtaining the “Smithian result,” to use North’s turn of phrase, requires the appropriate formal and informal institutions.²⁰

The social sciences—most remarkably economics—seems to have forgotten this basic insight as neoclassical economic theory, with its emphasis on technical analysis, rose to enjoy status as academic orthodoxy from the 1930s onwards.²¹ Indeed, institutions do not feature at all in neoclassical economics (embodied in welfare economics and the Arrow-Debreu general equilibrium models) because of a set of assumptions that the standard models make about the framework in which economic exchange takes place. As Boettke puts it, the “formalistic rendering of economic reasoning in the 1930 –1960s strove to be *institutionally antiseptic*” (his emphasis).²² Crucially, the models assumed what Weingast has dubbed “the neoclassical fallacy,” which is the view that (1) agents performing economic transactions enjoy perfect security; (2) a system of well-defined property rights with complete contract enforcement exists; and that (3) there is no arbitrary state (or any other actor with a monopoly on violence) to predate on productive economic exchanges.²³ In other words, neoclassical economics assumes that the transaction costs involved in economic exchange are zero.

Transaction costs are crucially dependent on the institutional framework in which they accrue. This insight comes from Coase’s seminal work in which he “opened the black box”²⁴ of the firm and asked why they exist as a form of governing hierarchy in market settings. His answer:

¹⁸ Burki & Perry (1998).

¹⁹ Smith (1976).

²⁰ North (2005, p. 80).

²¹ Boettke (2018).

²² *Ibid.*, (p. 161).

²³ Weingast (2016, p. 192).

²⁴ Coase (1937, 1960); Hodgson (2005).

Since the costs of finding potential trading partners, communicating with them, negotiating contracts, monitoring performance and ensuring enforcement are lower in a single hierarchical organization than they would be in a decentralized market setting, economic actors organize themselves in ways that lower these costs and thus allow them to realize higher economic gains from economic exchange. Following this literature, I define transaction costs as the sum of coordination costs (i.e., the costs involved in search, communication, negotiation, and monitoring) and transaction risks (i.e., the costs involved in enforcement including losses incurred due to non-enforcement):

$$\text{Transaction costs} = \text{coordination costs} + \text{transaction risks.}^{25}$$

A subset of transaction costs is closely related to the problem of trust.²⁶ These are generally the costs involved in transaction risks. Transaction risks become a problem for modern complex economies that are characterized by a division of labor and the attendant high levels of productivity and economic growth because trade is often conducted amongst strangers as opposed to close kin or family. These economies usually feature transactions that are governed by impersonal rules of exchange where there is a time lapse between when a promise is made and when it is executed.

These transactions have the structure of a one-sided prisoner's dilemma (or a Trust Game). One sided prisoner's dilemmas are games where the payoff structures for the players are asymmetric: Player A (trustor) has the option to cooperate (to invest or not invest). If Player A cooperates (invests), player B (trustee) has the option to cooperate (not exploit) or not cooperate (exploit). Player A's optimal choice is to not invest, and Player B's optimal strategy is to exploit. The result is a suboptimal Nash equilibrium that is Pareto-inferior where the gains from mutually beneficial exchange go unrealized.²⁷ This is one way to think about the effects of high transaction costs in a given economy: Player A's "fear" of being exploited by Player B (and the associated social losses) represent a social dilemma; high transaction costs represent the costs of overcoming that social dilemma in any given economy.²⁸

This Trust Game indicates the importance of *generalized* trust in resolving the social dilemma that Player A and Player B find themselves in. Generalized trust is the trust that individuals have in anonymous others or in impersonal formal institutions. How would trust emerge in such a situation such that Player A and Player B can realize mutual gains? Consider games that allow for long-term repeated interactions between two players (Repeated Trust Games). These games are structured to study the effects of "dyadic embeddedness," which allow for two mechanisms through which trustfulness and trustworthiness can be induced: Player A can

²⁵ Clements et al. (1993).

²⁶ Beccerra & Gupta (1999).

²⁷ Greif (1994, p. 918).

²⁸ Buskens (1998, 2002); Buskens & Raub (2002, 2013).

learn (from the past behaviour of Player B) and *control* (by withholding trust in the case that Player B decides to exploit during the games).²⁹ The experimental results from these games provide a somewhat intuitive result: Player A's decision to trust is largely conditional on whether Player B has decided to not exploit in previous iterations of the game.³⁰ With a long enough time-horizon, Player B has incentives to build a reputation in order to secure investments from Player A.³¹

A further variant of Repeated Trust Games provides Player A information about Player B's interactions with other trustors. These games are structured to study the effects of "network embeddedness."³² Games of this sort will typically provide Player A with a reputation score about Player B as well as allow for Player A to provide feedback (which in turn will have an effect on Player B's reputation score). The general experimental findings show that network embeddedness tends to promote trustfulness and trustworthiness.³³ Interestingly, a recent study found that the size of the effect is stronger when embeddedness is chosen endogenously (i.e., players in the game invest in an information mechanism that provides a reputation score or its equivalent) than when it is provided exogenously.³⁴ These experimental findings provide clues for mechanisms behind the purported links between social capital and economic growth. Many of the specific risks facing any given individual (or firm) stem from information asymmetries between the parties to the transaction. These generally result in adverse selection problems before the execution (precontractual opportunism) and principal-agent problems after the execution (postcontractual opportunism).

Consider an example of a simple bilateral employment contractual relationship. Contractual relationships between employers and employees typically involve information asymmetries of the following sort. Potential employees know their abilities to perform on a particular job better than potential employers do. Moreover, employees know more about their performance on the job than their employers do. Firms thus face transaction risks from potential employees misrepresenting their abilities (precontractual opportunism as an adverse selection problem) or shirking while on the job (postcontractual opportunism as a moral hazard problem). Firms are generally aware of such risks and tend to invest resources to minimize their exposure (such as screening mechanisms to shield against precontractual opportunism and monitoring mechanisms to shield against postcontractual opportunism).

The particular attributes of the good or service at the heart of the transaction might add a further level of transaction risk (and thus attenuation cost). For example, consider the 'lock-in

²⁹ Buskens & Raub (2002).

³⁰ Ibid.

³¹ Ibid.

³² Buskens & Raub (2013).

³³ Ibid.

³⁴ Frey et al. (2019). Note that the choice to invest in an information mechanism that promotes embeddedness in this context is a voluntary contribution toward a public good (since the information mechanism promotes trustfulness and trustworthiness which benefits all).

effect,' which refers to a contractual arrangement where one party to the contract has more of an interest in its eventual realization than the other. The lock-in effect is related to the 'hold up problem,' which refers to the possibility for opportunistic behaviour created by the asymmetry of interests between the contracting parties. Contracts involving highly "transaction-specific" assets will be particularly susceptible to lock-in effects and hold up problems because asset specificity usually means the asset's value is tied up to the particular firm or transaction.³⁵ An example of this would be an employee who learns firm-dependent skills, as in the case of an employee who learns how to code software in a language that is used exclusively by a particular firm or a law associate who has specific knowledge of a particular client with a set of idiosyncratic legal issues. Contractual mechanisms to mitigate the risks of opportunism that flow from the lock-in effect include collective bargaining arrangements (e.g., labor unions) or mixed-ownership structures (e.g., law firms).

When generalized trust is low, transaction costs are high because contracting parties must take steps to shield themselves from heightened risks of defection. Low trust societies must divert resources into risk minimization that might be freed up and directed toward more productive uses in a high trust environment. In already resource constrained environments, high transaction costs due to low generalized trust may prohibit markets from developing, producing failures of cooperation and higher costs for collective action. This is one reason why developing societies conspicuously lack sophisticated, advanced markets that require high levels of trust (such as primary and secondary insurance markets). It is precisely the lack of such integral markets that leaves individuals in the developing world, especially the poor, in a fragile and precarious position, often one small setback away from financial ruin.³⁶

Moreover, societies with low generalized trust allow for economic markets that rely on high levels of particularized trust. Indeed, societies with endemic corruption often have high levels of particularized trust, since iterated corrupt transactions rely upon informal rules such as reciprocity, loyalty, and honesty to emerge and structure trust specific to *particular* others. This also generates demand for intermediaries ("middlemen") to ensure that individuals aiming to contract with corrupt individuals will find the "honestly dishonest" party they are looking for.³⁷ These intermediaries are often individuals placed within a network of trust who have information of where honestly dishonest individuals may find and contract with each other.³⁸ As the number of corrupt market participants increases, specialization occurs through "investments in know-how and networks" that produces "positive feedback loops" signaling to others that corruption is both cheap and profitable, and carries low risk of sanction.³⁹ High transaction costs for legitimate contracts also allows for arbitrage opportunities, where, e.g., frontline bureaucrats can raise and

³⁵ Riordan & Williamson (1985).

³⁶ Banerjee & Duflo (2011).

³⁷ Lambsdorff (2007)

³⁸ Ibid.

³⁹ Ibid., 53.

lower transaction costs to perform functions that should be free or outside of their legitimate discretion.⁴⁰ In other words, individuals with power—whether on a large (grand) or small (petty) scale—create islands of low transaction costs to secure selective cooperation and collective action, at the expense of the welfare interests of society at large.

Finally, societies with low generalized trust tend to generate political markets that rely on high levels of particularized trust. These are political arrangements characterized by clientelistic non-programmatic politics,⁴¹ where political elites (patrons) or their representatives (brokers) selectively target particular voters (clients) or their representatives (brokers) with inducements aimed at securing the targeted voters’ political support or raising the probability of securing their political support. I turn now to a more thorough discussion of clientelism.

3. Clientelism and vote markets

a. Clientelism: a definition

Clientelism is notoriously hard to define. Since clientelist relationships are socially embedded in a matrix of informal institutions that structure political and economic life, conceptually delineating between different practices is a key task for scholars investigating distributive politics in the developing world. Moreover, the characteristics of a clientelistic change are prone to ambiguity – sometimes intentionally so – which makes conceptual clarity even more difficult. As Nichter put it: “[S]tudies of clientelism often report survey or fieldwork evidence about the relative prevalence of clientelist vote buying, but it is sometimes unclear what specific attributes of benefits are considered.”⁴²

Despite this “conceptual diversity,” there does seem to be broad agreement on two features that are in most, if not all, accounts: (1) the *contingency requirement*, which is that clients view inducements as conditional on some action or obligation (political support) on their part; and (2) the *iterative requirement*, which is that clientelist exchanges are not one-shot encounters, but interactions where all parties hold the mutual expectations to the effect that transactions will continue into the future.⁴³ With these requirements in mind, I define clientelism as follows:

Clientelism describes governance arrangements where political elites (patrons) or their representatives (brokers) selectively target particular voters (clients) or their representatives (brokers) with inducements aimed at securing the targeted voters’ political support or raising the probability of securing their political support.

⁴⁰ Ibid.

⁴¹ I contrast clientelist non-programmatic politics with non-clientelist programmatic politics.

⁴² Nichter (2014).

⁴³ Hicken & Nathan (2019).

Inducements are (1) iterative, (2) distributed in accordance with informal and non-transparent rules or discretion, and (3) may be positive or negative.

Positive inducement here refers to offers of benefits broadly construed in exchange for electoral support. Positive inducements are distributed by the informal discretion of the political elite, even if what is being distributed privately is in fact public. An example of a positive inducement would be offering a prospective voter a job in order to secure their political support.

Negative inducement here refers to sanctions broadly construed (e.g., the principal threatens the agent with violence to elicit a particular voting outcome) in exchange for electoral support. Negative inducements are imposed by the informal discretion of the political elite, even if the resources or institutions involved are public. An example of a negative inducement would be withholding a welfare entitlement to a prospective voter in order to secure their political support.

My definition aims to capture a central feature of clientelism which is often overlooked, namely, the distinction between positive and negative inducements. This distinction is crucial to understanding vote markets because of three reasons: (1) it allows us to capture a set of unobservable characteristics that structure all vote market transactions and will, in some cases, be essential in explaining or predicting whether a successful transaction is achieved; (2) a natural taxonomy flows from it, providing a richer picture of the possible repertoire of electoral strategies available to the patron, allowing an investigation into the circumstances under which particular mix will be employed (see *Table 1*). This is important as a matter of public policy, especially since negative inducement includes violence at its extreme, and reduction in electoral violence is a clear and important public policy goal; and finally (3) it captures a set of essential differences between markets in which violence is routinely (and successfully) employed strategically and those in which it is not, which has implications for a number of nearby contemporary debates on (a) the moral permissibility of legalized vote markets,⁴⁴ and (b) whether corruption is “efficient” and a “Political Coase Theorem” exists.⁴⁵

⁴⁴ Freiman (2014); Archer & Wilson (2014); Taylor (2016, 2017, 2018).

⁴⁵ Acemoglu (2003); Parisi (2003); Munger (2019). For a brief overview, see Medema (2017, p. 80).

	Types of inducement	
Strategies aim to...	Positive inducements	Negative inducements
Increase the number of votes for a particular candidate	<p>Type (A) Positive inducements in the form of benefits (broadly construed) are distributed in a way that is targeted (at particular voters), contingent (upon electoral support), subject to non-public and non-transparent rules of distribution, and temporally specific (i.e., offered close to actual elections). For example, offering direct cash for individual voters to vote for a particular candidate.</p>	<p>Type (D) Negative inducements in the form of threatened or actual sanctions (broadly construed) in a way that is targeted (at particular voters), contingent (aimed at eliciting electoral support), and temporally specific (i.e., offered close to actual elections). For example, threatening individual voters with violence to induce them to vote for a particular candidate.</p>
Increase voter turnout	<p>Type (B) Positive inducements in the form of benefits (broadly construed) are distributed in a way that is targeted (at particular voters), contingent (upon showing up to the polls on election day), subject to non-public and non-transparent rules of distribution, and temporally specific (i.e., offered close to actual elections). For example, offering direct cash for individual voters to turn up to the polls on election day (without stipulating who to vote for).</p>	<p>Type (E) Negative inducements in the form of threatened or actual sanctions (broadly construed) in a way that is targeted (at particular voters), contingent (aimed at increasing voter turnout), and temporally specific (i.e., offered close to actual elections). For example, threatening individual voters with violence to induce them to turn up to the polls on election day (without stipulating who to vote for).</p>
Decrease voter turnout	<p>Type (C) Positive inducements in the form of benefits (broadly construed) are distributed in a way that is targeted (at particular voters), contingent (upon not showing up to the polls on election day), subject to non-public and non-transparent rules of distribution, and temporally specific (i.e., offered close to actual elections). For example, offering direct cash for individual voters to not turn up to the polls.</p>	<p>Type (F) Negative inducements in the form of threatened or actual sanctions (broadly construed) in a way that is targeted (at particular voters), contingent (aimed at decreasing voter turnout), and temporally specific (i.e., offered close to actual elections). For example, threatening individual voters with violence to induce them not turn up to the polls on election day.</p>

Table 1. Categories of electoral clientelism.

If, for example, withholding a benefit is more likely to induce the required behavior than offering it because of asymmetry in the ways in which voters might perceive the costs and benefits of a thing taken away against its being offered, then knowing whether the voter holds expectations for

that good is critical as well.⁴⁶ This is primarily due to the fact that individuals tend to perceive gains and losses asymmetrically.⁴⁷ This implies that behaviour—in terms of whether inducements are successful and cooperation is achieved—may critically depend on whether inducements are perceived as gains (positive) or losses (negative).

Whether an inducement is considered positive or negative might depend on non-observable features of the particular situation, such as highly localized norms of reciprocity that structure a set of underlying beliefs and expectations on the part of the voter. Consider an example of a politician who offers a potential voter a job in the bureaucracy in exchange for electoral support. If the voter expects the job, threatening to withhold the job would be a negative inducement. If the voter does not expect the job, offering the job would be a positive inducement. It's entirely plausible that politician's offer will elicit a different response from a voter who was expecting the job against one who was not—if, for example, the perceived net loss in utility from the withheld benefit is enough to secure cooperation in the former but the perceived net gain from receiving the benefit is not enough to secure cooperation in the latter.⁴⁸ The fact that the two cases are indistinguishable in terms of the observable features of the situation underscores the claim that the underlying beliefs and expectations are integral to understanding the practice.

b. Vote markets beyond the principal-agent problem

Much of the quantitative rational choice literature on clientelism is centered around a principal-agent problem at the heart of the informal agreement (vote contract) between patron (buyer) and seller (client) in illicit vote markets under conditions where ballot secrecy is maintained.⁴⁹ One core element of clientelist exchange is that the *quid pro quo*: some good is targeted at particular votes in a way that all parties understand as conditional upon the execution of some obligation typically tied to political support. In the case of vote markets, the targeted benefit is exchanged for votes. Patrons as buyers in vote markets face substantially high transaction costs because rules that safeguard ballot secrecy ensure non-simultaneity of exchange and prohibitively high monitoring costs, placing them in a principal-agent problem: it's easy for voters to “take the money and run” since they benefit from the exchange before they discharge their obligation which are done in secret anyway. Patrons must solve the principal-agent problem for clientelism to be an efficient political strategy. On the assumption that resolving the commitment problem is required to make clientelism an efficient political strategy, the existence of vote markets suggests that the patrons succeed in holding voter-sellers “perversely accountable,”⁵⁰ and thus overcoming

⁴⁶ On the distinction between positive and negative inducements, and its implications, see Mares & Young (2016, 2018, 2019); Justesen & Mares (2019).

⁴⁷ Kirstin et al. (2011).

⁴⁸ Mares & Young (2016).

⁴⁹ The typical case of clientelism is where the patron is a member of the political elite or its representative who contracts with a client who is typically a member of the vulnerable poor through a broker who serves as an intermediary.

⁵⁰ Stokes (2005).

the information asymmetries inherent in vote market transactions. Thus, one of the central puzzles of scholarship on clientelism asks how clientelism is an efficient political strategy in light of the commitment problem.

A particularly popular line of research in responding to this puzzle emphasizes the distributive choices patrons make when distributing clientelist goods, with the assumption that patrons will target specific groups to help them overcome the commitment problem. This research is motivated in large part by a gap between a widely held theoretical intuition and a counter-intuitive empirical finding. The theoretical intuition, formalized in a seminal article by Susan Stokes is the expectation that patrons should target voters with weak political preferences in order to maximize electoral chances, since those with strong political preferences will vote favorably without inducement (party loyalists) or will be insensitive to positive inducement (opposition loyalists).⁵¹ The empirical finding is that most clientelist transfers are in fact targeted as party loyalists, raising a second central puzzle that animates scholarship on clientelism: why do patrons target their own loyalists for clientelist goods?

An unfortunate consequence of the emphasis on the commitment problem is that insufficient attention has been paid to the supply side of the interaction, namely, the beliefs, expectations, and motivations of the vote-seller. Instead, much of the analysis of remains fixed on the patron's strategy in terms of distributive choices in relation to the commitment problem. As some scholars have recently noted,⁵² what seems to be driving on the focus on the commitment problem is the assumption that it binds in every transaction and its "solution" is always monitoring. It is worth considering the theoretical conditions that are taken to be constitutive of the orthodox view of vote market transactions:

1. Clientelist transfers involve non-simultaneous exchange of a contingent benefit.
2. Non-simultaneity and contingency entails that all transfers occur under conditions of asymmetric information.

Note that (1) and (2) are equivalent to the conditions under which a principal-agent problem *may* arise. Now consider the following three conditions, which I take to be false:

3. The commitment problem is a constitutive feature of all vote market transactions.
4. The commitment problem "binds" parties with patron as principal and client as agent.
5. The solution to the commitment problem is monitoring compliance.

⁵¹ Ibid.

⁵² Hicken & Nathan (2019).

Conditions (3), (4), and (5) go further and claim that a principal-agent problem *does* arise and is solved in a very specific way. It is these last three claims, (3), (4), and (5), motivate the widespread view that clientelism is a feasible electoral strategy only when the commitment problem is overcome, meaning that a vote-buying equilibrium emerges only when monitoring and enforcement is successful, defection is deterred, and transactions are complete.⁵³ This line of reasoning motivates the centrality of the principal-agent problem and provides a rationale for the tendency to privilege the patron's perspective.

The rationale fails on both theoretical and empirical grounds. Conditions (1) and (2) hold insofar as they can be read as a recognition that information asymmetries are “built into” the transaction in light of (a) the non-simultaneity of exchange and (b) the contingent nature of the benefit. These properties are fixed to the exchange and their presence will generate information asymmetries between buyer and seller along with the potential for ensuing agency problems. But conditions (3), (4), and (5) need not be true, since there are no theoretical reasons to assume that conditions which provide ripe grounds for agency problems will in every case produce them. Moreover, there are a number of empirical studies documenting what look very much like clientelist exchange – with an iterated and contingent *quid pro quo* – without the subsequent monitoring and enforcement activities that we would expect if the commitment problem truly did bind in all cases.⁵⁴ In other words, the rationale fails on the ground that it mistakes the conditions under which a principal-agent problem might arise with the actual existence of a principal-agent problem in every case.

The implication is that clientelism looks like a feasible electoral strategy without the need for complete transactions, and so without the need for post-contractual monitoring. This raises a new puzzle: How it is possible for a clientelist equilibrium to persist *without* post-contractual monitoring and enforcement? Clientelist transfers represent significant expenditures in resource-scarce environments, often involve intermediaries (such as brokers), and require substantial organizational capacity. Without monitoring and enforcement, the substantial room for opportunism suggests high opportunity costs for vote buying relative to other possible electoral strategies. Yet a mounting body of evidence suggests that “clientelism-without-monitoring” is widespread, raising significant questions about the central assumptions of the discipline. I now briefly sketch a model that illustrates the conditions under which clientelism-without-monitoring becomes a viable strategy.

