Title
A Political, Economic and Social Agent Based Model of Migration (MAPES)

Author
Lim, Daniel Kibum

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

A Political, Economic and Social

Agent Based Model of Migration

(MAPES)

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Political Science

by

Daniel Kibum Lim

2019

ABSTRACT OF THE DISSERTATION

A Political, Economic and Social
Agent Based Model of Migration
(MAPES)

by

Daniel Kibum Lim
Doctor of Philosophy in Political Science
University of California, Los Angeles, 2019
Professor Michael F. Thies, Chair

In this study, I present and validate the first version of MAPES, an agent-based model (ABM) that facilitates the study of international migration (IM) by providing a framework within which theories from different fields can be easily modeled and reconciled. To validate this model, I examine three scenarios that test the capabilities of the model and provide opportunities to make substantive contributions to the field of international migration studies. I use the first of these scenarios, the bilateral economic gradient scenario, to show how rational decisionmaking by individuals gives rise to anti-immigrant sentiment and the flow patterns predicted by macro-economic migration theory. The second scenario, which I call the trilateral equal scenario, exhibits the theoretical phenomenon of path dependence, in which small random changes lead to large differences in migration system outcomes. Finally, I use a scenario that models the real-world border between Mexico and the US to predict the impacts of raising the cost of border crossings from south to north. Through these efforts, I show that MAPES is a research tool which can used to make valuable contributions to the field of IM studies.

The dissertation of Daniel Kibum Lim is approved.

Roger Waldinger

Lorrie Frasure Yokley

Michael F. Thies, Committee Chair

University of California, Los Angeles

2019

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

It's often said that pursuing a PhD is a lonely journey, yet the journey is one that cannot be completed without the support and indulgence of numerous people. Mine has been no exception, despite the self-contained nature of my dissertation and unorthodox path I took through graduate school. I owe thanks to three groups of people who have made it all possible and I hope that this present dissertation and my future endeavors make them proud.

First, I would not have made it to the point of dissertating had it not been for my mentors and classmates in the UCLA political science department. Foremost is my committee chair, Michael Thies, who drilled me through the bootcamp that is PS240, and guided me through two field papers and a dissertation - a better mentor I could not have asked for. My next thanks go to Jeffrey Lewis and James DeNardo who molded a mechanical engineer with identity crisis into a methodologically focused political scientist. Of course, my deepest gratitude also goes to my two other committee members, Roger Waldinger and Lorrie Frasure, my guides in international migration and REP. Last but not least, I cannot overstate the importance of the camraderie I shared with the late night commiseration crew – Jesse Acevedo, Steve Palley, Paasha Mahdavi, Galen Jackson and Mike LaCour – as well as the entire 2009 cohort.

The bulk of this dissertation was completed while I served in the U.S. Army and I am grateful to the many persons, especially in my chain of command, who indulged my academic predilections and granted me the time to see this through to the end. In particular, I am grateful to COL Ian Rice, LTC Benjamin Sangster, MAJ Matt Lacy, and Nicole Camarillo who all encouraged me to stay the course.

Finally, to my friends and family: needless to say, I could not have done this without your ongoing love and support. I love you all and I hope that you are proud of me and this dissertation.

VITA

2006	B.S. (Mechanical Engineering) and (International Studies), Yale University, New Haven, Connecticut.
2012	M.S. (Statistics), UCLA, Los Angeles, California.
2013	M.A. (Political Science), UCLA, Los Angeles, California.
2015	Commissioned 2^{nd} Lieutenant, United States Army.

PUBLICATIONS

A Simulation Approach to Predicting College Admissions. Master's thesis for the degree of Master of Science in Statistics. University of California, Los Angeles. 2013.

Immigration Policy, Native Institutions and Incorporation Among Marriage Migrants in Korea. Unpublished paper. Presented at the Midwest Political Science Association Conference 2012, Chicago, IL. Demonstrated that marriage migrants to Korea become more politically incorporated than guest laborers as a byproduct of policies formulated for economic and cultural aims. Based on a