⁵³ Ibid.

⁵⁴ For example, Kramon (2017), Hicken et al. (2019), Mares & Young (2020).

4. The Clientelist Trap: contracting for votes

a. Why not programmatic politics?

To understand how and why clientelism-without-monitoring is a viable political strategy for corrupt politicians, it is first necessary to consider the political options faced by potential voters in societies where political corruption is endemic:

Option set 1: the ideal programmatic set

(a) *A preference for programmatic politics.*

Vote out dishonest candidates and instead vote for an honest candidate through formal political channels.

(b) *A preference for collective action politics.*

Coordinate and act collectively with other potential voters to create their own institutions.

(c) *A preference for clientelism.*

Sell their vote and engage in informal, non-programmatic politics.

(d) *A preference for nothing.*

Remain politically powerless or leave the polity.

A reasonable assumption would be that most voter preferences run $(a) > (b) > (c) > (d)$. That is, most voters would prefer standard programmatic politics where honest politicians deliver public goods in a transparent and accountable fashion. Short of that, voters would prefer to create their own institutions and parties that deliver public goods in a transparent and accountable way. If these two options are not available, then voters would prefer clientelist politics, where political power is viewed as private power and public resources are distributed in accordance with informal and non-transparent rules and elite discretion. And lastly, the worst option of all (from the voter point of view), would be to refrain from the political system altogether through political non-participation or exit.

But this is not the actual option set faced by voters. Rather, it is a highly idealized version of it. Once we include costs associated with exercising each preference, it becomes clear that the real option set consists solely of (c) and (d), i.e., the *real* choice is between clientelism or nothing. So the “real,” non-idealized option set is actually:

Option set 2: the clientelist dilemma

(c) *A preference for clientelism.*

Sell their vote and engage in informal, non-programmatic politics.

(d) *A preference for nothing.*

Remain politically powerless or leave the polity.

The upshot is that clientelism emerges as the most attractive option even if the voter's preference ordering is (a) > (b) > (c) > (d):

(a) *A preference for programmatic politics.* We would expect almost everyone to have a preference for programmatic politics characterized by a political system where distributive rules are fair and transparent so as to ensure fair delivery of a wide range of public goods. There are three reasons why this is not a viable choice in societies where corruption is endemic.

First, when corruption is widespread, potential candidates face a *credibility problem*. As the standard principal agent model above suggests, politicians *qua* agents often have wide discretion after being elected because of prohibitively high monitoring costs faced by citizens *qua* principals. If a conflict of interest exists, which it does in the case that a politician pursues their own self-interest over the public interest, then standard moral hazard agency issues ensue.

Second, public resources that political offices give individuals or parties access to often raise a common pool problem because everyone who has access expects others with access to spoil the resource. In that sense, politicians may face their own somewhat perverse collective action problem vis-à-vis one another in a way that would make even the most honest politicians political promises non-credible. Pressures toward corruption might be at play both pre- and post-election; if they expect other political candidates to buy votes, for example, then they would lose out if they did not themselves buy votes. And once they are in office, if they expect other political elites to loot public resources, they themselves would lose out if they refrained from looting public resources.

Lastly, honest politicians who resist both pre- and post-election corruption pressures still face yet another problem. Shifting a society from high corruption low honesty to high honesty low corruption incurs significant transition costs, since there are often powerful interests who benefit from—that is, are able to extract rents from—the existing status quo institutions and the distribution of holdings. Even if these powerful interests would be better off under a high honesty low corruption system (and they almost certainly would), they have little reason to trust that they would be compensated for the loss of existing holdings and foregone future rents. The problem is exacerbated by the fact that high corruption low honesty societies are often low trust societies, where contracts—especially the kind of long-term compensatory contracts that would be required in this case—are routinely violated without compensation. Powerful and elite interests thus have

strong incentives to keep corrupt and inefficient systems corrupt and inefficient for they face a *commitment problem* with regard to post-transition compensation for lost rents and lost holdings. This commitment problem then exacerbates the credibility problem faced by potential voters, for even the most honest politician would need to overcome significant opposition of powerful entrenched elites in order to make good on pre-election political reform promises.

(b) A preference for collective action. If formal programmatic politics is out of the question, then why do citizens not organize with each other in order to solve the second-order collective action problem they face? In other words, if they cannot trust agents to be honest because the rules of the game are rigged in a way that rewards or encourages dishonesty, then why not change the rules of the game?

The conspicuous lack of collective action is due to a number of factors related to the class of individuals — as a political group — who would most benefit from the action. First, this is a large group in terms of sheer size, and, as Olson predicted many years ago, cooperation amongst large groups is lower due to the increased risks of free riding as well as the increased coordination and cooperation costs inherent in large group collective organizing. Second, this large group is often composed of the poorest and least educated individuals in developing societies, which means they lack resources and elite political connections. Third, there is a high probability of intra-group economic competition, since the number of poor and low-skilled laborer's often outstrip the available jobs. Fourth, unions—if they exist—exacerbate economic competition between laborers since surplus labor will tend to suppress the union's bargaining power and lower wages. Finally, there is a high probability of intra-group political competition, since various communities in dire socioeconomic straits are vying for public goods distributed at the informal discretion of the political elite.⁵⁵ This fractured, large, and often heterogenous group of individuals are often in competition with a highly organized highly resourced political elite who have an interest in keeping the status quo the status quo for the reasons mentioned above.

Organizing large groups of heterogenous individuals is thus often prohibitively costly. Other, related community wide beliefs raise the costs of cooperation. Consider, for example, the phenomenon of pluralistic ignorance. Pluralistic ignorance occurs when individuals mistakenly believe their own attitudes, motives, thoughts, and so on, are different from other individuals even every individual's public behaviour is identical.⁵⁶ Since there is a gap between private beliefs and public behaviour, people believe others implicitly endorse a behaviour that almost everyone privately holds negative beliefs about. In other words, the fact that many people engage in widespread corruption makes it seem as though many people endorse the practice or at least do not hold negative beliefs about it, even when most people do in fact hold negative views about it.

⁵⁵ Rudra (2018).

⁵⁶ Bicchieri (2016).

Pluralistic ignorance exacerbates collective action problems. Widespread discord in personal normative beliefs ('I think it is wrong for law enforcement officers to ask for a bribe'; 'law enforcement officers should not ask for a bribe') and second-order normative expectations ('I think other people think law enforcement officers should ask for a bribe') means that many people will be left with the false impression that there are not enough likeminded individuals to effect collective action. This results in a fatalism and sense of futility surrounding corrupt practices.

Moreover, pluralistic ignorance feeds into what's been called the *demoralization effect*.⁵⁷ When people begin to believe that the corruption that surrounds them is inevitable and insurmountable, they might stop engaging in the political process entirely. This is particularly problematic in democracies that rely on a well-informed and engaged electorate to 'vote out' bad candidates. Indeed, a strong enough demoralization effect severs the very link of accountability between the public and the candidate. Taken together, the effects of pluralistic ignorance sustain the vicious cycle of corruption by keeping the benefits of engaging in corruption higher than the costs of doing so. The false beliefs that drive pluralistic ignorance must be exposed and the illusion dispelled in order to begin to facilitate anti-corruption collective action.

b. Why non-programmatic politics?

(c) *A preference for clientelism.* If formal programmatic politics and collective action is out of the question, then what options are voters left with? They are left with the non-idealized option set that faces many citizens in developing societies, namely, they can choose to either engage in "selective collective action,"⁵⁸ which are often self-serving networks of political elites aiming to gain access to the spoils that high office gives them, or not engage in the political process (through non-participation or exit).

Clientelist politics generates its equilibrium stabilization pressures primarily through the incentives it produces for corrupt politicians. For example, if citizens believe, rightly in many cases, that both formal programmatic politics and organizing for collective action are not viable options, then they may themselves demand clientelist transfers. Demand of this sort may then place politicians, honest or not, into a Prisoner's Dilemma vis-à-vis one another described above, namely, and produce pressures for them to buy votes pre-election or to supply highly localized, targeted benefits post-election. In cases like these, clientelist demands are generated because of the dire prospects for engagement in formal programmatic politics and the costs associated with organizing collectively, which in turn create pressures for politicians to engage in corrupt politics. Note also that this dynamic would stabilize *even if* the politicians *and* voters had a preference for standard, run-of-the-mill programmatic politics, so long as pluralistic ignorance is pervasive, and individuals are systematically mistaken about others' preferences for corruption.

⁵⁷ Stephenson (2015, p. 95).

⁵⁸ Mungiu-Pippidi & Hartman (2019).

A further equilibrium stabilization pressure comes from the characteristics of the goods that corrupt transactions between politicians and potential clients. This point bears directly on the nature of the inducement, as suggested in the definition.⁵⁹ One incentive-compatible good that politicians often offer voters in exchange for political support are state jobs. Exchange of political support for state jobs is incentive compatible for both the patron (the politician) and the client (the voter) because the politician is able to withdraw the benefit “at-will” (by removing the client from the job, therefore ensuring continued political support), and the client is given what is often the best form of employment relative to jobs available in a highly inefficient and underdeveloped private sector. But since the offer of state employment must remain attractive relative to private sector employment, corrupt politicians then have a perverse incentive to keep poorly functioning contract enforcement institutions functioning poorly, for these are the bedrock upon which advanced economic activity depends. This is an example, of course, of a corrupt elite generating rents through weakly functioning institutions, which in turn generate commitment problems and raise transition costs in the dynamic described above. Moreover, it is no surprise, in light of this reasoning, that developing societies often have highly inefficient economies with bloated state governments, and where people are often vying for a job in the government rather than the private sector. Inefficient economies and bloated state governments are a byproduct of the link that exists between political and economic power for corrupt elites; in that sense, talking of “weak state capacity” is often a misplaced. Rather, a more accurate understanding of institutions in the developing world suggests that what is often at play is *selective* state capacity, where institutions are highly effective for some (the wealthy political class) at the expense of others (the poor voters), a state of affairs that reproduces itself through the perverse incentives it generates for powerful interests.

A key implication of this dynamic is that there need not be a conflict of interest between the patron and the client. Indeed, clients exhibit a demand for clientelist transfers, since informal clientelist politics is their sole avenue to express political preferences in a political system where formal politics is exclusively reserved for the elite. And without a conflict of interest, i.e., without a misalignment of incentives, principal-agent models and the agency problems they capture become largely irrelevant to the analysis. The key point here is that while the *conditions* for agency problems exist, the actual agency problems do not. The upshot is that clientelism-without-monitoring because an efficient political strategy either because there is clientelist demand given the client’s other viable options, or because patrons are forced into a Prisoner’s Dilemma with one another, where the costs of deviating from a corrupt “clientelist-transfer” equilibrium are, or are at least perceived to be, higher than the costs of compliance.

Strategic monitoring is not required in cases like this because the agency issues underlying the commitment problem are absent when political interests of the patron and the

⁵⁹ For more on incentive compatibility and employment relationships, see Baland & Robinson (2007, 2008).

(short-term) political interests of the client align. The cost-benefit changes in the long-term, of course, given the societal costs that everyone burdens that stem from clientelism in general. Alignment is possible if clientelism offers the best chances for effective political participation: If clientelism is the most viable means to solve local problems that require solutions on a “public goods” scale, or even just to receive redistributive private goods from public funds, then electing the patron who provides those goods is in the interest of the client. Our survey results indicated that the majority of people who had positive normative beliefs about vote selling did so on purely prudential grounds, that is, because of practical reasons.

Indeed, the theoretical model presented here complements the empirical evidence which suggests that clientelism is often the only way in which some communities can satisfy policy preferences by ensuring clientelist provision of private club goods (effectively local public goods). Both theoretical and empirical evidence suggests that for some particularly vulnerable communities, clientelism is often the only way to satisfy political preferences by ensuring clientelist transfer of private club goods (effective local public goods). These groups exercise political power through informal channels facilitated by clientelist networks often because their socioeconomic status precludes from them participating in the formal politics. Formal political and economic exclusion means that excluded groups must turn to informal channels in order to secure provisions from the state, whether they be local public goods for communities or private goods for individuals.

The preferential provision of private club goods to a deprived community might reasonably be regarded as the provision of local public goods to individuals within that community. But it is precisely the selective provision subject to the informal discretion of political elites that places the poor in the “Clientelist Trap”—and keeps them there. Self-reinforcing clientelist politics, on this view, is the result of a pernicious equilibrium: the community enjoying the local public good will not much care about the fact that its provision was the result of preferential distribution at the informal discretion of a political elite. But while the benefits are concentrated in one community, the costs of a system in which public resources are subject to informal discretion rather than public and transparent rules confers significant costs onto everyone subject to those distributive rules. Perhaps the greatest costs arise from the perverse incentive structures faced by both patrons and clients: patrons have incentives to keep poor people poor since their political power stems from the ability to selective distributive public resources to those that need it. The incentive structures on both sides of the patron–client relationship also provide a plausible story as to why clientelism stabilizes as a political equilibrium.

One empirical implication of these theoretical considerations is that we would expect individuals to have a preference for clientelism on generally prudential or pragmatic grounds.

This is precisely what we found during fieldwork in Nigeria.⁶⁰ Specifically, we found that a majority of respondents held negative normative beliefs toward vote selling (i.e., they thought that people *should not* sell their votes), there was a split between those who did so on non-prudential grounds (43%) and those who did so on prudential grounds (48%). However, due to the large size of this particular subset of respondents (n=4069) and their uneven distribution across states, we found significant variation in the relative frequency in types of normative beliefs across states.

Q3: DO YOU THINK THAT PEOPLE SHOULD COLLECT MONEY FOR A GIFT OR A VOTE			
	YES	NO	DON'T KNOW
ADAMAWA	7.4%	91.0%	1.6%
BENUE	7.3%	91.5%	1.2%
ENUGU	14.0%	80.5%	5.5%
LAGOS	19.2%	77.0%	3.8%
RIVERS	14.8%	83.0%	2.2%
SOKOTO	45.8%	50.1%	4.2%
FCT-ABUJA	20.4%	74.7%	4.9%

Table 2. *Personal normative beliefs – “should” across states*

Q5: WHY DO YOU THINK THAT PEOPLE SHOULD NOT COLLECT MONEY OR A GIFT FOR THEIR VOTE?			
	MORAL	PRUDENTIAL	OTHER
ADAMAWA	77.9%	13.9%	8.3%
BENUE	43.2%	55.5%	1.3%
ENUGU	26.1%	72.6%	1.3%
LAGOS	40.9%	54.2%	4.8%
RIVERS	45.4%	33.5%	21.1%
SOKOTO	44.0%	56.0%	0.0%
FCT-ABUJA	27.9%	52.0%	20.1%

Table 3. *Personal normative beliefs – “should not” across states*

⁶⁰ See Hoffman & Patel (2017, 2020).

Adamawa, for example, was the only state where the majority of respondents who thought selling votes was wrong thought so because of moral considerations (77.9%), with a minority considering it bad to sell votes because it is imprudent (13.9%). In Enugu, however, a sizeable majority thought that voters should not sell their votes because of prudential concerns with the minority considering it morally impermissible (26.1%). We found that those who held prudential normative beliefs against selling votes to be in the majority (amongst the particular subset of respondents who held negative normative beliefs toward vote selling) in all of the states except Adamawa and Rivers.

Q4: WHY DO YOU THINK THAT PEOPLE SHOULD COLLECT MONEY OR A GIFT FOR THEIR VOTE?			
	MORAL	PRUDENTIAL	OTHER
ADAMAWA	24.3%	58.1%	17.7%
BENUE	24.4%	75.6%	0.0%
ENUGU	5.6%	90.6%	3.9%
LAGOS	4.0%	85.8%	10.2%
RIVERS	3.4%	81.2%	15.4%
SOKOTO	5.1%	94.9%	0.0%
FCT-ABUJA	9.0%	77.3%	13.7%

Table 4. *Reasons for personal normative beliefs – “people should sell votes”*

In contrast, respondents who held positive normative views toward vote selling (i.e., they thought voters *should* sell their votes) did so mostly on prudential grounds (85.3%) rather than non-prudential grounds (7.8%). The distribution of this subset across states was such that over 75% of those who thought that voters should sell their votes thought so on prudential rather than moral grounds in all states except Adamawa.

This is significant because violations of shared prudential normative beliefs, unlike non-prudential normative beliefs, do not typically elicit sanctions. To illustrate, consider a community whose members believe that selling votes is good because they are much more likely to receive private clientelist transfers than they are to receive programmatic public goods from a politician after the election. If there is mutual consistency between individual personal normative beliefs and normative expectations, this particular community’s normative expectations are second-order prudential beliefs of the following sort: “I believe that others believe that selling votes is good because they believe they are much more likely to receive private clientelist transfers than they are to receive programmatic public goods from a politician after the election.” For any one individual in the community, the motivation for selling one’s vote does not come from the fact

that everyone in their community agrees that selling one's vote is the prudent course of action. Rather, it comes from the content of the shared belief itself.

This general point is true even if individuals in a community holds the opposite belief but for similar types of reasons. For example, consider a case where a community's members hold the normative belief that "I believe that selling votes is bad because if I sell my vote to one party, I may face violence from another party." If there is mutual consistency between individual normative beliefs and normative expectations, this particular community's normative expectations are second-order prudential beliefs of the following sort: "I believe that others believe that selling votes is bad because they may face violence from another party." Again, the motivation for not selling one's vote comes from a reason that is independent of the fact that many individuals in a community happen to hold it. Crucially, violators in this community (i.e., those who sell their votes) would perhaps be thought of as foolish for engaging in risky behaviour, *not* sanctioned for violating a shared moral rule.

The results suggest that these two cases – at least in terms of *types* of reasons – are reflective of vote selling in at least some contexts in Nigeria. In Sokoto, for example, 45.80% of respondents thought that individuals *should* sell their votes, and 94.9% this subset (n=339) thought so for prudential reasons. In Enugu, on the other hand, 80.5% of respondents thought that voters *should not* sell their votes, and 72.6% of this subset (n=377) thought so for prudential reasons.

Across most states, our conditionality measures revealed that respondents were generally insensitive to shifts in normative expectations, especially when exposed to those shifts under high empirical expectation condition. That is, when the hypothetical person in the vignette thought that everyone was selling their votes, it did not matter what everyone thought about selling votes. For example, in Lagos, shifting from high normative expectations to low normative expectations lead to a change of 0.01% in the target's predicted vote selling, whereas shifting from high empirical to low empirical lead to a change of 0.08%. Both of these are relatively small changes and thus are consistent with the idea that vote selling may be a shared prudential norm that is weakly, if at all, socially conditioned.

If vote selling is driven by shared prudential norms that confers private benefits to individuals whilst hiding the true social costs of their actions, vote selling is best thought of as a classic collective action problem. A potential intervention strategy to disrupt this dynamic would be to destabilize the prudential norm that lies at the heart of the collective action problem by exposing the tremendous externalities that such activities impose on political institutions. It should be clear, however, that the effectiveness of any intervention depends crucially on the content of the prudential belief as well as the conditions on the ground. No amount of information exposure will curb vote markets if, for example, individuals hold the belief that selling their votes is the only way to insulate themselves from political violence and that belief is in fact correct. This

is why the nature of the inducement—positive or negative—is integral to understanding the overall functioning of the vote market.

(d) *A preference for nothing*. The last option is to either directly disengage from the political process altogether, as indicated by the demoralization effect described above. This is, unfortunately, the option that some of the poorest and politically powerless citizens in the developing world choose through rejection of the only other viable choice (i.e., clientelism).

The other way to “do nothing” is, of course, exit. But exit is costly, both in economic and psychological terms. There is a well-established empirical link between corruption and immigration, and given the dynamic presented here, it is no wonder that people invest significant resources to emigrate from high corruption low honesty societies to high honesty low corruption societies. For exit opens up the better options in Option Set 1 (the Ideal Programmatic Option Set): it is just that the potential for programmatic politics and the benefits it generates lie elsewhere in the world. The vast majority of the poor in developing societies, however, face Option Set 2 (the Clientelist Dilemma), and thus become trapped in what I have described as the Clientelist Trap.

As should be clear, a number of these factors interact with each other to trigger vicious cycles and thus provide stabilization pressures for a pernicious high corruption low honesty equilibrium. Developing societies with high endemic corruption are often marked by poorly functioning formal contract enforcement institutions, resulting in high transaction costs across the relevant markets. High transaction costs preclude the type of compartmentalization of labor and complex chained economic relationships that are a precursor for a sophisticated, advanced economy and wealth creation. In this kind of economic environment, it is often easier to generate rents (by, e.g., selectively raising and lowering transaction costs or other corrupt activities) rather than create value through labor or investment, resulting in inefficient rent-generating political and economic institutions. Honest politicians wishing to reform inefficient extractive political and economic institutions then face high transition costs because of commitment problems faced by elites, and these commitment problems are themselves a result of low-trust due to the aforementioned poorly functioning formal contract enforcement institutions. These commitment problems then generate credibility problems for potential citizens, who must then turn away from programmatic politics to clientelist politics, in turn reproducing the entire vicious cycle.

5. Conclusion

In this chapter, I developed and defended a rational choice neo-institutional account of corruption. I explicated the theoretical connections between my account and economic and political development, paying particular attention to the role that institutions (both formal and

informal) play through transaction costs and trust. I then applied my account to political clientelism. I provided my own definition of clientelism, one that calls attention to the formal and informal institutions that structure political relationships that characterize clientelistic non-programmatic politics. I then provided a theoretical model of clientelism, the Clientelist Trap, which illustrated the conditions under which clientelism-without-monitoring might stabilize as an efficient political strategy (i.e., due to voter preferences for clientelist politics). I also presented novel empirical evidence to provide content to the theoretical account presented.

CHAPTER 3. Institutions out of thin air? A classical liberal case against charter cities

1. Introduction
2. What are charter cities?
3. Libertarianism and classical liberalism
 - a. Hobbes and the alienation contract
 - b. Liberalism, rights, and the separateness of persons
 - c. Nozick and the ultraminimal state
 - d. Classical liberalism, libertarianism, and the charter city
4. The colonialism objection
 - a. The wrongs of colonialism
 - b. Consent
 - c. Domination
5. Pragmatic arguments
 - a. Informal institutions matter
6. Conclusion

1. Introduction

Should the international community try to build a place like Singapore on the coast of a place like Somalia? Some people seem to think so.⁶¹ These people are proponents of what are called ‘charter cities,’ that is, cities created in an initial partnership between a host country and foreign guarantors that aim to effectively administer institutions that promote growth and wealth. The institutions of a charter city would ideally protect a set of rights (e.g., of property, of contract, and so on), uphold the rule of law, and provide otherwise business-friendly economic conditions. Economist Paul Romer first introduced the idea of charter cities back in 2010, and since then, it has enjoyed support from within libertarian and classical liberal circles.⁶²

Proponents of charter cities view them as a solution to difficult and complex policy problems, such as widespread corruption in the developing world and large-scale migration from the developing world to the developed world. These two problems are, as noted in Chapter 1, related. The thought is that charter cities will provide a set of fresh high honesty low corruption institutions and thereby provide people an opportunity to escape the high corruption low honesty institutions that are ubiquitous in the developing world. Supporters of charter cities go as far as to

⁶¹ In the rest of this paper, when I refer to proponents of charter cities, I am referring to, e.g., Romer (2009, 2010); Fuller and Romer (2012); Freiman (2013).

⁶² See, e.g., Freiman (2013).

claim that they provide the potential for a “new model of development that actually works.”⁶³ The rationale behind this view is that instead of attempting to change the formal and informal institutions that lead to ineffective governance in places like Somalia, the international community should instead focus on creating institutions that have worked elsewhere on the shores of places like Somalia, where people can “opt-in” by moving there and “opt-out” by leaving.

Proposals vary over the precise details of what a charter city would look like, but they almost all agree that it must (at least initially) be non-democratic in order to secure the kind of stability required for wealth generation and for trust in institutions to emerge.⁶⁴ Moreover, the development of charter cities is heavily tied to private interests, since one central justification for charter cities is that their institutions must be friendly to investment thereby facilitating economic development. It’s no surprise, then, that such proposals tend to be welcomed in self-described libertarian and classical liberal circles, since the thought is that pro-growth rules (that welcome investment and promote economic development) are simple rules that protect a minimal set of rights held dear by most liberals (e.g., of property, of contract, and so on).

In this paper, I present a classical liberal case against charter cities. My argument proceeds in two parts: (1) I claim that there are reasons to think that charter cities are not a morally justifiable form of government by the classical liberal’s lights, even if some form of the charter city (i.e., a private city) is an ideal form of libertarian government. Contrasting the positions of classical liberals and libertarians on this position is instructive. And (2) much of the pro-charter city arguments rest on a set of empirical claims that fuel pragmatic considerations, but these empirical claims are either false or they actually give us reason to question the very workability of charter cities.

The arguments in this paper have at least two interesting implications. First, I bring attention to a rift between libertarians and classical liberals that turns on separateness of persons constraints on liberal theories of justice, the justifiability of coercive takings for public goods,⁶⁵ and whether consent to a non-democratic state is possible. I show that the commitments of libertarianism and the commitments of classical liberalism are at fundamental odds with one another; this rift is so deep that morally justifiable governments on the libertarian view are morally unjustifiable on the classical liberal’s view, and vice versa. And second, my arguments tentatively support one prominent critique of charter cities, namely, that they are a form of objectionable colonial relations. Moreover, in addition to supporting the colonialism objection, my arguments go a long way in deflating the response to the colonialism objection from the proponents of charter cities.

⁶³ Fuller and Romer (2012, p. ii).

⁶⁴ See, e.g., Sagar (2016, p. 514).

⁶⁵ ‘Public goods’ here means goods that are nonrival and nonexcludable, that is, a good that can be effectively provided to all and one that individuals cannot be effectively excluded from. Classic examples of public goods are clean air (and other environmental goods), lighthouses, and national defense.

In §2, I present a brief outline of my understanding of charter cities. §3 elucidates the commitments of libertarianism and classical liberalism. I begin with a discussion of the Hobbesian social contract. I then show how concepts of inalienability of rights, the provision of public goods, and the separateness of persons constraints normatively hang together, and how these linkages differ for the libertarian and classical liberalism. The positions that the libertarian and classical liberal endorse differ on how these concepts hang together, and this difference has implications for whether charter cities are morally justifiable. In section §4, I outline one prominent critique of charter cities, namely the colonialism objection, and show how my account supports this critique. I also show that the standard response from proponents of charter cities do not answer the critique. In §5, I summarize a pragmatic argument that proponents of charter cities put forward in favor of their view, and show that it is untenable. §6 concludes.

2. What are charter cities?

Proponents of charter cities find their inspiration in ‘special economic zones’ (SEZs) found in developing countries around the world. SEZs typically have different business and trade laws than the rest of the country in order to attract foreign investment and provide a generally business friendly environment. Laws in SEZs are typically geared toward reducing the cost of doing business by providing a business climate with more lenient or flexible labor regulations, lower tax liabilities, lower customs duties, and so on. Importantly, charter cities differ from SEZs in that they are created (‘chartered’) through law by the legal authority that controls the sovereign territory in which the charter city will be located.

There are four features that all charter cities would share are:

1. *Size.* First, charter cities are best thought of as expansive SEZs in literal terms of their physical characteristics. They must be “large enough to eventually accommodate a city with millions of residents,” and that a “good target size is 1,000 square kilometers, roughly the size of Hong Kong and Singapore.”⁶⁶ Moreover, since market economies rely on trade, an essential feature of proposed charter cities is that they are in a position to exchange goods and services with friendly trade-partners in order for their complete independence from their neighbouring countries. This is usually interpreted as a requirement that they have access to the sea.
2. *An actual charter.* Secondly, and perhaps one of the most important features of charter cities, is the idea that an actual charter that “pre-specifies the broad rules” that would apply within its borders. These “reforms must extend to all the rules

⁶⁶ Fuller and Romer (2012, p. 3).

needed to structure interactions in a well-run city and to support exchange in a modern market economy.”⁶⁷ These formal rules ought to be broad enough to allow for experimentation, whereby some experiments will persist and others perish through market competition.

3. *Freedom of choice.* Thirdly, there must be a political commitment to choice, which is usually interpreted as “voluntary entry and free exit for all residents, employers, and investors.”
4. *Rule of law.* Fourthly, and finally, charter cities must contain a “commitment to the equal treatment of all residents under the law,” which is best interpreted as a commitment to the rule of law.

Charter cities require participating sovereign nations to provide three things: the land on which the charter city is built, the people that will move to the charter city, and the credible guarantee that the charter’s rules will be enforced fairly and reliably. What this in effect means is that there will be a foreign or private (or both) element typically in the enforcement and guarantee of the charter’s rules: for example, when China presumed control of Hong Kong from the British in 1997, it agreed that justices for the highest courts in Hong Kong could be recruited from other commonwealth nations (and indeed they were, from places such as New Zealand and Australia).⁶⁸

	Who creates ?	Who controls ?	Given Justification?	Public goods?	Governance Structure ?	Example
Special Economic Zone (SEZ)	Sovereign state	Sovereign state	Economic Growth	Public	Democratic or non-democratic	Shenzhen
Trusteeship	Intl. Community	Intl. Community	Transitional Assistance	Public	Democratic or non-democratic	Kosovo
Colonial	Foreign State	Foreign State / Private Interests	Self-Interest / Paternalistic	Public and/or Private	Non-democratic	Diego Garcia

⁶⁷ Ibid., p. 4.

⁶⁸ Fuller and Romer (2012, p. 12).

Charter City	Sovereign state	Self-governing/ Sovereign State / Foreign State / Private Interests	Economic Growth	Public and/or Private	Democratic or non-democratic	Honduran RED (Regiones Especiales de Desarrollo)
--------------	-----------------	--	-----------------	-----------------------	------------------------------	--

Table 5. Distinctions between types of cities.⁶⁹

The broad characteristics of a charter city outlined by their staunchest proponents has led to some confusion as to the precise definition of a charter city. For example, Rahul Sagar finds it puzzling that Romer and Fuller take Shenzhen to be an example of a charter city since Shenzhen is an SEZ administered completely by the Chinese government and, as such, Shenzhen does not have autonomy from the Chinese government. As Sagar explains, this means that “its credibility is indistinguishable from that of the national [Chinese] authority” which in turn makes the “concept [of the charter city] redundant.”⁷⁰

In order to avoid these ambiguities, I present my own taxonomy of cities in *Table 5*. As *Table 5* suggests, there are a number of governance structure possibilities in the creation of a charter city. For example, a charter city might be controlled completely by private interests, a foreign state, and a sovereign state that created it, or be self-governing. It could be a combination of all four. It might have a public provision of public goods (i.e., its own legitimate police authority), or it may have a completely private provision of public goods (i.e., private security services take the role of enforcing contracts and protecting property). What this means is that there is a range of possible governance structures in charter cities, which theoretically run the gamut from a completely private city to a completely public one, to a combination of both. Private cities refers to a city where important governance decisions are undertaken by private actors (such as developers or large corporations, and so on), and essential public goods (such as police protection, national defense, and so on) are provided privately. Public cities refers to a city where important governance decisions are undertaken by public actors (such as a democratically elected legislative authority), and essential public goods (such as police protection, national defense, and so on) are provided publicly.

It is important to note that some governance structures in charter city are, of course, more likely, since the justification of the charter city is that it attracts investment and is generally “business friendly.” What this means is that governance structures are non-democratic (to ensure stability) and private interests are heavily involved in the creation of rules (or are in complete

⁶⁹ This is an edited version of the table found in Sagar (2016).

⁷⁰ Sagar (2016, p. 512).

control, as in a full private city). This has been the experience of real life charter cities to date; Paul Romer, citing transparency concerns, was forced to distance himself from the Honduran *Regiones Especiales de Desarrollo* (RED)—a proposed charter city in Honduras—after the then Honduran president signed a memorandum of understanding with a group of private investors.⁷¹ What I will suggest in the coming section is that the completely private city represents the ideal form of libertarian governance, and that the classical liberal has good reasons to reject this libertarian utopia. One the main points of contention starts with the question of whether people are able to consent to the government in a libertarian utopia, i.e., the private city. It is to this I turn to now.

3. Libertarianism and classical liberalism

a. Hobbes and the alienation contract

Perhaps the most famous instantiation of consent to non-democratic governance comes from Thomas Hobbes. In this section, I will outline Hobbes's argument, which will be instructive for contrastive purposes later in this paper.

Hobbes claimed, *contra* Locke, that a state of nature (a state in which there is no government) is equivalent to a mutually destructive state of war. To sum up briefly, Hobbes's argument goes something like this: (1) there is a natural equality in the physical and mental endowments of individuals in the state of nature (any differences between individuals in terms of these are negligible); (2) this natural equality leads to competition and a state of distrust between individuals; moreover, (3) given that individuals may be driven by "pride and vainglory" and that no one individual or group of individuals have a monopoly on violence to enforce contracts, individuals have an incentive to 'strike first' in anticipation of attacks by others to provide for their security; and therefore (4) this constant insecurity and incentive to strike first leads to the state of nature—that is the state of war—means that individuals are in "continual fear, and danger of violent death," which in turn means that life is "solitary, poor, nasty, brutish, and short."⁷² In sum, the state of nature is a state in which individuals are in a "war of all against all."⁷³

In order to avoid this state of nature, Hobbes contends, it is rational for individuals to institute a contract to authorize an absolute sovereign that will provide stability necessary for society itself to exist. The nature of this authorization has been the subject of discussion in political philosophy, with at least two prominent positions defended: (1) an *alienation* social contract, whereby individuals give up ('alienate') at least some rights (such as rights to basic self-governance) to the sovereign through consent; and (2) an *agency* social contract, whereby individuals delegate at least some rights to a sovereign. From the former view, individual's

⁷¹ See <https://www.nytimes.com/2012/10/01/world/americas/charter-city-plan-to-fight-honduras-poverty-loses-initiator.html> and <https://foreignpolicy.com/2014/09/01/under-new-management/>

⁷² Hobbes, *Leviathan*, I.XIII.9 (1998/1642).

⁷³ Hobbes, *De Cive*, I.XIII (1984/1651).

surrender their rights to the sovereign, thus the terms of the authorization constitute a type of authoritarian relationship.⁷⁴ From the latter view, the sovereign's power is *loaned* to that sovereign, thus the terms of the authorization constitute a type of principal-agent relationship.⁷⁵

Notwithstanding the longstanding debate about the nature of authorization in Hobbesian political philosophy, it remains true that the former view is typically seen as the Hobbesian alienation social contract, and the latter as the Lockean agency social contract.⁷⁶ Consider some representative quotes that motivate this reading of Hobbes:

“Since therefore the *conspiring* of many wills to the same end doth not suffice to preserve peace, and to make a lasting defence, it is requisite that in those necessary matters which concern Peace and selfe-defence, there be but *one will* of all men ... This *submission* of the *wils* of all those men to the *will of one man, or one Counsell*, is then made, when each of them obligeth himself by contract to every one of the rest, not to resist the *will* of that *one man, or counsell*, to which he hath submitted himselfe[.]”⁷⁷

“The only way to erect such a common power, as may be able to defend them from the invasion of foreigners, and the injuries of one another, and thereby to secure them in such sort, as that by their own industry, and by the fruits of the earth, they may nourish themselves and live contentedly; is, to confer all their power and strength upon one man, or upon one assembly of men, that may reduce all their wills by plurality of voices, unto one will ... *I authorize and give up my right of governing myself, to this man, or to this assembly of men, on this condition, that thou give up thy right to him, and authorize all his actions in like manner.*”⁷⁸

“In all cities or bodies politic not subordinate, but independent, that one man or one council, to whom the particular members have given that common power, is called their SOVEREIGN, and his power the sovereign power. which consisteth in the power and the strength that every of the members have transferred to him from themselves, by covenant. And because it is impossible for any man really to transfer his own strength to another, or for that other to receive it; it is to be understood: that to transfer a man's power and strength, is no more but to lay by or relinquish his own right of resisting him to whom he so transferreth it. And every member of the body politic, is called a SUBJECT, (viz.) to the sovereign.”⁷⁹

⁷⁴ Hampton (1986, p. 114).

⁷⁵ Ibid., pp. 114 – 115.

⁷⁶ Ibid.

⁷⁷ Hobbes, *De Civ*, V.VII (1984/1651).

⁷⁸ Hobbes, *Leviathan*, II.XVII.13 (1998/1642).

⁷⁹ Hobbes, *Elements of Law*, I.XIX.10 (2017/1650).

The point is that the social contract Hobbes has in mind authorizes a sovereign whose power is permanent and absolute. The justification for the absolute power of the sovereign comes from Hobbes's regress argument,⁸⁰ which runs as follows: any limited political power in a society must have its limits placed by another, more powerful political authority. If that more powerful political authority is itself limited in any sense, then those limits must be credibly imposed by a greater political power. In the end, for there to be any credible political power, the buck must stop somewhere, so to speak, in the form of the ultimate political power that can set the limits on other political power but cannot itself be limited by any other political power. This ultimate political power guarantees the stability that motivates Hobbes's political thought, indeed, for him, it is *the* definitive feature of government itself.

Thus, on Hobbes's view, individuals in the state of nature consent to the creation of an absolute sovereign through an authorization that alienates some of their basic self-governance rights. That is, they consent to a non-democratic form of government. This is because the stability that is necessary for civil society itself springs only from the creation of an ultimate authoritative sovereign whose power is not under the supervision or control of the individuals who have authorized its power. The type of power Hobbes envisions the sovereign to have is *not* a fiduciary power; the power is *not* held in trust by the sovereign and the sovereign and the individuals who authorized the sovereign do not stand in a principal-agent relationship with one another.⁸¹

a. Liberalism, rights, and the separateness of persons

In the last section, I outlined Hobbes's famous argument for an authoritarian Leviathan. I now turn my focus onto liberalism and liberal institutions, since it is the dispute between two liberal views (libertarianism and classical liberalism) that is the focus of this paper. Hobbes was no liberal, and it's no surprise that classical liberals have good reason to reject the Hobbesian alienation social contract. But, as we shall see, libertarians might have more in common with the Hobbesian alienation contract than they think.

Liberalism is a philosophical view that stresses liberty and equality. Political liberalism holds that people are equal, that they ought to be as free as possible to pursue their own conceptions of the good, and that there are many conceptions of the good that reasonable people may disagree about. A minimally just liberal society is a society in which people are able to pursue their own conception of the good while respecting others' rights and freedoms. Gaus and Mack identify at least nine core commitments of what they call the 'liberty tradition':⁸²

⁸⁰ Hampton, (1986, p. 122).

⁸¹ Hampton, (1986, p. 123).

⁸² Mack and Gaus (2004) actually outline twelve doctrinal commitments, but I highlight nine here for simplicity. This is because it seems that at least two of their doctrinal commitments collapse into each other, or are, for our purposes here at least, indistinguishable from each other.

- i. Normative individualism (i): interpreted as a separateness of persons requirement.
- ii. Individual liberty to be a core political or legal norm.
- iii. Respect for an individual and her liberty requires respect for the individual's control of extra-personal objects (e.g., property).
- iv. A desirable social order allows individuals to pursue their own *individual* conceptions of the good without legislating an overall common goal.
- v. A social order where individuals are free to make choices within the constraints of an expansive set of rights is preferred over a centrally planned system (this commitment stems from an appreciation of Hayekian knowledge problems).
- vi. Coercion is permissible only in the case that it is used to stop impermissible infringements on the rightful claims of others.
- vii. Political institutions in the liberty tradition authorize the use of force and threaten and use force against citizens legitimately.
- viii. Liberty tradition rejects a basic distinction between morality that applies to individuals generally and the public.
- ix. A general presumption against political power in the form of a suspicion that political regimes have historically 'engaged in extensive acts of unjustified aggression, plunder, and meddlesomeness.'

Clearly the liberty tradition is vast and encompasses a range of political thought, from minimal-state libertarianism to classical liberalism to expansive state high liberalism. I take Robert Nozick to be the paradigmatic minimal state libertarian, and libertarians of this ilk endorse a state that is limited to the enforcement of a minimal set of rights (most importantly of property and of contract) and to the resolution of disputes. I take David Hume and Adam Smith, and Friedrich Hayek to be the paradigmatic classical liberals, and classical liberals of this ilk endorse a relatively limited state that protects a set of rights and resolves disputes, but also provides a set of non-controversial public goods. I take Immanuel Kant, John Stuart Mill, John Rawls to be the paradigmatic high liberals, and high liberals of this ilk endorse a state that protects a set of rights, resolves disputes, funds public goods, and allows for redistribution to preserve liberal institutions.

Those working within the liberty tradition generally accept some form of the above doctrinal commitments. But the thought here is that to the extent that there are intellectual disputes between proponents of views *within* the liberty tradition, they are disputes that, generally speaking, adopt some variant of most (if not all) of the above commitments.

Most standard contemporary liberal theories of rights invoke some notion of 'separateness of persons'—which is a combination of a mundane descriptive claim and a more important normative claim. The mundane descriptive claim is that people are separate—they have their own lives, values, goals, and so on. The normative claim is that each individual's life, values,

goals, and so on, are intrinsically important and ought to be respected to the extent that we have grounds to reject most aggregative political theories, such as utilitarianism.⁸³ The thought behind the notion of separateness of persons is simply that we ought to respect the moral status of individuals *qua* individuals: it implies that there are some ways in which we may never treat an individual, that indeed, in some important respect, ‘individuals are inviolable.’⁸⁴

An acceptance of the separateness of persons is usually built-in to a liberal theory of rights because it identifies the ways in which a society ought to construct interpersonal rules and claims that facilitate each individual pursuing their own ends to the extent allowable without unjustifiably impeding upon another person’s pursuit of their ends. In other words, rights, in this sense, create boundaries in which one may legitimately pursue their own ends without trampling upon the rights of others.

Different moral theories may interpret a separateness of persons requirement more or less seriously. For illustrative purposes, consider the fact that both Nozick and Rawls reject utilitarianism on separateness of persons grounds,⁸⁵ but Nozick interprets separateness of persons requirements more stringently than Rawls. Nozick claims that a proper appreciation of separateness of persons results in a side-constraint theory of morality, which reflects the ‘underlying Kantian principle that individuals are ends and not merely means.’⁸⁶ A person’s property is an inviolable part of their person, on Nozick’s view, such that coercive takings that result in redistributive policies violate the separateness of persons requirement and thereby some longstanding commitments of the liberty tradition (in particular commitments (i) to (iv) above). On a less stringent view, e.g., Rawlsian view, a person’s wealth is not considered part of their personhood, which in turn means coercive policies that are redistributive do not violate the separateness of persons requirement. Nozick’s view involves an absolutism about property rights that is supported quite naturally by his strict interpretation of the separateness of persons requirement. On a third, but still less stringent view, e.g., a classical liberal view, even if we accept that a person’s wealth is considered part of their personhood, public goods might still be justified as a form of coercive takings if the provision of the public good leaves everyone better off than they otherwise would be. This is by hypothesis true of public goods, and so public goods are morally justifiable on this classical liberal view.

There is an important distinction, on Nozick’s view, between the structure of morality and the content of morality. An appreciation of separateness of persons informs Nozick’s “side-constraint” theory of morality. In other words, accepting normative individualism means that we must accept that there are “morally binding, absolute, [and] exceptionless side constraints”⁸⁷ on how we may treat them.

⁸³ Not always; see Gaus and Mack 116, on the ‘highly qualified’ liberal utilitarianism.

⁸⁴ Nozick (1974, p. 31).

⁸⁵ See, e.g., Nozick (1974, p. 33); Rawls (1999, p. 28).

⁸⁶ Nozick (1974, p. 30).

⁸⁷ Arneson (2011, p. 1).

But this side constraint theory of morality is independent of the content of morality. The content of morality is a minimal set of rights that protect life, liberty, property, and contract. Call these ‘libertarian rights.’ The argument for these libertarian rights, for Nozick, runs as follows: (1) an agent has a moral right to act however she may choose to with whatever she owns (provided that her ownership is legitimate), insofar as the exercise of her rights does not encroach on another agent’s rights. (2) Each person has a right that others do not encroach on their rights. Violations of rights include theft, fraud, use of force, and so on. (3) Each person enjoys full self-ownership. (4) Each person is able to acquire ownership over previously unowned resources as long as ownership of unowned resources does not leave anyone worse off than if the resources were left unowned. (5) Ownership rights an agent has over herself and her property are transferable through voluntary transactions (such as contracts or gifts).⁸⁸

These rights are ‘state of nature’ rights in that they exist independently of any political institution that enforces them and independently of others’ beliefs and political views. The moral justifiability of any given state turns on whether it violates these rights. I refer to Nozick’s view as collapsing into an absolutism about property rights because there is no case in which conformity to libertarian rights might be relaxed even if the consequences of conformity might be quite dire. For example, respecting libertarian rights as stringently as Nozick conceives of them rules out the very possibility of the public provision of public goods, and this *would* lead to situations that one might consider quite dire (or so I shall argue in §4d). Classical liberals have good reasons to reject the stringency with which Nozick and other libertarians conceive of libertarian rights, and accepting these reasons has implications on what sorts of governments each position deems morally justifiable. Before I bring attention to these differences, I will now turn to Nozick’s justification of the ultraminimal state.

c. Nozick and the ultraminimal state

In *Anarchy, State and Utopia*, Nozick provides what he calls an ‘invisible-hand explanation’ for how a non-democratic ultraminimalist state might arise without violating the rights of anyone who comes to live under that state.

For Nozick, individuals living under no state would likely seek protection services in order to protect themselves from those seeking to harm their person or property. These services would be offered under market like conditions, that is, with “protective associations” competing in order to secure clientele to whom they could provide their services.⁸⁹ A natural monopoly would likely emerge in the protection services market because conflicts between clientele of different agencies would require different property claims to be adjudicated by a set of rules that applies to all. Either a single protective association would overpower others, or it would be

⁸⁸ Arneson (2011, p. 3).

⁸⁹ Nozick (1974, p. 12 – 17).

rational for protective associations to cooperate with one another to prevent costly disputes. Moreover, this dominant protective association would prohibit people within its geographical territory from exercising their own rights to defensive violence in order to mitigate the risks exposed to their clientele that would arise from the independent exercise of such rights. While these would-be independent exercisers of rights would be prohibited from the sort of defensive violence required to enforce their own rights, they would be compensated for the loss of these rights through the free protection services offered by the dominant protective association.

Therefore, for Nozick, a single dominant protective association would emerge, and it would have two necessary features of a state: (1) a monopoly on violence in a given geographical territory; and (2) protection of rights of everyone in the territory.⁹⁰ It's important to note that Nozick's invisible-hand explanation of the ultra-minimalist state is a response to an anarchist challenge, that is, to the question of whether the state can be justified at all. And what Nozick claims to have shown is that an ultra-minimalist state can be justified on the grounds that it would arise "by spontaneous groupings, mutual-protection associations, division of labor, market pressures, economies of scale, and rational self-interest."⁹¹ Importantly, it would also arise without violating anyone's fundamental rights.

It should be stressed that there is no requirement for this libertarian state to be democratic. Indeed, for the libertarian, there is nothing special about democracy as a form of government; a non-democratic government that preserves the minimal set of rights that Nozick favors is more desirable than a democratic government that violates them. In other words, a democratic government that raises funds through taxation to fund non-controversial public goods (other than those governed by the dominant protective association) is morally inferior to an autocratic government that does not tax beyond what is necessary to provide the minimal set of protective services. Nozick's interpretation of separateness of persons is so stringent that it results in an absolutism about property rights, and so there is no room left for comparative judgments about the relative merits of democratic forms of governance versus other forms of governance.⁹²

One final point is worth mentioning here. Not only is there no requirement that the libertarian state justified by Nozick be democratic, but libertarians tend to be openly skeptical or even hostile to democratic forms of government.⁹³ This skepticism of democracy is fueled by an empirical claim that democracies may be more likely to routinely violate private property rights for, e.g., redistributive purposes. Classical liberals such as Hayek share this worry, but libertarians ought to be even more worried since even coercive takings to fund public goods are unjust on their view (a view that is not shared by the classical liberal, but more on this in the next section).

⁹⁰ Nozick (1974, p. 22 – 23).

⁹¹ Ibid., (1974, p. 16 – 17).

⁹² Arneson (2011, p. 3 – 5).

⁹³ See Arneson (2017) for a good overview.

The thought is that a democratic state simply provides more opportunity for transfers of wealth over a state that limits itself to the protection of a set of libertarian rights.

d. Classical liberalism, libertarianism, and the charter city

In many ways, one form of the charter city (i.e., the private city) represents a real-life instantiation of a perfect libertarian society: (1) Government institutions are limited to the set of institutions that one would find in the perfectly libertarian society: to protect against infringement of a minimal set of rights, to enforce contracts, and to resolve disputes. (2) There are no democratic institutions, and so the libertarian skepticism about the tendency for democratic institutions to upset existing distributions of holdings need not apply. And (3) Public goods are all or in part provided privately. In this section, I present some reasons to question the moral justifiability of this libertarian paradise.

Before I move on, a caveat is in order. My interpretation of a charter city in this section might strike some proponents of charter cities as alien to theirs. The people I have in mind here make a distinction between charter cities and private cities, where the former has a set of public goods provided by a formal state, and the latter has its set of public goods provided privately. For example, Paul Romer explicitly rejects the idea of a private city.⁹⁴ But my characterization of private cities as a form of charter cities is justified because of three reasons: (1) there is nothing in the proposals put forward by proponents of charter cities that theoretically rules out charter cities as private cities, and there is nothing to say that a private city cannot be a type of charter city. This goes even for non-libertarian proponents of charter cities. Further, (2) there are reasons to think that a charter city might very well end up as a private city, given the private interests that play a heavy role in the creation of the charter city. And (3) I do refer to this as a *libertarian* charter city, and I claim that *libertarians* must be committed to the idea that the private city, on their view, is a justifiable form of government. Proponents of charter cities who do not consider themselves libertarians may think that the criticisms in this section do not apply to them, though they must still provide reasons for why we shouldn't expect charter cities to end up as libertarian private cities (or even something like libertarian private cities), given reasons (1) and (2).

I now turn to two features of a libertarian charter city that are morally objectionable on most liberal views, including those of the classical liberal. These are the private provision of public goods, and the private nature of political power in the charter city.

Private public goods. Recall that, in the last section, I claimed that libertarians have no theoretical room to make comparative judgments about the relative merits of democratic governance against competitors other than on the grounds that one respects a set of minimal

⁹⁴ See, e.g., <https://paulromer.net/interview-on-urbanization-charter-cities-and-growth-theory/>.

rights (i.e., of property, of contract, and so on) more than another. I claimed that, on the libertarian view, a democracy that funds uncontroversial public goods through taxation is unjustly coercive and morally inferior against an autocracy that does not fund public goods but otherwise protects and maintains rights, enforces contracts, and resolves disputes. A libertarian might respond to this claim by pointing out that I am begging the question in claiming that they would reject funding *non-controversial* public goods through taxation since it's precisely the moral permissibility of public goods that is at stake. For the libertarian, *all* public goods provided publicly (i.e., by the state), other than the protective services provided by the dominant protective association, are controversial, since *all* public goods provided by the state require unjust coercive takings. That I refer to them as non-controversial seems to imply that an important question is settled, that I am in fact assuming, in other words, what ought to be shown.

But this libertarian view on public goods is an outlier amongst liberal positions, since most plausible liberal views reject the idea that the funding of public goods violates non-redistributive norms. Indeed, most liberal views accept the idea that respecting separateness of persons rejects redistributions whereby some gain and others lose; but the provision of public goods does not involve redistributive coercive takings since everyone is left better off than they otherwise would be. Note that this line of reasoning rules out the strict Nozickian interpretation of separateness of persons and of property rights as absolute; instead, the stringency of the right-claims surrounding property ought to be as stringent as necessary to serve the particular interests that serve as the justification of the property right in general. As Gaus and Mack put it:

“If the rationale for rights is their service in protecting people’s basic interests, the rationale for rights would point to rights of the less stringent sort. In the case of public goods, agents’ basic interests are served by possessing rights of the less stringent variety. For this allows coercive takings that are, by hypothesis, necessary for public goods to be produced whereas more stringent rights would forbid those takings and leave agents worse off in their basic interests.”⁹⁵

I take the paradigmatic classical liberal position to accept something along these lines. The claim that the government is justified in using coercive takings to fund public goods is a line of argument that stretches back to classical liberals such as David Hume and Adam Smith. This is a stark point of contrast between the classical liberal and libertarian position: the classical liberal position does not interpret the separateness of persons in a way that precludes a morally justifiable state in funding things like public goods through coercive takings.

Note that the justification for coercive takings for the provision of public goods, on this particular classical liberal view, rests on the idea that the rationale for rights is that they serve an

⁹⁵ Mack and Gaus (2004, p. 123).

agent's basic interests. Libertarians reject this view. While Nozick himself admits that he provides no "precise theory of the moral basis of individual rights,"⁹⁶ it's clear that the stringency with which he regards an agent's rights precludes this sort of rationale. If conformity to his particular interpretation of libertarian rights is required at *all times* under *all circumstances*, coercive takings are, *ipso facto*, unjust. The justification of public goods, even if they leave everyone better off, still runs afoul of the underlying non-consequentialism of Nozick's view.

What this means is that the libertarian state does not provide public goods because the provision of public good involves unjust coercive takings. Libertarians emphasize this as one of their positions: "[I]n the libertarian society ... everything will be privatized, there will be no public property. There will be no public police[.]"⁹⁷

But, as the above quote suggests, a society cannot function without a minimal set of public goods; even a libertarian state, where there must be an institutional apparatus to protect rights, enforce contracts, and resolve disputes, requires what would normally be taken to be a set of public goods (e.g., security services, national defense infrastructure, and a system of laws and courts, etc.) In response to this, libertarians either claim that markets will provide adequate provision of public goods in the absence of government intervention, or they reject the very idea of public goods as a coherent concept.⁹⁸ Note that while it's possible to hold both of these views independently of each other, many libertarians hold both (indeed, for people like Walter Block and Hans Herman-Hoppe, the former implies the latter).

But there are at least two reasons to question the private provision of public goods. The first reason is roughly instrumental: in many cases, public goods are provided privately at the expense of some other important liberal value, such as equal protection before the law. This is because of a clash of private and public interests. A society without public goods must leave important functions of the state, such as its police power, to the market to provide. But market incentives reward private interests, and private interests come apart from social interests quite easily. Paul Romer himself writes about the internal handling of sexual assault by university police departments and religious authorities, where the interests of the institution (the university and the religious community respectively) are in direct conflict with the interests of the individuals whose rights have been violated by persons within that institution. Universities may be hesitant to pursue sexual assault charges against, e.g., members of a sports program (whether it be players or coaches) because of how lucrative those sports programs might be to the overall coffers of the university. Similarly, religious authorities may be hesitant to pursue sexual assault charges against, e.g., Cardinals, in order to preserve the integrity of the institution in the public's eye. Cases such as these are ubiquitous and extend to other areas in which normal government functions have been outsourced to private or semi-private actors; the privatization that fuels the

⁹⁶ Nozick (1974, p. xiv).

⁹⁷ Block (2011a, p. 52 – 53).

⁹⁸ See, e.g., Block (2011b) and Herman-Hoppe (1989).

military-industrial complex and the prison-industrial complex in the United States are two particularly disturbing examples.

Lessons from ‘real life’ private cities provide further support for how private interests depart from social interests. Consider Gurgaon, which is something like a private city,⁹⁹ just outside of New Delhi. Many public goods, such as the provision of electricity, water, sewage, and security, are provided privately in Gurgaon; in 2011, for example, there were 35,000 private security guards and just 3,100 public police officers.¹⁰⁰ As one would expect, security provision is quite good for those who can afford to pay for it (such as the middle class and above), but quite bad for those who cannot (the poor).¹⁰¹ Further, even though the security provision is good for those in the middle class and above, there is still a perception that crime is quite high in Gurgaon (a perception that departs from reality), where “members of the middle class do not feel secure and instead see themselves as living on islands of private security surrounded by a sea of criminality.”¹⁰² Security services do not have an incentive to deflate this general perception, and indeed, their incentives run in the other direction. This is yet another example of the clash between private and public interests, since a general perception of safety (when security provision is adequate) is in the interest of the common good. It’s worth adding here that it’s precisely this sort of “siege mentality” that the Hobbesian Leviathan was intended to remedy, since it’s this siege mentality that leads to the instability and cruelty of life in the state of nature.

Libertarians might respond to this last line of reasoning by pointing out that the sort of violations of equal protection I am pointing out are routine even when governments hold the monopoly on the provision of public goods, in this case, of police protection. It’s not as if, they might argue, poor people in the rest of Indian society outside of Gurgaon (or outside of any other private city) enjoy an adequate amount of security along with their wealthier counterparts. However, the important difference is that when a public government is in charge of providing public goods and fails to do so in a way that violates equal protection, that government is in some important way failing its citizens: this failure represents a *deviation* from the relationship that ought to be enjoyed between a citizen and their government. But private provisions that result in the uneven provision of the public good and the violation of equal protection for all is a *feature* of a system of market relationships; the important good in question is doled out on the basis of a person’s effective demand and relative bargaining position.¹⁰³

The second reason why we might object to the private provision of a public good is noninstrumental: the private provision of public goods is objectionable because some exercises of political power ought to be inherently public, and whether a political power is exercised in a public or private manner hinges, in some important way, on the actor exercising the power. In

⁹⁹ Rajagopalon and Tabarrok (2014).

¹⁰⁰ Ibid., p. 212.

¹⁰¹ Ibid., pp 212 – 213.

¹⁰² Rajagopalon and Tabarrok (2014, p. 213).

¹⁰³ Freeman (2002, p. 142 – 143).

these cases, the political actor *must* be a public authority in order to secure the good the action is intended to procure.

One natural way to interpret this claim is to accept that some exercises of public power are incapable of outsourcing. Avihay Dorfman and Alon Harel hold this view.¹⁰⁴ Their argument runs roughly as follows: Consider privatized criminal punishment, such as private prisons, which are pervasive in the United States. If criminal punishment is intended to express, beyond retributive purposes, *public* condemnation of the act the criminal is being punished for, then the private provision of criminal punishment necessarily involves the violation of the equality that exists between the private person (or persons, in the case of a private company) providing the punishment and the criminal receiving it as persons before the state. In this case, the state has outsourced its own privilege—to speak and act in its name—to a private citizen; and that private citizen is then entrusted with meting out the punishment in the name of the public good on another private citizen. The state, in effect, chooses to privilege one citizen’s judgment for how a public good ought to be realized over another’s, thereby undermining the moral equality it ought to afford to both people in question.¹⁰⁵ From this view, a private actor who metes out criminal punishment cannot realize the public good at stake, since the private actor’s actions are not undertaken in the name of the state.¹⁰⁶

Private political power. There is a broader theme underlying the last point here which provides a stark contrast between the libertarian position and other liberal positions. It is that, in the libertarian state, political power is not treated as a public power entrusted to the state on behalf of its citizens. Political power is private power, and the beneficiaries of its exercise are those that have paid for it. Like the Hobbesian Leviathan, the relationship between the citizens and the state is *not* fiduciary.

Things are no different in the libertarian charter city. Since even non-libertarian proponents of charter cities stress that charter cities will not, at least at first,¹⁰⁷ be democratic, inhabitants of the charter city will have no way to press their interests in the running of the government or any comparable rights to self-governance beyond entry and exit. Some proponents of charter cities have claimed that they might eventually transition into liberal democracy, but as I claimed in §2c, libertarians have reason to be hostile to such an idea. Since consent does all the normative work in the moral justification of the libertarian state, all that is required for the charter city government to be morally justified is that its inhabitants have freedom of entry and exit and that its government respect a minimal set of libertarian rights. Inhabitants are thus asked to enter into an alienation social contract since they do not enjoy any self-governance rights in the

¹⁰⁴ Dorfman and Harel (2013).

¹⁰⁵ Dorfman and Harel (2013, p. 95).

¹⁰⁶ *Ibid.*

¹⁰⁷ E.g., Fuller and Romer talk about an eventual transition to democracy (2012, p. 10).

charter city; they relinquish them through consent. Nozick explicitly accepts this view when he writes that

“[I]f someone starts a private “town” on land whose acquisition did not and does not violate the Lockean proviso, persons who chose to move there or later to remain there would not have a right to a say in how the town was run, unless it was granted them by the decision procedures for the “town” which the owner had established.”¹⁰⁸

It’s worth noting, at this point, that libertarians reject inalienability of particular rights ‘all the way down,’ so to speak, since in addition to the alienation social contract (what might be called a political alienation contract), libertarian society also permits personal alienation contracts (such as the voluntary slave contract). In other words, persons are free to contract themselves into slavery and such contracts are enforceable by the minimal libertarian state. Nozick explicitly states that persons can fully relinquish their rights of self-ownership through consent when he asks “whether a free system will allow [a person] to sell himself into slavery.” He answers in the affirmative: “I believe that it would.”¹⁰⁹

The claim that some basic rights, such as those to self-governance and self-ownership, can be alienated even with consent and through contract represents a stark departure from standard liberal thought. This rejection of the alienability of particular rights (call them inalienability arguments) has a long history stretching back to at least Spinoza, but perhaps one of first classical liberal treatments of the inalienability of some rights comes from Adam Smith’s teacher, Francis Hutcheson. Hutcheson claimed that:

“Rights are also divided into the alienable, and such as cannot be alienated or transferred. These are alienable, where the transfer can *actually be made*, and where some interest of society may often require that they should be transferred from one to another Unless both these qualities concurr, the Right is to be deemed unalienable. ’Tis plain therefor, for instance, that for defect of both these qualities, our opinions in matters of Religion and worship are unalienable; and so are our internal affections of devotion; and therefor neither of them can be matters of commerce, contract, or human laws.”¹¹⁰

Hutcheson, therefore, presents two conditions for the alienability of a right: that it (1) is *actually* transferrable (call this the non-transferability condition), and that (2) there is some overriding

¹⁰⁸ Nozick (1974, p. 270).

¹⁰⁹ Nozick (1974, p. 331).

¹¹⁰ Hutcheson (1747).

social interest for that right *to* be transferred (call this the social interest condition). Liberty of conscience fails these conditions, not least because of the fact of privacy of individual judgment deems it impossible for a person to give up their liberty of conscience.¹¹¹

Inalienability arguments tend to emphasize the first condition by positing that certain rights are inalienable because they simply cannot be transferred. Consider personal alienation contracts, which most liberal views take to satisfy this non-transferability condition. Contracting with another party to sell oneself into slavery asks the contracting parties to recognize one party of the contract as something they are not: a non-person, or a piece of property. But a person is a person and not a piece of property, and no amount of consent can change this basic fact. A contract through which someone sells themselves into slavery, then, is a kind of contract fraud because, as Kant claims, “[n]o one can bind himself by a contract to the kind of dependency through which he ceases to be a person, for he can make a contract only insofar as he is a person.”¹¹² From this view, libertarian society in which people are able to make personal alienation contracts which are recognized by the legal apparatus of the state is a society in which such contract fraud is institutionalized.¹¹³

Liberal inalienability arguments against personal alienation contracts are rooted in the fundamental equal moral status afforded to individuals *qua* individuals that lies at the heart of the liberty tradition. This moral equality is what drives the normative individualism and the separateness of persons requirement that flows from this appreciation of normative individualism. Some classical liberals, such as James Buchanan, have referred to this as a “natural equality.”¹¹⁴ From this view, persons ought to be treated as moral equals in the organization of society because any individual differences that may exist between them are simply normatively irrelevant.¹¹⁵ Libertarian societies in which inalienable rights are treated as alienable with consent and through contract offend this basic moral equality because the state recognizes (and asks the citizens of the state to recognize) some people as less than people, indeed, as property, or as Kant puts it, as having “enter[ed] into the class of domestic animals.”¹¹⁶ Citizens of the state are asked to recognize persons as less than persons on this view *not* because someone has consented to being treated like a slave, but because the personal alienation contract that posits the fiction that one party is a piece of property is enforceable by the entire legal apparatus of the state.¹¹⁷

This liberal class of inalienability arguments generalizes from the personal alienation contract to the political alienation contract, since the political alienation contract involves the surrendering of what are taken to be inalienable rights: self-governance rights. Political alienation contracts undermine the moral equality of persons because they give some people a place in a

¹¹¹ Ellerman (2014, p. 15).

¹¹² Kant (1797, p. 98).

¹¹³ Ellerman (2010).

¹¹⁴ Buchanan (2005, p. 5, 67).

¹¹⁵ Buchanan (2005, p. 68).

¹¹⁶ Kant (Perpetual Peace).

¹¹⁷ Freeman (2001, p. 111 – 112).

social hierarchy over and above others; in the case of charter cities, inhabitants of the libertarian charter city—most probably the workers—are a part of a political class that has no say in the administration of the charter city or in the governance structures under which they live. They are necessarily placed lower, on the social hierarchy, than the private political powers that govern the libertarian charter city. Such an arrangement is at odds with liberal thought:

“[The] postulate of natural equality places the classical liberal directly and specifically at odds with all those who, explicitly or implicitly, accept the Platonic postulate of natural hierarchy. To Plato there are natural slaves and natural masters, with the consequences that follow for social organization, be it economic or political. To Adam Smith, by contrast, who is in this as in other aspects the archetype classical liberal, the philosopher and the porter are natural equals with observed differences readily explainable by culture and choice.”¹¹⁸

And the implications for accepting this natural equality are quite clear for Buchanan:

“The postulate of natural equality carries with it the requirement that genuine classical liberals adhere to democratic principles of governance; political equality as a necessary norm makes us all small ‘d’ democrats.”¹¹⁹

Accepting this natural equality thus means respecting a form of governance in which there are no natural masters and no natural slaves. This allows persons to *delegate* their self-governance rights but not *alienate* them, since to alienate self-governance rights is to, in effect, enter into a master and slave contract. And no amount of consent can change the fact that, as individuals, there are no masters and there are no slaves. Thus, a government is morally justifiable, on this classical liberal view, only if it respects the natural equality between persons and is structured around the implications for social organization that flows from this respect.

4. *The colonialism objection*

One of the principal criticisms levied against charter cities is that they are an instance of one morally problematic political relation of the past: that of the colonizer and the colonized.¹²⁰ In response to this line of criticism, proponents of charter cities typically emphasize the voluntary entry and freedom of exit enjoyed by inhabitants of charter cities. This does not answer the colonial critique for two reasons: first, in the non-ideal world, it’s not clear that the voluntariness of entry and

¹¹⁸ Buchanan (2005, p. 67).

¹¹⁹ Buchanan (2005, p. 69).

¹²⁰ See, e.g., <https://www.theguardian.com/science/2010/jul/27/paul-romers-charter-cities-idea>

freedom of exit is as ‘clean’ as would be required for the response to hold normative weight. But second, and perhaps most importantly, even if we do grant that residents have freedom of entry and exit, there seems to be something wrong with the basic associative structure between the governed and the government in charter cities, i.e., that the associative structure is based on the ostensibly voluntary alienation of the residents’ right to be ‘sovereigns’ or ‘principals’ in matters of social organization (to use James Buchanan’s turn of phrase).¹²¹ It is the *nature* of this basic associative structure that motivates the claim that charter cities represent an objectionable instance of colonialism. Before I turn to these two arguments, it is instructive to briefly outline Laura Ypi’s theoretical account of colonialism on which my critique draws.¹²² It is to this I turn to now.

(a) The wrongs of colonialism

Ypi’s account of what’s wrong with colonialism draws on Kant’s “cosmopolitan” critique of commercial and colonial practices as they were practiced by European colonial powers from about the sixteenth century onwards. Kant claims that each human being has a right of hospitality, i.e., a right to not be treated with hostility when they “present [themselves] for society.”¹²³ This “cosmopolitan right” of “universal hospitality” belongs to “all human beings by virtue of the right of possession in common of the earth’s surface.”¹²⁴ This right is limited to conditions of universal hospitality which makes it distinct from more expansive rights, such as the “right to be guest” in a foreign land. What this means is that all human beings have the right to present themselves on foreign lands and extend offers of political association without being treated with hostility (as long as the person presenting themselves also behaves “peacably”). Note that this right is about how the initial presentation and offer of political association is made: it governs the conditions that structure interaction between two people foreign to each other (or groups of people foreign to each other) without specifying what the outcome of the interaction out to be: “The other can turn him away, if this can be done without destroying him[.]”¹²⁵

The cases that Kant uses to illustrate violations of this cosmopolitan right are instructive in that they show another feature of his account that go beyond simply specifying just conditions as regards initial contact to also providing the outlines of an ideal political association that respects “the claims of all those involved in the exchange.”¹²⁶ For example, he claims that while people have the right to settle on any unclaimed land they please (and this right flows from the fact that all of humanity holds natural resources in common), a special contract is required if unclaimed open land is required by a nearby group for their sustenance (as is the case with hunter-gatherer or shepherding groups).¹²⁷ In cases such as these, settling in these areas “may not

¹²¹ Buchanan (2005, p. 23 – 24).

¹²² For an overview, see Ypi (2013).

¹²³ Kant, *Towards Perpetual Peace* (8:358).

¹²⁴ *Ibid.*

¹²⁵ *Ibid.*

¹²⁶ Ypi (2013, p. 175).

¹²⁷ Kant, *Towards Perpetual Peace* (6:353).

take place by force but only by contract that does not take advantage of the ignorance of those inhabitants with respect to ceding their lands.”¹²⁸ What is required in cases like these is a just political association that takes into account the varying interests at stake in such an interaction and potential settlement. Importantly, Kant argues, the need for these interests to be articulated and defended cannot be silenced by other supposedly overarching considerations as the ones that were used to justify the use of force during the height of Europe’s colonization efforts, such as force is necessary because “it is to the world’s advantage, partly because these crude peoples will become civilized ... and partly because one’s own country will be cleaned of corrupt men[.]”¹²⁹

As Ypi points out, Kant’s cosmopolitan critique rests on an analysis of colonialism as a departure of a political association of a particular sort, i.e., one that treats everyone in that association—*all* parties whose interests are at stake—in accordance to norms of reciprocity and equality. This ideal political association requires “political institutions that allow people to relate to each other as equals, guaranteeing that their voice will be heard and their claims will be equally taken into account when decisions affecting both are made.”¹³⁰ In concrete terms, what this means that the political association that results from two different groups of people coming into contact with each other must have equality and reciprocity built into its design; it is in this sense that colonial political relations are wrong because they are constituted by precisely the kind of political relations that violate these associative norms.

b. Consent

Proponents of charter cities respond to charges of colonialism by stressing the various choices involved in establishing a charter city. For example, Romer claims that,

“[C]harter cities are based entirely on voluntary actions. Only a country that wants to establish a charter city will do so. Only people who want to live and work under the rules specified in the city's charter will move there. Free choice is essential for the legitimacy of the rules in a charter city. It is also what makes a charter city very different from colonial occupation.”¹³¹

Others echo this sentiment.¹³² The problem with this response is that it seems Polyannish to think that the consent given will be as free and as voluntary as is required by the argument that claims that free and voluntary consent quells concerns of objectionable colonial political relations. Consent can be compelled, particularly when there is a large discrepancy between the relative

¹²⁸ *Ibid.*, 8:358.

¹²⁹ *Ibid.*, 6:353.

¹³⁰ Ypi (2013, p. 175).

¹³¹ Romer (2010), see here <https://www.cgdev.org/article/charter-cities-qa-paul-romer>

¹³² Freiman, e.g., writes that “[c]harter cities differ from colonial regimes because they do not coerce any party into participating” (2013, p. 4).

bargaining power of two parties negotiating a contract. Romer is sensitive to this fact and so explicitly claims that a “country that is subject to a military intervention has little true freedom of action and choice.”¹³³ Presumably, then, if the United States wants to be a foreign guarantor of a charter city in Belize and the government of Belize “consents” to opening up a charter city, under threat of military intervention, one can hardly say that Belize truly consented to opening up the charter city in any meaningful way. Indeed, even for Romer, then, that *would* look dangerously like a neo-colonial pursuit on the part of the United States.

But we need not go as far as to talk about military intervention with respect to the tools that states have to compel or coerce consent. Anybody with a cursory knowledge of international relations can attest to the fact that powerful states have a number of ways to put pressures on states, particularly in the cases of large discrepancies in power, where the term power does not only refer to military might. For example, policy tools might run the gamut of economic sanctions from boycotts, to trade restrictions or barriers, to embargoes, to full scale blockades. The United States is the largest export market and leading import source for countries such as Barbados and Belize; and if the United States decided to stop trading with Barbados and Belize, such a move would surely cripple those economies. And this problem is not limited to countries with relatively small economies: indeed, even countries with economies as large as Japan depend upon imports to provide enough food to sustain their population. To put it clearly, a globalized economy where countries play to their relative comparative advantage means that countries with large economies will be able to enact policies that might cripple other economies without having to revert to military intervention.

Moreover, charter cities will typically be proposed in places where there is a pressing need for capital investment. This presents a very real concern that host countries and charter cities will be reliant on the foreign guarantors in an objectionable way, since it is the fact that foreign guarantors guarantee the administration of the city and the credibility of its institutions that facilitate the investment in the first place. In cases such as these, consent may be given freely and voluntarily, but it is the relationship of dependence of the host country and charter city to the foreign guarantor that does the normative work in charges of objectionable subordination of a poor country (the host country) and the rich country (the foreign guarantor) that motivates the underlying charges of colonialism.¹³⁴

Finally, aside from ‘national’ consent issues, there are obvious problems with individual consent for the people who will end up populating and working within the charter city. Proponents of charter cities themselves pitch the idea as a solution for policy problems such as widespread corruption in the developing world and the large-scale migration from the developing world to the developed world. The people moving to charter cities, then, will likely be people escaping societies in which corruption or violence is rife, and it’s easy to see why moving to the

¹³³ See <https://paulromer.net/charter-cities-versus-humanitarian-military-occupation/>

¹³⁴ See Sagar (2016) for a similar point.

charter city might be done out of sheer necessity rather than free choice. The underlying point is that background conditions under which the choice is made might undermine whether the choice is made freely, and this fact might frustrate the normative work done by the notion of consent.

c. Domination

One theme of Kant's critique of colonialism is that colonial relations are characterized by domination and subjugation of one group over another. One way to avoid a political association characterized by domination and subjugation is to design institutions such that they reflect norms of reciprocity and equality. There are good reasons to suspect that charter city institutions do not reflect these norms.

Consider again the fact that charter cities rely explicitly on the idea that the people subject to the laws of the charter city do not have a say in the administration of key governance issues that are likely to arise, such as labor laws, taxation, and the general workings of the justice system. It should be emphasized that proponents generally see this as a *virtue* of charter cities, since what lends credibility to the new institutions is the promise of political stability and business friendly rules. Political stability in this case is secured by institutions that have reliable and stable rules that favor capital (i.e., investment), and these rules are not subject to democratic change.

The arguments that motivate the classical liberal critique of the libertarian position in §2d identify precisely what is objectionable about the political relationship prospective inhabitants of charter cities are asked to agree to. There's no need to repeat the arguments here, but suffice to say that, in a nutshell, charter cities represent an objectionable political relationship because they violate norms of equality when they ask inhabitants to enter into a contract through which they relinquish their basic rights to self-governance. Such a contract violates the moral equality afforded to persons because it means that some people are placed lower on the social hierarchy than others, and they are placed in a position such that they have no way to press their interests even with respect to decisions that will affect them. Conceiving of some rights as basic means that they cannot be overridden for precisely these sorts of considerations, i.e., that they will yield, on aggregate, desirable outcomes even if some people must bear a burden (of a loss of rights) in the short term.

Recall that proponents of charter cities tend to stress the consent of the inhabitants of charter cities as a response to the colonialism objection. An interesting feature of Ypi's account is that it deflates this response, since what drives the charge of colonialism does not turn *only* on whether individuals have consented to the charter city. There is something objectionable about the governance structure of the (non-democratic) charter city itself, and what makes this structure objectionable stands independently of whether the inhabitants have consented to it.

6. Pragmatic arguments

Proponents of charter cities often cite pragmatic considerations that they claim weigh heavily in favor of the establishment of charter cities. The literature contains what might be called feasibility arguments, that is, arguments that contain the claim that charter cities are more practicably able to deliver on the goals of development (i.e., increases in human welfare through prosperity-creating institutions) than other feasible alternatives. In this section, I highlight one potential problem with this view, namely, that it does not take into account the importance of informal rules with respect to economic development.

a. Informal institutions matter

Proponents of charter cities claim that one reason why they are more feasible in delivering their stated aims is because they provide a fresh set of institutions that promote economic growth which will in turn lead to vast improvements in welfare. It's best to think of 'institutions' here as sets of rules that promote economic growth. For example, it's generally well accepted that a system of rules that protects private property backed by effective coercive enforcement mechanisms will provide the right kind of incentives that induce sustainable and long-term investment. It's also generally well accepted that sustainable and long-term investment is a necessary condition for raising living standards. This requires a number of well-functioning institutions: financial institutions that are able to direct capital to its most productive uses; legal institutions that specify property rights and adjudicate property disputes fairly; and law enforcement agencies that enjoy a monopoly on violence which allows them to effectively enforce property rights. What is important about these institutions is that they must be legitimate in the descriptive sense, i.e., the populace that these institutions serve must have trust in these institutions such that they abide by their decrees and respect their judgments.

A lot turns on whether these institutions will be effective as the fundamental point underlying charter cities is that the primary constraint on development is the inability to create and implement the right sets of rules. Romer explicitly states that the "central task of reducing global poverty is to find ways for developing countries to adopt new rules that are known to work better than the ones they have."¹³⁵ The idea that Romer is getting is what is known as the institutional view of development, a view that is perhaps best defended by Daron Acemoglu and James Robinson.¹³⁶ The institutional view of development stresses the importance of economic institutions, such as institutions that protect property rights, to economic development; the best way to explain why one society is wealthier than another society is through analyzing their relevant institutions and the impact these institutions have on comparative development. The

¹³⁵ Romer (2010, p. 1).

¹³⁶ Acemoglu and Robinson (2012).

view states that what determines whether a society develops is whether it creates the right sorts of incentives for investment, and whether this happens turns on the sorts of incentives that are created by institutions (or sets of rules) that people in that society obey.

The thought is that charter cities promise to deliver these elsewhere-tested institutions by bypassing the domestic institutions of the local host countries as well as evading the collective action issues that inhibit domestic institutions from becoming prosperity-creating institutions. Broadly speaking, these collective action issues arise when individuals have a preference to act in accordance with good rules (e.g., act honestly) and it's in everyone's interest to do so, but each individual has an incentive to act dishonestly because they believe that others will also act dishonestly. Developing countries with widespread corruption are paradigmatic examples of this, where almost each individual has a preference to live in a high-honesty low-corruption society, but they are stuck in a 'bad equilibrium' of a high-corruption low-honesty society because acting honestly in such a society is costly relative to acting dishonestly (as I argued in Chapter 1).¹³⁷ Economic development is stunted in societies like these where people view economic interactions as a zero-sum game, levels of trust are low, and transaction costs are high.

Note that this is a comparative judgment about the relative likelihood of inducing compliance to prosperity-inducing rules between a new charter city's institutions against their domestic counterparts. Bryan Caplan captures this idea perfectly when he claims that,

“In principle, Third World countries could put nationalistic prejudice aside and ‘import’ the written and unwritten rules that have made the West rich. But this is extremely difficult. Intense populist opposition aside, it is hard to graft one country's institutions on to another's - especially when entrenched interests fight you every step of the way. This is true in the business world as well. Competitors often try and fail to adopt leading firms' ‘best practices.’ Corporate culture is notoriously stubborn. In both business and politics, success often requires a clean slate. It is easier to open a new WalMart than to make the Kmart chain better. Advocates of charter cities argue that is also easier to bring in “outside management” to make a new city that works than to reform existing countries that don't.”¹³⁸

But there are at least two reasons to question this line of reasoning. First, modern economies are characterized by a host of both formal and *informal* institutions. In many cases, what stunts economic growth in developing countries is not the “nationalistic prejudice” in importing the rules that have worked elsewhere, but rather the informal institutions that make compliance with the formal institutions costly, or the perverse incentives that are created at the intersection of

¹³⁷ Fisman and Golden (2017); see also Hoffman and Patel (2017).

¹³⁸ See Caplan, *A Quick Case for Charter Cities: Memo to the Gates Foundation* http://www.econlib.org/archives/2011/03/a_quick_case_fo.html

informal and formal institutions. Indeed, many developing countries have formal rules about the illegality of corrupt practices, property right protection, and the like, but there are a set of informal rules that render these rules redundant. To continue with Caplan's analogy, it's just not clear that opening a new WalMart but populating it with the people who used to work at Kmart is better than trying to change the culture—the informal institutions—at Kmart better.

The general point is this: what motivates the comparative judgment that a “blank slate” of institutions will be better able to deliver sustainable and long-term economic growth is that domestic institutions suffer from a number of collective action issues that stunt growth. But what drives these collective action issues are not the domestic institutions themselves, but the informal norms that underwrite the costs and benefits of compliance with these formal institutions. These informal rules will not suddenly disappear once they are presented with a “blank slate” of formal rules. This is well illustrated by the fact that pernicious social norms, such as forced marriage or honour killings, do not suddenly disappear from the communities in which they are practiced once those communities migrate to countries in which those practices are illegal. Think, for example, about the South Asian diaspora in the United Kingdom: forced marriages and honour killings persist in parts of the United Kingdom amongst the South Asian diaspora *even in the face of the credible institutions and under threat of enforcement of laws that ban these practices*.

One might respond to this line of reasoning by pointing out that there is evidence that formal rules also influence informal rules. This is the move that Christopher Freiman makes.¹³⁹ The problem with this response is this: recall that the comparative judgment about charter cities being better than domestic institutions is that they promise growth-friendly rules that the domestic institutions lack. But the domestic institutions typically do not formally lack growth-inducing rules; they have a set of unwritten rules that renders them redundant. If it were the case that formal rules on their own would be able to “control corruption, secure personal property, and stabilize expectations,” or influence informal institutions to the degree required to do these things, then we wouldn't need the charter cities in the first place. In the end, comparative judgments about the relative efficacy of bad domestic institutions of host countries against the good foreign institutions of charter cities smack of a kind of selective idealization: proponents of charter cities tend to idealize institutions promised by the charter city but are non-idealistic about the domestic institutions. But the same reasons that motivate the suspicion of changing bad domestic institutions also seem to apply in the case of new foreign institutions.

Of course, proponents of charter cities are sensitive to the importance of informal institutions to the functioning of modern, well-functioning economies. Indeed, Freiman himself writes that “development economists stress that we cannot discount the contribution of *informal* norms and customs to a society's stability, efficiency, and cohesion” and, as such, the “indispensability of informal norms should temper our enthusiasm about the ability of charter

¹³⁹ Freiman (2013, p. 45).

cities' formal institutions to bring about prosperity by themselves."¹⁴⁰ My point here is that, taking the indispensability of informal norms seriously means that we should *question*, rather than just *temper*, our enthusiasm with respect to charter cities.

6. Conclusion

In §2, I introduced a contrastive taxonomy of types of cities in order to motivate my understanding of the paradigmatic charter city. In §3, I discussed the Hobbesian social contract before outlining the minimal commitments of classical liberalism and libertarianism. I showed that the inalienability of rights, the provision of public goods, and separateness of persons constraints hang together normatively in both views, but the way they relate to each other differ significantly on either view. In §4, I outlined Ypi's account of colonialism and showed that the arguments in this paper, taken with Ypi's account, support the colonialism objection against charter cities. I also showed that the standard response from proponents of charter cities (i.e., to stress consent) fails against this objection. In §5, I showed that pragmatic arguments in favor of charter cities are unconvincing because they do not take into account the informal institutions that drive the pernicious practices that proponents of charter cities wish to eradicate.

¹⁴⁰ Ibid.

CHAPTER 3. The natural and the social in the metrics of justice

1. Introduction
2. Resourcism and the capabilities approach
 - a. Quality of life determinants
 - b. Purely personal heterogeneities
 - c. Compensation claims for natural inequalities
3. The status of natural inequalities
 - a. Natural and social causes of inequality: a view from disability studies
 - b. Interactionalism all the way up: gender and norms
4. Measuring global justice
5. Conclusion

I. Introduction

Should an appropriate metric of justice be sensitive to internal individual differences? Rawlsian resourcists claim that the answer to this question is no.¹⁴¹ These resourcists claim that persons are entitled to a kind of standardized package of resources provided by the basic structure of society as a matter of justice.¹⁴² On this view, individuals might still appeal for a greater share of resources in virtue of a relevant individual difference (such as being born blind), but this appeal cannot take the form of a demand of justice.¹⁴³ Capability theorists claim the answer to this question is yes. These capability theorists insist that the provisions of the basic structure of society ought to be sensitive to the way individual differences interact with forms of social organization and other what might called “environmental” factors.

Whether an appropriate metric of justice should be sensitive to internal individual differences rests in part on whether principles of justice ought to be sensitive to these internal individual differences. Further, whether principles of justice ought to be sensitive to internal individual differences is also informed, at least in part, by how we conceive of these differences. How these questions are resolved rely on an underlying understanding of a natural/social inequality distinction, one that is much invoked in theories of justice.¹⁴⁴ Prominent resourcists

¹⁴¹ There are many different conceptions of resourcism. In this paper, I focus exclusively on Rawlsian resourcism unless otherwise explicitly stated. The position I have in mind is articulated by Rawls himself (1999), Pogge (2002, 2010), and Freeman (2006, 2018).

¹⁴² The capabilities approach, or capability theory, also has much “internal diversity.” I aim to set aside this internal diversity as much as possible and instead try to focus on features of the approach common to all, or most, of its proponents. The position I have in mind is articulated by Sen (1999), Nussbaum (2011), and Robeyns (2017).

¹⁴³ Pogge (2010).

¹⁴⁴ See Lewens (2010) for good discussion on this point.

hold that the distribution of natural endowments (innate traits and dispositions) is not just or unjust in itself. This provides motivation for the view that a metric of justice ought not to be sensitive to particular types of individual internal differences, since a distribution that is neither just nor unjust does not require us to rectify or justify the inequalities that arise from it.

The literature on resourcism and capability theory, and the central disputes between them, is vast. I do not attempt to square away all issues here. Rather, I focus narrowly on a critique of the capabilities approach levied by prominent resourcists which turns on how the respective theories treat “natural” inequalities in their preferred institutional orders. I call this the *compensation for natural inequalities critique*:

- (1) *Compensation for natural inequalities critique*. The capability approach metric is committed to compensatory claims of justice for distributions of natural endowments—distributions that are, on the resourcist view, neither just nor unjust.¹⁴⁵

In this paper, I argue that capability theorists are able to accept both that distributions of natural endowments are neither just nor unjust while also maintaining that a theory of justice’s metric ought to be sensitive to individual variation (even when this variation has, at bottom, a “natural” (innate, biological) component). This is because capability theorists are able to adopt a more sophisticated understanding of the natural/social distinction, one that stresses the *interaction* between the two through the mediating role of both formal and informal institutions. What lies at the heart of this understanding is a particular conceptualization of the nature of disability, namely, the interaction model.

The arguments in this paper do not conclusively settle the debate between resourcists or capability theorists. My aims are more modest. Instead, the contributions of this paper are (1) to pinpoint the where the two approaches disagree but also, importantly, where they agree; (2) to introduce some of the complexities of the natural/social distinction, for while it is much-invoked across political philosophy, it is often drawn in simple causal terms at odds with our best understanding as provided by theoretical and empirical developments in both disability studies and biology; and (3) to relate this theoretical dispute about the appropriate metric of justice to the more practical concerns regarding how we ought to make comparative evaluations (e.g., of poverty and other welfare-outcomes) across countries. The goal of the arguments in this paper is not to tip the hat decisively toward the resourcism or the capabilities approach, but rather to clarify the nature of the dispute and move the debate forward on common ground.

In §2a, I briefly outline resourcism and the capabilities approach, paying special attention to the the wide agreement between the two approaches in their more sophisticated forms.¹⁴⁶ §2b

¹⁴⁵ Pogge (2010).

¹⁴⁶ I take Rawlsian resourcism, with its focus on social primary goods, as a more sophisticated version of the resourcist position. This view might be contrasted with what might be thought of as cruder forms of resourcism, such as those that focus exclusively on, e.g., income metrics as a matter of justice.

attempts to pinpoint the precise point that the views putatively diverge, namely, with respect to the distribution of natural endowments (and the natural inequalities they may give rise to). §2c then formulates and explicates the resourcist's central critique of the capability theorist's metric: the compensation for natural inequalities critique. §3a shows that the capability theorist might adopt an understanding of the nature of disability, the interaction model, that largely deflates the resourcist critique. §3b relates the underlying characteristics of the interaction model, its focus on the interaction between putatively natural traits and their environment, to broader understandings of the role that formal and informal institutions play in development outcomes. I pay special attention to the problem of gender in this subsection. §4 then shifts gears from a focus on what kinds of inequalities might trigger compensation claims under the two respective views to how their respective metrics fare with respect to comparative evaluative judgments of welfare-relevant outcomes across countries. §5 concludes.

2. Resourcism and the capability approach

A central point of contention between resourcists and capability theorists concerns the kind of metric that should be endorsed as a public criterion of social justice—a criterion that is to be used to make comparative evaluations of individual advantage. The public criterion of any theory of justice is of significant practical import, since it is precisely the public criterion that forms the basis of evaluating how feasible alternatives to institutional arrangements measure up to the demands of justice (however they end up being specified).

Resourcists claim that the appropriate metric of justice ought to be defined in terms of social primary goods,¹⁴⁷ which Rawls specifies as follows:

- (1) The basic liberties (freedom of thought and liberty of conscience, etc.) are the background institutions necessary for the development and exercise of the capacity to decide upon and revise, and rationally to pursue, a conception of the good. Similarly, these liberties allow for the development and exercise of the sense of right and justice under political and social conditions that are free.
- (2) Freedom of movement and free choice of occupation against a background of diverse opportunities are required for the pursuit of final ends as well as to give effect to a decision to revise and change them, if one so desires.
- (3) Powers and prerogatives of offices of responsibility are needed to give scope to various self-governing and social capacities of the self.

¹⁴⁷ Rawls (1993, p. 310).

- (4) Income and wealth, understood broadly as they must be, are all-purpose means (having an exchange value) for achieving directly or indirectly a wide range of ends, whatever they happen to be.
- (5) The social basis of self-respect are those aspects of basic institutions that are normally essential if citizens are to have a lively sense of their own worth as moral persons and to be able to realise their highest order interests and advance their ends with self-confidence.

According to Rawls, social primary goods are those goods that are reflective of “various general facts about human needs,”¹⁴⁸ and, as such, are requirements of free and equal individuals living a complete life. They are best thought of as *valuable resources* that a just political order affords its citizens access to. It is the difference in individuals’ (reasonable) expectations of primary goods over a lifetime that constitute the inequalities that generate justice-relevant advantage and disadvantage, and so are the normatively salient for comparative assessment of various institutional schemes. From this view, a person is the least advantaged in a well-ordered society—a society in which their basic rights and liberties are secure, and they enjoy fair opportunities—in the case that they “belong[] to the income class with the lowest expectations [of primary goods].”¹⁴⁹ The implication is that alternative political arrangements should then be judged in terms of individuals’ *expectations of* (as Rawls puts it), or their *access to* (as Pogge puts it), the valuable resources constitutive of social primary goods.

Capability theorists, on the other hand, claim that the appropriate metric of justice ought to be defined in terms of what they take to be substantive freedoms—what they call capabilities—that one would need in order to choose a life one has reason to value.¹⁵⁰ Rather than focus on the expectations of or access to the external resources required to exercise those freedoms, the capability theorist shifts the normative focus to a person’s ability to meaningfully promote their own ends given the resources they are entitled to. This necessarily implicates each individual’s ability to convert social primary social goods into the functionings or capabilities that constitute their own ends.

This disagreement that gives rise to the difference in choice of metric is about the extent to which principles of justice ought to be sensitive to internal differences, broadly defined as “natural endowments.” These natural endowments include internal qualities like the set of genes a person is born with. The resourcist view calls for a kind of standardized package of social primary goods to be provided for as a matter of justice, which is based on a “selection, formulation, and weighting of valuable resources on some account of *standard* human needs” (emphasis in original).¹⁵¹ Capability theorists claim that provisions provided by the basic structure

¹⁴⁸ Rawls (2005, p. 283).

¹⁴⁹ Rawls (2001, p. 49).

¹⁵⁰ Sen (1999); Nussbaum (2011); Robeyns (2017).

¹⁵¹ Pogge (2010, p. 27).

of society, in the service of securing important capability sets (sets of functionings), ought to be sensitive to individual variation. This is because different persons might require different amounts of resources to secure the same capability or set of capabilities. Pogge sums up the central difference: “[c]apability theorists assert, while resourcists deny, that a public criterion of social justice should take account of the individual rates at which persons with diverse physical and mental constitutions can convert resources into valuable functionings.”¹⁵²

One reason why capability theorists take their account to be superior in terms of its public criterion is that it fares better at capturing key determinants of a quality of life that, so capability theorists allege, resourcist accounts are not—and, in principle, cannot be—sufficiently sensitive to. That resourcist accounts fail to capture these key determinants of a quality of life *in principle* is an important caveat since capability theorists sometimes contrast their views with impoverished resourcist accounts when more sophisticated ones are readily available. In what follows, I provide a brief overview Sen’s original proposed list of key determinants and contrast capability theory with what I take to be a sophisticated resourcism so as to accurately reflect the real divergence between the two positions, as well illustrate the significance of this divergence.

- (1) *Environmental diversity*. Variation in the environment (such as temperature and likelihood of disasters) will levy different costs and benefits subject to what Sen refers to as specific “climatic circumstances.”¹⁵³ For example, costs to maintain comfortable temperatures and reasonably manage environmental risk will be higher in some parts of the world than in others; Alaskans spend more on heating than do Arizonans and Mississippians spend more on flood insurance than do New Mexicans. If the resourcist’s standardized package is not sensitive to this variation, then Alaskans and Mississippians will bear more of a burden to have a comparably similar quality of life to their Arizonan and Mississippian counterparts simply because more of their income is spent on mitigating the externalities of environmental variability.

On this point, the capability theorist and resourcist contrast most sharply if we assume that the resourcist’s standardized package admits of only a fixed, non-variable income. But if we take the *standardized human needs*, rather than the income required to achieve those needs, as fixed, it becomes clear that (at least some level of) sensitivity to environmental variation is easily built into a plausible resourcist account. This is because what is troublesome about problematic climatic circumstances in question, from the resourcist view, is that they force individuals who live under those conditions to expend more resources to have access to the “standard” level of resources (e.g., expenditure on a coat to achieve a temperature within a normal range in which human beings can

¹⁵² Ibid., p. 29.

¹⁵³ Sen (1999, p. 149)

comfortably work, to use Pogge's example).¹⁵⁴ In other words, a person's expenditure on "extra" resources in order to achieve access to the level of resources they *ought* to have access to should properly be included in the overall inter-personal evaluative accounting.

- (2) *Social climate*. Salient aspects of the social conditions in which people live, such as "public educational arrangements ... prevalence or absence of crime and violence ... [i]ssues of epidemiology and pollution ... [and] the nature of community relationships"¹⁵⁵ will play a significant role in the way personal incomes and resources are converted into quality of life. One way to capture this concern is to think about the role that public goods (such as a basic level of education, low crime, a clean environment, a healthy population, and a sufficient population-wide level of social capital) play in allowing individuals to convert personal resources into what they take to be a life worth living. Access to a standard package of personal resources does not mean much in terms of quality of life if, for example, one lives in a high crime area where property rights are routinely violated.

Resourceists might respond to this concern in two ways. First, a society which lacks the provision of a basic set of public goods—the sort that would give rise to the set of concerns Sen is raising here—would violate its citizens basic liberties, liberties that a standard resourceist account ought to be sensitive to.

Second, sophisticated resourceists view the provision of public goods as constitutive of a well-ordered society since justice itself is not a normatively neutral cooperative enterprise, but rather a "social union"¹⁵⁶ in which "people have their shared ends and regard other members' success and flourishing as complementary to their own."¹⁵⁷ A "private society" in which "each person assesses social arrangements solely as a means to his private aims ... [and] no one takes account of the good of others, or what they possess"¹⁵⁸ is unjust on this view, even if such a society succeeds in the provision of public goods through the selfish contributions of its citizens. The upshot is that a society which lacks provision of basic public goods is incompatible with both the resourceist and capability approach account, even if the underlying rationale for the incompatibility may differ.

- (3) *Relational perspective*. The advantage and disadvantages individuals face may turn on specific community-wide factors, such as others' resource levels. For example, the poorest Americans may be many times wealthier than the poorest Nigerians in absolute terms, yet

¹⁵⁴ Pogge (2010, p. 23).

¹⁵⁵ Sen (1999, p. 150)

¹⁵⁶ Rawls (1999, p. 456)

¹⁵⁷ Miller (2012, p. 72); Rawls (2001, p. 529).

¹⁵⁸ Ibid.

they may nonetheless face very real capability deprivation (such as being unable to participate in their community's civic and political life).¹⁵⁹

Resourcists are easily able to accommodate this relational perspective by acknowledging that the value of an individual's personal resources is often dependent on the value of the relevant community's resources.¹⁶⁰

(4) *Interfamilial distribution*. The conversion of social primary goods into freedoms enjoyed by individuals is somewhat complicated if the basic unit of analysis is the family. It is often the case, especially in patriarchal societies, that men are compensated for their labor in a normal market exchange whereas women—undertaking “invisible work” within the family, such as caretaking and domestic labor—are not. A metric that focuses exclusively on the income generated by the family unit alone misses the complications that arise with the problem of *intrafamilial distribution*, and the (typically gendered) inequalities they bring the fore.

Since resourcist accounts generally fall within the liberty tradition outlined in Chapter 2, they all share a commitment to normative individualism (typically interpreted as a separateness of persons requirement). Separateness of persons means that we ought to respect the moral status of individuals *qua* individuals, and that it is an *individual's* goals, values, freedoms, and so on, that are of intrinsic importance. In other words, taking the familial unit as the unit of analysis for the conversion of primary goods into freedoms is at odds with a fundamental commitment of the resourcist account.¹⁶¹

On the four key determinants mentioned so far—environmental diversity, social climate, relational perspective, and interfamilial distribution—both the resourcist and the capability approach offer adequately sensitive accounts. Sen's critique of resourcism is rather a critique of a rather crude sort of resourcism, one that focuses income measures as proxy for welfare assessments. As Rawls's list of social primary good shows, resourcism can (and does) go well beyond that, even if non-governmental multinational institutions do not.¹⁶² But the key divergence between the capability approach and these more sophisticated forms of resouricism, and therefore the site of the “real contrast,” as Pogge puts it, is how the two views treat what Sen calls *personal heterogeneities*. It is to this I turn to now.

¹⁵⁹ Sen (1999, p. 151).

¹⁶⁰ Pogge (2010, p. 21).

¹⁶¹ *Ibid.*

¹⁶² For a critique normative benchmarks (e.g., a “dollar a day” for poverty assessments) by the World Bank, etc., see Robeyns (2005).

b. Purely personal heterogeneities

(5) *Personal heterogeneities*. People's needs are diverse in light of "disparate physical characteristics connected with disability, illness, age or gender, and these make their needs diverse."¹⁶³ A metric of justice must be sensitive to differences that arise due to personal heterogeneities, so the capability theorists argue, because differences in these sorts of characteristics will affect the conversion of social primary goods (such as income) into the "real" freedoms that people ought to be able to enjoy.

The capability theorist critique is that the resourcist framework, because of an indifference to the variance in conversion rates of primary goods and resources into whatever good that is valued, remains indifferent in cases where such conversion rates matter importantly with respect to justice: "Of two persons with identical incomes and other primary goods and resources (as characterized in the Rawlsian ... framework[]), one may be entirely free to avoid undernourishment and the other not at all free to achieve this."¹⁶⁴ And the resourcist account fails to be sufficiently sensitive to a set of cases in which differential conversion rates affects how the ends of the 'standard package' of Rawlsian resources are *actually* realized, because the standard package—being what it is—remains entirely indifferent to.¹⁶⁵

On the one hand, this critique does in fact point to a central disagreement between the two views for, as Pogge claims, resourcists, *contra* capability theorists, "define and consider individual shares without regard to the particular features of the persons whose shares they are."¹⁶⁶ The implication is that the defining the resourcist metric consists in "weighting selected [primary social] goods relative to one another ... guided by some conception of the *standard* needs and endowments of human beings,"¹⁶⁷ rather than the specific disadvantages felt by individuals' (in light of their natural endowments) ostensibly required by a metric endorsed by capability theory.

On the other hand, we must be careful not to exaggerate the difference in terms of sensitivity between the two views. This is because the resourcist account can in fact accommodate a set of cases that the capability theorist has in mind—a set that is, admittedly, smaller than the broader set of cases that fall under "personal heterogeneities," but which nonetheless goes some way in deflating the capability theorist's critique. Sensitivity to this particular set of cases is triggered in light of two principles that are compatible with sophisticated resourcist accounts: the *normative individualism principle* (mentioned above) and the *prior wrongdoing principle*. Since

¹⁶³ Sen (1999, p. 148).

¹⁶⁴ Sen (1992, p. 33).

¹⁶⁵ *Ibid.*

¹⁶⁶ Pogge (2010, p. 35)

¹⁶⁷ *Ibid.*

sophisticated resourcists might accept these two principles as core commitments of their view, they are able to go quite a way to deflate the capability theorist's critique.

Consider first the normative individualism principle. Once we accept normative individualism, any "standard package" must also be sensitive to the fact that a standard package catered for a *single* adult human is not adequate for, say, a mother and child—whether during pregnancy or breastfeeding.¹⁶⁸ Further, future children have pre-birth needs, these needs give rise to extra needs for a pregnant woman.¹⁶⁹ Children, in other words, are individuals and generate their own justice-relevant needs and claims, and this fact (along with its metric-relevant implications) are quite easily accommodated in light of the resourcist commitment to the normative individualism principle.

Second, the prior wrongdoing principle reflects the fact that "the members of a society or other social system have a duty of justice to mitigate the harms caused by their prior wrongdoing."¹⁷⁰ Sensitivity to at least some set of the cases that capability theorists have in mind are captured by this principle in the case that the severe mental or physical disabilities that give rise to specific special needs are caused by the "effects of severe past (and present) resource deprivation: lack of effective civil and political rights and inadequate access to water, food, shelter, health, care and education."¹⁷¹ The prior wrongdoing principle allows the resourcist to distinguish between justice-relevant and justice-irrelevant disabilities, where the former is implicated into the resourcist metric whereas the latter is not. Justice-irrelevant needs and claims are those that arise because of what Pogge calls *pure* personal heterogeneities—inequalities that arise due to differences in natural endowments, self-caused factors (such as drug abuse), or pure luck. Inequalities caused by these kinds of factors do not give rise to claims of justice. Rather, they are viewed as misfortunes of a "natural lottery,"¹⁷² one which society does not have a *justice-relevant* duty to rectify or remedy even if there may be "quite stringent" duties of humanity or solidarity.¹⁷³ The prior wrongdoing principle reflects the view that heterogeneities that give rise to compensation claims couched in terms of justice are justice-relevant in virtue of their causal origins. A person whose disability is the result of some prior wrongdoing on the part of society (e.g., injury on a public transport system) has a justice-relevant claim whereas person whose equivalent disability is the result of, e.g., bad genetic luck does not.

c. *The compensation for natural inequalities critique*

In the last section I claimed that the resourcist responds to the capability theorist by pointing out that a set of cases—those captured by recognition of the *normative individualism principle* and the *prior wrongdoing principle*—can be accommodated under their account, and therefore are included under the insensitive-to-internal-difference metric they endorse. And so it is the set of

¹⁶⁸ Ibid., p. 36.

¹⁶⁹ Ibid.

¹⁷⁰ Ibid., p. 40

¹⁷¹ Ibid.

¹⁷² Ibid., p. 45.

¹⁷³ Ibid., p. 44.

cases that fall outside of the grasp of these two principles where the real disagreement lies, that is, in cases where the inequalities between persons are caused by differences in natural endowments and that have their causal origins in pure personal heterogeneities.

These so-called natural inequalities generate no claim to justice because society “did not contribute to their emergence and does not benefit from their existence,”¹⁷⁴ and so, from the resourcist point of view, it is a *virtue* of the account that their metric is able to exclude them. One reason for taking this as a virtue is because of the institutional implications of the divergence of the views: resourcists “seek an institutional order under which the distributive pattern of resources access satisfies their preferred criteria” and, allegedly unlike the capability theorist, “they pay no attention to how this distribution correlates with the distribution of natural features.”¹⁷⁵ This line of reasoning forms the underlying rationale for the central resourcist critique of the capability theorist:

- a) *The compensation for natural inequalities critique.* The capability approach metric is committed to compensatory claims of justice for distributions of natural endowments—distributions that are, on the resourcist view, neither just nor unjust.

Compensatory claims of justice here refer to complaints that aim to compensate for disadvantages that might be chalked down to natural inequalities. These are distinct from what Pogge refers to as *intrinsic discrimination complaints* and *extrinsic discrimination complaints*. Intrinsic discrimination complaints are complaints against an institutional order that overtly or covertly disadvantage persons based on ostensibly natural characteristics (e.g., race or sex) through the execution of its rules. Extrinsic discrimination complaints are complaints against an institutional order that fails to compensate individuals disadvantaged because of historically discriminatory rules (e.g., failure to pay reparations for past injustices).¹⁷⁶

Resourcists claim that institutional schemes inspired by the capability approach run into an attendant set of problems in light of its commitment to compensation claims for natural inequalities:

- b) *The vertical inequality problem.* By mandating an institutional arrangement in under which distributions of natural endowments are taken to be justice-relevant, the capability approach conceptualizes of natural inequalities (and natural human diversity) in *vertical* terms. Thus, the capabilities approach views human beings as “better or worse endowed.”¹⁷⁷

¹⁷⁴ Pogge (2010, p. 44).

¹⁷⁵ Ibid., p. 45.

¹⁷⁶ Ibid.

¹⁷⁷ Pogge (2010, p. 56).

- c) *The imprecision problem.* By mandating an institutional arrangement in under which distributions of natural endowments are taken to be justice-relevant, the capability approach fails to specify a criterion of justice aimed at achieving “equity among their beneficiaries ... and equity also among their contributors.”¹⁷⁸
- d) *The demandingness problem.* By mandating an institutional arrangement under which distributions of natural endowments are taken to be justice-relevant, the capability approach generates the “highly demanding” and psychologically unrealistic requirement that “persons to share the fruits of their cooperative efforts equally with those who do not participate.”¹⁷⁹

These four concerns are, of course, intimately linked. A natural way to read the above is that *because* of (a), *then* (b), (c), and (d). For example, (b) vertical inequality is, on this view, but a symptom of an institutional arrangement that unduly generates compensation claims for distributions that are not just nor unjust in themselves. Inequalities borne of natural endowments (“purely natural inequalities”) do not generate justice relevant disadvantages, and therefore do not stand in need of rectification or justification by a minimally just institutional arrangement.

Further, (c) imprecision and (d) demandingness might appropriately be regarded as two sides of the same coin. Imprecision here refers to the precision with which the normative rules specify *who* gets *what*, and *who* pays for it—the beneficiaries and contributors to whatever it is that is the subject of those rules. The administration of these rules simply *is* the administration of justice, since it is these rules that take from some and give to others (under threat of legitimate coercion) as a means to achieve an overall just distribution. In any egalitarian theory of justice, which of course includes both resourcist and capabilities approach views, a specified relevant set of inequalities ought to be justified or rectified according to the normative rules specified by the theory in question. Since the conceivable natural inequalities that exist between individuals may be—and probably are—many, their rectification would impose severe obligations on the non-beneficiaries of that rectification. Imprecision means under-specification of the relevant rules (e.g., distributive rules), thus resulting in a theory of justice that generates (many) demands in the form of compensation claims on behalf of some beneficiaries at great cost, if realized, to other contributing non-beneficiaries.

In what follows, I leave aside the problems of imprecision and demandingness, and focus instead on the central critique (a) and the vertical inequality problem (b). But it is worth mentioning two points before moving on here. First, the arguments in this paper rely on an underlying presumption that capability theorists are interested in an institutional scheme that does not take purely natural inequalities themselves as a matter for our justice-relevant

¹⁷⁸ Ibid., p. 61.

¹⁷⁹ Ibid., p. 67.

obligations. Capability theorists need not accept this presumption, however, and instead endorse a type of radical egalitarianism whereby natural inequalities *per se* do in fact give rise to claims of justice. One way to read the resourcist's critique is in these terms, i.e., that capability theorists are committed to a radical egalitarianism without the theoretical resources to delineate between inequalities appropriate for rectification or in need of justification (e.g., those arising from artificial restrictions on wages) and those that are not (e.g., individuals who are insufficiently attractive to attract a sexual partner) in a principled way.¹⁸⁰ It seems clear that imprecision and demandingness would be most acutely felt in an institutional scheme that reflected this kind of radical egalitarianism. I take this in itself to be a compelling reason to reject this kind of radical egalitarianism. I also take it that capability theorists are not necessarily committed to such a view. Indeed, some capability theorists, such as Elizabeth Andersen, explicitly reject the view.¹⁸¹

Second, despite the fact that I leave aside the problems of imprecision and demandingness, they are still relevant for the overarching arguments in the paper. This is because what I aim to show in the next section is that some inequalities that resourcists take to be purely natural have a justice-relevant social component, therefore raising them to candidacy for legitimate government intervention (in terms of either rectification or in need of justification). But whether these inequalities should in fact be rectified would require a balancing of the justice-relevant practical concerns that are part-and-parcel of policymaking, such as the distribution of costs and benefits of the government action, which I take as the concerns that the problems of imprecision and demandingness raise to the fore. The argument, in other words, is not that showing that an inequality has a justice-relevant social component means that that inequality *must* be rectified or justified, but rather that the question is not a closed one as if it would be if the inequality was shown to be a purely natural inequality in a way that renders it justice-irrelevant.

4. The status of natural inequalities

a. *The natural and the social: the interaction model*

One way capability theorists might respond to the resourcist critique is by denying that they are indeed committed to compensatory claims for “purely” natural inequalities. Rather, armed with a more sophisticated understanding of underlying distinction between natural and social inequalities, they might further claim that at least some of the cases resourcists have in mind as purely natural are actually social and therefore trigger legitimate compensation claims under a just institutional scheme—on *both* the resourcist *and* capability theorist understandings of what a just institutional scheme requires. The key point is that many inequalities—even those that seem paradigmatically natural in the relevant sense—actually arise from the *interaction* between “internal endowments, [] external resources, and the social and physical environment which

¹⁸⁰ Robeyns (2005, p. 39)

¹⁸¹ Anderson (2010).

[they] find themselves.”¹⁸² Thinking about the natural/social distinction in terms of the relevant interactions has a long history, particularly in theoretical debates as to how to conceptualize the nature of disability, and so it is worth briefly outlining the three prevailing “models” in this literature to situate the subsequent discussion.

- (i) *The medical model.* One way to define disability is as a deviation from biomedical norms. This is what is known as the *medical model*, or the *individual pathological model*, since the focus is on the *inability* of the particular individual to achieve some species-wide functioning. A person born blind, on this view, is unable to read because of an individual functional limitation; their eyes are not “working” the way they ought to, and we know that because of the way the relevant community’s eyes do in fact “work.”¹⁸³ The gap between a person’s functional limitation and the functional norms of the community are thus treated as a medical phenomenon requiring medical correction or prevention through, e.g., biological screening.

- (ii) *The social model.* In stark contrast to the medical model of disability, the *social model*, or the *social pathological model*, rejects the view that disability is best thought of in exclusively medical terms as an individual’s functional limitations. Instead, proponents of the social model view disability as a problem of social organization and social institutions (both formal and informal). Social model theorists make a distinction between an individual’s impairment, which may be thought of purely in individual and biological terms, and disability, which may be thought of purely in terms of social organization. Impairment itself is not viewed as a problem but rather “it is the way difference and impairment manifest themselves in our social institutions that results in a problem.”¹⁸⁴ Where the medical model takes an individual’s impairment as a sufficient condition for disability, the social model rejects the idea that impairment is even a necessary condition for disability. Disability is, in other words, purely a matter of social organization and social oppression even if impairment is purely a matter of individual traits and biological endowment.¹⁸⁵ On the social model, a person born

¹⁸² Anderson (2010, p. 108)

¹⁸³ Riddle (2013).

¹⁸⁴ Ibid.

¹⁸⁵ Impairment on the social model might be read in a more neutral way, implying difference without any further evaluative judgment as to the status of that difference (see, e.g., “mere-difference” views of disability as in Barnes (2014, 2016)). Critiques of the very idea of mental healthcare and psychiatry, traditionally associated with Thomas Szasz and, more recently, Bryan Caplan, take the view that what we ordinarily take to be mental health problems (such as addiction and schizophrenia) are but terms that scientists and mental health professionals use to categorize and pathologize—and thus render abnormal in an *evaluative* rather than *descriptive* sense—ordinary human diversity. This, I think, would be considered a more extreme version of the social model applied to mental disability. See Szasz (2001) and Caplan (2006).

blind may have what is thought of as a difference or an impairment—that is, their eyes do not operate in the same way as others’ do—but the inability of that person to, e.g., read is solely a matter of whether the requisite social infrastructure has been considered and put into place (i.e., whether there is an adequate availability of braille books).

- (iii) *The interaction model.* In contrast to both the medical and social model of disability, the *interaction* model views disability as best understood as a complex interaction between an individual’s impairment and the *mediating* role of formal and informal institutions—the environment—which provide the “disabling facts of one’s impairment.”¹⁸⁶ The interaction model, *contra* the medical model, rejects the idea that an individual’s impairment is a sufficient condition for disability, but also, *contra* the social model, rejects the idea that an individual’s impairment is *not* a necessary condition for disability. There is, in other words, a recognition of a causal relationship between one’s impairment and one’s disability in the interaction model that is absent in the social model. On the interaction model, a blind person’s inability to read is a matter of social organization (e.g., the availability of braille books) but it is the underlying “physical fact” of blindness that *interacts* with the “institutional fact” of braille book availability that *causes* the disability.¹⁸⁷ Interaction model theorists recognize, where social model theorists do not, the “inevitable physical foundation of the social phenomena [of disability].”¹⁸⁸

Which model an institutional scheme reflects has implications for the overarching resourcist and capability theorist debate since what is often taken to be the decisive factor as to whether an inequality is “justice-relevant” is often framed in simple causal terms. From this view, social inequalities are caused by factors such as social rules and regulations, and these are the sorts of inequalities a just institutional scheme should justify or rectify. For example, when all the desirable jobs with liveable wages are restricted to a particular ethnicity of the population, inequalities arising from this “artificial” restriction are due to a social cause. Natural inequalities, by contrast, stem from factors such as biological or psychological (or otherwise “natural”) dispositions or traits. These natural inequalities are the sorts of inequality that do not require justification or rectification under a just institutional scheme. When one group is better off than another group because of a difference in an innate (e.g., biopsychological) trait (such

¹⁸⁶ Riddle (2013, p. 378).

¹⁸⁷ *Ibid.*, p. 380.

¹⁸⁸ *Ibid.*

intelligence), inequalities that arise from the relative advantage the presence of that trait confers are said to arise because of a natural cause.

In this framework, the medical model views *all* inequalities stemming from disability as naturally caused whereas the social model views *all* inequalities stemming from disability as socially caused. The resourcist critique of capability theorist's metric is, in a sense, an accusation of an institutional scheme that incorporates the worst aspects of both of these models: first, the capability theorist aims for an institutional scheme that generates compensation claims for purely natural inequalities (an institutional scheme inspired by the social model); and second, the capability theorist aims for an institutional scheme that views *individuals* as having better or worse natural endowments (an institutional scheme arguably inspired by, or at least guilty of the most affinity with, the medical model). I return to this latter point below.

But capability theorists can comfortably reject both of these views. It is the interaction model, and its subsequent view of the nature of the *interaction* between the natural (innate traits) and social (formal and informal institutions), that sits most comfortably with the view of the distinction endorsed either explicitly or implicitly in the capability theory literature.¹⁸⁹ This allows the capability theorist to retain what they take to be a virtue of their account—sensitivity to individual variation—without sliding into a radical egalitarianism that would be a natural consequence of an institutional scheme that endorsed the social model (and so recognized any and all disabilities to be legitimate candidates for justification or rectification in the form of compensation complaints in virtue of their being social caused).

To illustrate the interaction between the natural and the social as characterized by interaction model, consider phenylketonuria (PKU). PKU is a genetic disorder that causes severe intellectual disability. Individuals who have PKU are unable to convert the amino acid phenylalanine into other amino acids. Since high levels of phenylalanine cause severe brain damage, individuals who have PKU, if untreated, will develop severe intellectual disability. However, in most industrialized societies, new-born infants are routinely screened for phenylalanine levels. New-borns with PKU (who screen for higher-than-normal phenylalanine levels) are kept on a special low-phenylalanine diet, and do not develop intellectual disability. This is a clear case where a social factor (appropriate medical screening and subsequent dietary restrictions) directly influences, in a causally relevant way, the phenotypic expression of the natural factor (gene endowment) as it relates to an outcome that may or may not generate a compensation claim (inequalities borne from having an intellectual disability). From the view of the interaction model, the relevant intellectual disability is only realized if the requisite forms of social organization (appropriate medical screening and subsequent dietary restrictions) allow for it, but the underlying physical fact of the gene endowment also plays a causally efficacious role.

¹⁸⁹ See, e.g., Anderson (2010); Nussbaum (2002, 2006, 2009).

Many resourcists take variation in genetic endowment to be a matter pure personal heterogeneity and so the inequalities that arise in light of such variation as justice-irrelevant. It is here where the PKU case, and many other cases that have a similar causal structure, are particularly illuminating.¹⁹⁰ This is because these cases show that it is difficult to separate out, in the relevant way resourcists may want to, the distinction between inequalities caused by a natural lottery in terms of gene endowment and inequalities caused by formal and informal institutions. Are individuals who have severe intellectual disabilities as a result of untreated PKU simply unlucky, for they were born with the wrong genes, or is society “on the hook” so to speak, for inadequate social organization and unjust distribution of medical resources?¹⁹¹ Capability theorists who endorse sufficientarian standards for, say, intellectual outcomes can accommodate the complex interdependencies involved in cases like these and reflected by the more general interaction model in a way that less sophisticated versions of resourcism cannot. They can do this, crucially, without viewing the distribution of natural endowments as a matter of justice. And it is worth noting that this nuanced understanding also leaves open the possibility that there are “pure” natural inequalities, such as those experienced by individuals whose (severe) disability would not be mitigated under varying social institutions.

Sophisticated resourcists seem well aware of these complex interactions. For example, in the context of inequalities that arise from disabilities, Pogge claims that “special needs and disabilities in which social causes play no role at all are rare.”¹⁹² Indeed, this point is part of a broader argument for why resourcism can take into account a wide range of personal heterogeneities against Sen’s initial criticism (outlined above). So with this in mind, it is not clear why these complex interdependencies between natural and social do not flow down to internal natural endowments as well, considering that genes can be considered to be the “canonical natural resources” and that “[i]t is a platitude of behavioural genetics that the causal action of a given gene is dependent on the environment in which it is found.”¹⁹³ The point here is that it seems as though, even by a sophisticated resourcist’s lights, the inequalities that would count as those legitimately able to generate compensation claims are larger than what resourcists themselves take to be the case. Both views are able to accept that “purely” natural inequalities do not generate claims of justice on their own, but that in many cases what look like purely natural inequalities are actually mediated by socially controllable forms of social organization.

Recognition that the interaction model is the underlying view of disability in the capability theorist’s preferred institutional scheme also goes some way in deflating the *vertical inequality problem* identified by Pogge. From the capability theorist’s view, persons are not better or worse endowed than others since what is salient with regard to institutional arrangements is the forms of socially controllable social organization that mediate the causal relationship between

¹⁹⁰ For a good overview, see Lewens (2010).

¹⁹¹ *Ibid.*, p. 272.

¹⁹² Pogge (2010, p. 30).

¹⁹³ Lewens (2010, p. 270).

an individual's impairment and the disability. There does not seem to be a stigma attached to the claim that children with PKU are entitled to the some form of medical attention early on in life lest they develop severe intellectual disabilities later on in life, for example. Rather, the vertical inequality problem seems to be most acutely felt by an institutional scheme that endorses the medical model of disability. Indeed, it is true that the medical model characterized much of the assessment and conceptualization of disability in the twentieth century, and this characterization did indeed lead to much stereotyping and stigma of disabled persons as a result. This is because the medical model places the focus on the *individual* by positing that the innate trait or otherwise natural disposition is the sufficient cause of the relevant inequality in question. But capability theorists need not, and in fact do not, endorse the medical model, and thereby are not guilty of viewing individuals as better or worse endowed.

It is worth mentioning here that some social model theorists argue against both the medical model and the interaction model on the grounds that they lead to the unwarranted stigmatization of those with disabilities.¹⁹⁴ The claim here is that both views are guilty of recognizing a causal link between impairment and disability, even if medical model theorists take impairment to be sufficient and the interaction model theorists take impairment to be necessary. It is the fact that there is a theoretical link at all, so social model theorists allege, that triggers stigmatization and stereotyping. This argument fails to appreciate the distinction between what it means for a cause to be sufficient or necessary, for in this case the interaction model shifts the view from individual pathology to social organization in a way that mitigates the vertical inequality worry. Thus, this particular critique from social model theorists fails for similar reasons to why Pogge's critique fails.

These three models and their underlying characteristics provide instructive analogies that generalize beyond disability. What is particularly important for our purposes is the way the three models treat the causality with respect to innate or natural traits or dispositions and social organization. The medical model views disadvantages as being caused purely in terms of innate or natural traits; the social model views disadvantages as being caused purely in terms of social organization; and the interaction model, the one I think most accords with the capability approach, views disadvantages as being caused by a complex interaction between both innate, natural traits and the social organization in which they are found. In the next section, I outline briefly what this interactional way of thinking means for understanding of "natural traits" beyond disability.

b. Interaction, institutions, and gender

In the last section I claimed it is not clear how to distinguish between the relevant social and natural factors in the *simple* causal terms as required. Even what we ordinarily take to be purely

¹⁹⁴ For a brief overview, see Riddle (2013).

natural traits of individuals or features their natural environment often crucially rely upon social conditions and rules (e.g., formal and informal institutions) for their relevant causal effects. I made the case that this interactional thinking goes “all the way down to genes,” that is, what we might take to be the clearest example of purely natural causes. In this section, I consider the way other forms of social organization may play a role in mediating “natural” facts of our environment beyond just genes.

Consider, for example, food distribution. It’s reasonable to think of the availability of food as a paradigmatic feature of the environment and so of a species’ natural endowment—directly related to the soil fertility, temperature, adequate water supply, and so on, in the relevant geographic area. Yet whether there is enough food for individuals in a given community is often determined by societal distribution networks or social rules determining who gets what, and when. As Sen has argued convincingly, even famine has less to do with Malthusian pessimism about the relative properties of population growth and agricultural productivity than it does with the distribution of existing food supplies and shifts in purchasing power due to speculation, hoarding, or inflation.¹⁹⁵

Social norms and other informal institutions that regulate the distribution of food and other resources may produce conditions that disadvantage some groups over others all while masking the subsequent disadvantage with the appearance of an underlying “natural” cause. For example, “table manners” in patriarchal societies often dictate that men and boys should eat first—over women and girls—leading to predictably disastrous outcomes for women and girls in resource constrained environments. In cases like these, it’s not that there is not enough food to go around or that women and girls are particularly susceptible to malnutrition or the diseases that come along with, but rather there are social rules governing who should eat first, and how much they are allowed to eat. Men and boys eat first and they are also given preference for second-helpings, if there are any to go around. Social rules such as these may confer disadvantage to some groups overtly but there may also be other subtler norms at work, as when women and girls must meet a higher threshold of pain or discomfort before a family decides it appropriate to seek medical attention for a potential underlying medical condition.

Note that this point is entirely compatible with the view that there *are* natural differences between individuals or groups. Consider biopsychological differences between the sexes. Evolutionary psychologists often claim that, over the course of human evolutionary history, different sexual selection pressures in ancestral environments have resulted in systematic differences between the sexes (i.e., sexual dimorphism). A primary difference posited in this regard concerns the respective parental investment (the investment by the parent in the offspring that increases the offspring’s chance of survival at the cost of the parent’s ability to invest in other offspring) between the two sexes.¹⁹⁶ Law professor Kingsley Brown, drawing inspiration from this

¹⁹⁵ Sen (1977, 1986).

¹⁹⁶ Trivers (1972).

line of reasoning, claims that women increase their reproductive success by “devoting the bulk of their parental energies to investment in children, through provision of milk and other forms of direct care taking,”¹⁹⁷ whilst men increase reproductive success “through the attainment of high status and control of resources,” giving them access to multiple and the “most desirable” mates.¹⁹⁸ As a result, Browne claims, men are, *inter alia*, more likely to take physical and nonphysical risks and be physically stronger than women, whereas women exhibit greater verbal skills and are more bound to their infants than men. This sexual dimorphism is then taken as an explanation for disparate societal outcomes (such as variation in occupational patterns between the sexes). For example, Browne claims sexual dimorphism is the reason why there are more male CEOs (since men prefer high-risk and/or high-reward occupations), and why there are more female nurses and stay-at-home-mothers (since they prefer caretaking occupations or prefer to forego the labor market altogether in order to raise children).

But even if systematic differences implied by this particular interpretation of sexual dimorphism *do* exist, there are still a host of social rules—such as rules governing the permissibility of bringing children to work or community-wide expectations about the appropriate level of risk a woman ought to take on as part of her working life—that explain disparate labor market outcomes in a causally efficacious way. Social technology also plays a causal role, of course, especially if particular innovations shift the relative costs and benefits of choices in a way that would interact with the suggested natural traits. For example, ride-hailing mobile applications for taxi cabs (such as Lyft or Uber) lower the transaction costs involved in the cab services market by providing relevant information about the trustworthiness of driver and rider to the relevant parties, thereby reducing the level of risk attached to driving a cab for a living. If it were true that women are naturally more risk averse, technological innovation exemplified by the ride-hailing application software created by Lyft and Uber would still render open previously “too risky” potential occupations. From this view, social rules and/or social technology play an important part—perhaps the important part—in the relevant labor market outcomes rather than the putatively natural underlying traits posited in a way that forecloses a parsing out of the natural and the social in simple causal terms.

4. Measuring global justice

The arguments in this paper so far have been concerned with what sorts of inequalities might legitimately generate compensation complaints under just institutional schema envisioned by two prominent liberal views: Rawlsian resourcism and the capabilities approach. What I have tried to show is that, despite what resourcists claim, capability theorists are not committed to the view that the distribution of natural endowments is just or unjust. Rather, at least some inequalities

¹⁹⁷ Brown (2004, p. 277).

¹⁹⁸ *Ibid.*

that resourcists take to view as purely natural are actually caused by socially controllable forms of social organization, and this becomes clear once the appropriate model of disability—the interaction model—is recognized.

In this final section, I move away from questions about legitimate compensation claims to another important implication for a theory's metric of justice, namely, issues pertaining to how we ought to make comparative judgments about well-being, poverty, and inequality within a global context. This practical concern of poverty and general welfare measurements in the service for comparative evaluative judgements between countries and/or regions of the world is, of course, distinct from the more theoretical questions that animate the debate between which inequalities are legitimate candidates for rectification or justification in a just institutional order.

Nonetheless, to the extent that evaluative metrics are indeed informed by overarching notions of justice, the two questions are linked.¹⁹⁹ The arguments in this paper so far shed some light on the connection. Specifically, what becomes clear in light of the arguments in this paper so far is that the importance of the choice between a resourcist metric (insensitive to individual difference) and a capability theory metric (sensitive to individual difference) turns, in large part, on the distribution of disability across the globe. If disability is randomly distributed, that is, the “rate” of disability is generally constant per capita, then the choice in metric will not make much of a difference in terms of measuring underlying poverty or welfare levels. Randomly distributed disability will be measured indirectly through its effects on the distribution and consumption of social primary goods under the resourcist metric, and it will be measured somewhat more directly in terms of its causal effects on the freedoms or functionings persons are able to enjoy (or not) under the capability theorist metric. Even with a difference in the direct and indirect measurement, the overall comparative evaluation between countries would be the same.

The problem, of course, is that there is no reason to think that disability will in fact be randomly distributed across the globe. Rather, in light of the interaction model of disability, we would expect there to be a non-random distribution of disability given that disability is, at least in many of the cases we are interested in, directly related to how underlying “natural” impairments interact with social organization. Non-random distribution is expected because it is a near truism in development economics and other nearby social sciences that the highly disparate development outcomes observed between countries is largely a function of the variation in the relevant institutions between those countries. Formal and (as I have argued both in Chapter 1 and elsewhere) informal institutions vary tremendously between different countries, and some institutions are better than others producing welfare-relevant desirable outcomes.²⁰⁰ This point is made *a fortiori* by looking toward “natural experiments” where institutional variation is present but demographic and environmental variables are held constant. Think about the difference in expected welfare outcomes for residents of Juarez and El Paso, or Nogales and

¹⁹⁹ Robeyns (2005).

²⁰⁰ Patel and Hoffman (2017, forthcoming).

Tucson, cities that fall on either side of the border between the United States and Mexico, for example. The residents of Juarez and Nogales do worse on every welfare relevant metric compared to the residents of Tucson and El Paso, who are but a few miles away from them separated by an international border and living under an entirely different set of institutions that the border reflects.

The interaction model gives us the theoretical resources to understand how macro- and meso-level formal and informal institutions interact with underlying natural traits and dispositions to produce development relevant outcomes, much in the way that macro-level comparisons of political institutions are often taken to be the explanatorily decisive factor in determining economic outcomes in standard development practice. Note, also, that the capability theorist's metric remains sensitive to non-random distribution of disability in way that helps us make comparative evaluative judgments across countries *even if* the underlying distribution of natural traits or dispositions (impairments) is randomly distributed. Non-random distribution of disability and random distribution of impairment are entirely compatible on this view, for impairment is recognized as a necessary but not sufficient condition for disability. The capability theorist's measurement metric will therefore be able to make micro-comparisons of effects of the effects disability has on key functionings in a way that is only indirectly measured by the resourcist approach, even on the more sophisticated versions of that approach.²⁰¹

5. Conclusion

In §2a, I outlined resourcism and the capabilities approach, paying special attention to the wide agreement between the two approaches in their more sophisticated forms. In §2b, I aimed to expose the precise point of divergence between the views, namely, with respect to the distribution of natural endowments (and the natural inequalities they may give rise to). In §2c, I formulated and explicated the resourcist's central critique of the capability theorist's metric: the compensation for natural inequalities critique. In §3a, I argued that the capability theorist can adopt the interaction model of disability, a strategy that largely deflates the resourcist critique. In §3b, I related the underlying characteristics of the interaction model, namely, its focus on the interaction between putatively natural traits and their environment, to broader understandings of the role that formal and informal institutions play in development outcomes. I paid special attention to the problem of gender in this subsection. In §4 I moved from a focus on what kinds of inequalities might trigger compensation claims under the two respective views to how their respective metrics fare with respect to comparative evaluative judgments of welfare-relevant outcomes across countries.

²⁰¹ Robeyns (2005).

APPENDIX 1. Methodology (measuring informal institutions)

The empirical findings in this dissertation are based on a specialized social norms survey, '*Local Understandings, Experiences and Expectations Survey*' developed by the Chatham House Africa Programme and the University of Pennsylvania's Social Norms Group (PennSONG). Survey administration was carried out in 2018 with Nigeria's National Bureau of Statistics (NBS) as well academics and practitioners from around the country. 5,600 surveys were carried out across urban and rural areas in Adamawa, Benue, Enugu, Lagos, Rivers, Sokoto, and the Federal Capital Territory of Abuja (FCT).

The survey instrument asked about five behaviours: power sector corruption, examination malpractice, vote selling, and embezzlement. The survey was designed to uncover underlying beliefs that drive corrupt practices. Using the theoretical resources from Bicchieri's (2006; 2016) definition of social norms, the survey asked the following questions:

1. Behaviour – what did you do? How many times do you think it happened?
2. Prudential – what are the practical reasons for engaging or not engaging in a practice?
3. Empirical expectations – what do you think other people do in the relevant situation?
4. Normative expectations – what do you think other people think you should do in the relevant situation?
5. Personal normative belief – what do think about a particular practice (in a moral evaluative sense)?
6. Legal knowledge – do you think this practice is illegal?

The survey instrument was designed to measure the beliefs and expectations that accompany voter markets in seven states across the country: Adamawa, Benue, Enugu, Lagos, Rivers, Sokoto, and FCT-Abuja.

(a) What kind of beliefs did we measure?

(i) Personal normative beliefs

What are personal normative beliefs?

Personal normative beliefs are beliefs individuals hold about what people *should* do. There are two different types of personal normative beliefs: (1) *prudential* personal normative beliefs, which are beliefs about what one should do that appeal to a purely practical or instrumental reasons; and (2) *non-prudential* personal normative beliefs, which are beliefs about what one should do that appeal to non-practical or non-instrumental reasons, such as moral or religious reasons.

Measuring personal normative beliefs can be complicated because many languages have terms that denote normativity but are ambiguous as to the type of normativity denoted. In English, “should” operates in this way. Consider the following examples:

(1) I believe I should sell their votes because

- a. Prudential: I don't trust the politician to deliver on campaign promises so I will take what I can get.
- b. Non-prudential: The politician is from my district, and if he offers something to secure my support, it is my obligation to give it.

(2) I believe that people should not sell their votes because

- a. Prudential: Taking money from one political party will invite violence from another political party.
- b. Non-prudential: Each vote sold is a private benefit at the expense of a social cost. We have obligations to each other to avoid acting selfishly in order sustain fair political institutions.

Distinguishing between prudential and non-prudential beliefs is crucial because intervention strategies for beliefs that are sustained by prudential beliefs will differ from those sustained by non-prudential beliefs. For example, prudential beliefs may be responsive to empirical evidence; moral or religious beliefs, especially deeply held ones, are not. This is why information campaigns aimed at exposing the false prudential beliefs might be effective if the target behaviour is driven by false prudential beliefs.

How did we measure personal normative beliefs?

To measure personal normative beliefs with respect to vote markets, respondents were asked “Do you think that people should collect money for a gift or a vote?” and “Do you think that it is acceptable that people should collect money or a gift for a vote?” Depending on their answers to these questions, they were then asked why they thought people should/should not collect money for a gift or a vote and why it was acceptable/not acceptable to collect money for a gift of a vote. The responses to these questions were then coded by the survey administrator as prudential, non-prudential, or other.

(ii) Empirical expectations

What are empirical expectations?

Empirical expectations are beliefs about what others – typically a relevant reference network – do with respect a particular behaviour. “I believe that 8 out of 10 people in my community exchanged their vote for a money or a gift in the last election” is an example of an empirical expectation.

Empirical expectations might sometimes be taken as an imperfect proxy for the frequency of a particular behaviour, especially in cases where direct questioning will likely lead to social desirability bias in response. “How many people in your community do you think sold their vote?” is more likely to elicit a truthful response relative to “Did you sell your vote?”

How did we measure empirical expectations?

To measure empirical expectations with respect to vote markets, respondents were asked “Think about the people in your community, such as your friends, family, neighbours, and colleagues. Out of 10 people in your community who voted in the last election, how many of them do you think collected money or a gift for their vote?”

(iii) Normative expectations

What are normative expectations?

Normative expectations are beliefs about the personal normative beliefs of other people. Since personal normative beliefs might be either prudential or non-prudential, an individual’s normative expectations may be composed of prudential or non-prudential beliefs. Consider the following examples:

- (1) I believe other people believe that individuals should sell their votes because
 - a. Prudential: They don’t trust the politician to deliver on campaign promises so they should take what they can get.
 - b. Non-prudential: The politician is from our district, and if he offers something to secure our support, it is our obligation to give it.

Violations of normative expectations that are composed of non-prudential normative beliefs will typically illicit sanctions. This is because expectations driven by non-prudential normative beliefs suggest a *shared rule* within a community (typically a moral or religious rule). However, widely shared prudential beliefs do not suggest a shared rule within a community. Rather, a community may come to hold the same set of prudential beliefs because they have access to the same relevant information and interpret it in the same way.

Note that individuals need not hold the normative beliefs that compose their normative expectations. Normative beliefs may thus be *inconsistent*. A community is said to be suffering from *pluralistic ignorance* when there is widespread mutual inconsistency in normative expectations, e.g., when most individuals in the community share the same personal normative beliefs but believe that others do not.

How did we measure normative expectations?

To measure personal normative beliefs, respondents were asked “Think about the people in your community, such as your friends, family, neighbors, and colleagues. Out of 10 people in your community who voted in the last election, how many of them do you think said that it is acceptable to collect money or a gift for their vote?”

What kind of collective behaviors do these beliefs contribute to?

Measuring personal normative beliefs, empirical expectations, and normative expectations is key to understanding the following collective behaviors:

Moral rules

If a collective practice is sustained by a moral rule, individuals usually have a *socially unconditional* preference to follow the rule because of *non-prudential personal normative beliefs*. For example, a devout Muslim would not eat pork even if they lived in a society where the consumption of pork was widespread.

Shared prudential norms

If a collective practice is a shared prudential norm, individuals usually have a *socially unconditional* preference to follow the rule because of *prudential personal normative beliefs*. For example, the frequency of smoking cigarettes may be low in a number of communities because each individual holds prudential beliefs about the health-related risks posed by cigarette smoke.

Descriptive norms

If a collective practice is a descriptive norm, individuals usually have a *socially conditional preference* to follow the rule based on their empirical expectations. Descriptive norms are typically solutions to coordination problems. For example, individuals will generally drive on a certain side of the road because others drive on that side of the road.

Social norms

If a collective practice is a social norm, individuals usually have a *socially conditional preference* to follow the rule based on both their empirical and normative expectations. For example, individuals in the United States will usually tip after a meal in a restaurant because they expect other individuals to do so and because they expect others to expect them to do so.

Bibliography

- Acemoglu, D., & Robinson, J. A. (2013). *Why nations fail: The origins of power, prosperity, and poverty*. Broadway Business.
- Acemoglu, D., & Robinson, J. A. (2019). *The narrow corridor: States, societies, and the fate of liberty*. Penguin Press.
- Anderson, E. (2010). Justifying the capabilities approach to justice. *Measuring justice: Primary goods and capabilities*, 81-100.
- Arneson, R. J. (2011). Side constraints, Lockean individual rights, and the moral basis of libertarianism. *The Cambridge companion to Nozick's Anarchy, state, and utopia*.
- Arneson, R. (2017). Democracy Versus Libertarianism. In *The Routledge Handbook of Libertarianism* (pp. 73-87). Routledge.
- Batenburg, R. S., Raub, W., & Snijders, C. (2003). Contacts and contracts: dyadic embeddedness and the contractual behavior of firms. *Research in the Sociology of Organizations*, 20(1), 135-188.
- Bates, R. H. (1988). Contra contractarianism: some reflections on the new institutionalism. *Politics & Society*, 16(2-3), 387-401.
- Beccerra, M., & Gupta, A. K. (1999). Trust within the organization: Integrating the trust literature with agency theory and transaction costs economics. *Public Administration Quarterly*, 177-203.
- Bicchieri, C. (2005). *The grammar of society: The nature and dynamics of social norms*. Cambridge University Press.
- Bicchieri, C. (2016). *Norms in the wild: How to diagnose, measure, and change social norms*. Oxford University Press.
- Block, W. E. (2011). Libertarianism vs. Objectivism: A Response to Peter Schwartz.
- Block, W. E. (2011). Public goods and externalities: The case of roads.
- Brighouse, H., & Robeyns, I. (Eds.). (2010). *Measuring justice: Primary goods and capabilities*. Cambridge University Press.
- Browne, K. R. (2004). Women in the workplace: Evolutionary perspectives and public policy. *Evolutionary Psychology, Public Policy and Personal Decisions*, 257.
- Buchanan, J. M. (2006). Why I, too, am not a conservative. *Books*.
- Caplan, B. (2006). The economics of szasz: Preferences, constraints and mental illness. *Rationality and society*, 18(3), 333-366.
- Caspi, A., McClay, J., Moffitt, T. E., Mill, J., Martin, J., Craig, I. W., ... & Poulton, R. (2002). Role of genotype in the cycle of violence in maltreated children. *Science*, 297 (5582), 851-854.
- Caspi, Avshalom, et al. "Moderation of breastfeeding effects on the IQ by genetic variation in fatty acid metabolism." *Proceedings of the National Academy of Sciences* 104.47 (2007): 18860-18865.

- Dorfman, A., & Harel, A. (2013). The case against privatization. *Philosophy & public affairs*, 41(1), 67-102.
- Ellerman, D. (2010). Inalienable rights: A litmus test for liberal theories of justice. *Law and Philosophy*, 29(5), 571-599.
- Fisman, R., & Golden, M. A. (2017). *Corruption: What everyone needs to know*. Oxford University Press.
- Freeman, S. (2001). Illiberal libertarians: Why libertarianism is not a liberal view. *Philosophy & Public Affairs*, 30(2), 105-151.
- Freeman, S. (2006). Frontiers of justice: the capabilities approach vs. contractarianism. *Texas Law Review*, 85(2), 385.
- Freeman, S. (2007). *Rawls*. Routledge.
- Freeman, S. (2018). Cognitive Disabilities. *Disability in Practice: Attitudes, Policies, and Relationships*, 174.
- Freiman, C. (2013). Cosmopolitanism within borders: on behalf of charter cities. *Journal of Applied Philosophy*, 30(1), 40-52.
- Fukuyama, F. (2014). *Political order and political decay: From the industrial revolution to the globalization of democracy*. Macmillan.
- Fuller, B., & Romer, P. (2012). *Success and the city: How charter cities could transform the developing world*. Macdonald-Laurier Institute for Public Policy.
- Grief, A. (2003). Comparative and historical institutional analysis: A game theoretic perspective.
- Grief, A. (2006). *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. Cambridge University Press. Cambridge UK.
- Greif, A., & Kingston, C. (2011). Institutions: rules or equilibria?. In *Political economy of institutions, democracy and voting* (pp. 13-43). Springer, Berlin, Heidelberg.
- Guala, F. (2016). *Understanding institutions: The science and philosophy of living together*. Princeton University Press.
- Hampton, J. (1988). *Hobbes and the social contract tradition*. Cambridge University Press.
- Hindriks, F., & Guala, F. (2015). Institutions, rules, and equilibria: a unified theory. *Journal of Institutional Economics*, 11(3), 459-480.
- Hobbes, T. (1998). On the citizen (de cive). *ders.: Man and Citizen (De Homine and De Cive)*, 3, 87-386.
- Hobbes, T. (2016). *Thomas Hobbes: Leviathan (Longman Library of Primary Sources in Philosophy)*. Routledge.
- Hodgson, G. M. (2005). Introduction to the inaugural issue by the Editor-in-Chief. *Journal of Institutional Economics*, 1(1), i-ii.
- Hoppe, H. H. (1989). Fallacies of the public goods theory and the production of security. *Journal of Libertarian Studies*, 9(1), 27-46.

- Hoffmann, L. K., & Patel, R. N. (2017). *Collective action on corruption in Nigeria: A social norms approach to connecting society and institutions*. London: Chatham House.
- Hoffmann, L. K., & Patel, R. N. (forthcoming). *Local Understandings, Experiences and Expectations of Corruption in Nigeria*. London: Chatham House.
- Hutcheson, F., & Leechman, W. (1755). *A system of moral philosophy* (Vol. 1). Millar.
- Kant, I., & Gregor, M. J. (1996). *Practical Philosophy* (The Cambridge Edition of the Works of Immanuel Kant). Trans. M.J Gregor. *The Cambridge Edition of the Works of Immanuel Kant*. Cambridge: Cambridge University Press.
- Keefer, P. (2018). Collective action and government: still a mystery. In *A Research Agenda for New Institutional Economics*. Edward Elgar Publishing.
- Lewens, Tim. "What are 'Natural Inequalities'?" *The Philosophical Quarterly* 60.239 (2010): 264-285.
- Mack, E., & Gaus, G. F. (2004). Classical liberalism and libertarianism: The liberty tradition. *Handbook of political theory*, 115-130.
- Mallaby, S. (2010). The Politically Incorrect Guide to Ending Poverty. *The Atlantic*, 306(July/August), 93-102.
- North, D. C. (1981). *Structure and change in economic history*. Norton.
- North, D. (1990). Institutions and their consequences for economic performance. *The limits of rationality*, 383-401.
- North, D., & Weingast, B. R. (1990). The role of institutions in the revival of trade: The law merchant, private judges, and the champagne fairs. *Econom. Politics*, 2(1), 1-23.
- North, D. C. (1993). The new institutional economics and development. *Economic History*, 9309002.
- North, D. (1997). Institutions, institutional changes and the functioning of the economy. *M.: Nachala*.
- North, D. C., Wallis, J. J., & Weingast, B. R. (2009). *Violence and social orders: A conceptual framework for interpreting recorded human history*. Cambridge University Press.
- Nozick, R. (1974). Anarchy, Utopia and the state. *New York: Basic Book*.
- Nussbaum, Martha Craven. *The Quality of Life*. Oxford [England: Clarendon ;, 1993. Print.
- Nussbaum, Martha Craven. *Women and Human Development: The Capabilities Approach*. Cambridge: Cambridge UP, 2000. Print.
- Nussbaum, Martha Craven. *Frontiers of Justice: Disability, Nationality, Species Membership*. Cambridge, Mass.: Belknap ;, 2006. Print.
- O'Neill, Martin. "What should egalitarians believe?." *Philosophy & Public Affairs* 36.2 (2008): 119-156.
- Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. Cambridge university press.

- Ostrom, E. (1990). The evolution of institutions for collective action. *Edición en español: Fondo de Cultura Económica, México*.
- Ostrom, E. (2005). Doing institutional analysis digging deeper than markets and hierarchies. *Handbook of new institutional economics*, 819-848.
- Ostrom, E. (2010). Institutional analysis and development: Elements of the framework in historical perspective. *Historical developments and theoretical approaches in sociology*, 2, 261-288.
- Ostrom, E. (2014). Do institutions for collective action evolve?. *Journal of Bioeconomics*, 16(1), 3-30.
- Pogge, Thomas. *Realizing Rawls*. Ithaca: Cornell UP, 1989. Print.
- Pogge, Thomas W. "Can the Capability Approach Be Justified?" *Philosophical Topics* 30.2 (2002): 167-228.
- Pogge, T. (2010). A critique of the capability approach. *Measuring justice: Primary goods and capabilities*, 17.
- Rawls, J. (1999). *The law of peoples: with, the idea of public reason revisited*. Harvard University Press.
- Rawls, J. (2001). *Justice as fairness: A restatement*. Harvard University Press.
- Rawls, J. (2005). *Political liberalism*. Columbia University Press.
- Rawls, J. (2009). *A theory of justice*. Harvard university press.
- Rajagopalan, S., & Tabarrok, A. (2014). 10. Lessons from Gurgaon, India's private city. *Cities and Private Planning: Property Rights, Entrepreneurship and Transaction Costs*, 199.
- Riddle, C. A. (2013). Defining disability: Metaphysical not political. *Medicine, Health Care and Philosophy*, 16(3), 377-384.
- Riddle, C. A. (2014). *Disability and justice: The capabilities approach in practice*. Lexington Books.
- Riddle, Christopher. "Natural Diversity and Justice for People with Disabilities." *Disability and the Good Human Life*. Cambridge, UK: Cambridge UP, 2014. Print.
- Robeyns, Ingrid. "Assessing global poverty and inequality: income, resources, and capabilities." *Metaphilosophy* 36.1-2 (2005): 30-49.
- Robeyns, I. (2017). *Wellbeing, freedom and social justice: The capability approach re-examined*. Open Book Publishers.
- Romer, P. (2010). *Technologies, rules, and progress: The case for charter cities* (No. id: 2471).
- Sagar, R. (2016). Are charter cities legitimate?. *Journal of Political Philosophy*, 24(4), 509-529.
- Sen, Amartya. *Inequality Reexamined*. New York: Russell Sage Foundation ;, 1992. Print.

- Sen, Amartya, and James E. Foster. On Economic Inequality. Enl. ed. Oxford: Clarendon ;, 1997. Print.
- Sen, Amartya. The Idea of Justice. Cambridge, Mass.: Belknap of Harvard UP, 2010. Print./
- Szasz, T. (2001). Mental illness: psychiatry's phlogiston. *Journal of Medical Ethics*, 27(5), 297-301.
- Tan, K. C. (2004). *Justice without borders: Cosmopolitanism, nationalism, and patriotism*. Cambridge University Press.
- Tan, K. C. (2012). *Justice, institutions, and luck: the site, ground, and scope of equality*. Oxford University Press.
- Trivers, R. L. (1972). Parental Investment and Sexual Selection.
- Weingast, B. R. (2016). Exposing the neoclassical fallacy: McCloskey on ideas and the great enrichment. *Scandinavian Economic History Review*, 64(3), 189-201.
- Williamson, O. E. (1998). The institutions of governance. *The American Economic Review*, 88(2), 75-79.
- Ypi, L. (2013). What's wrong with colonialism. *Philosophy & Public Affairs*, 41(2), 158-191.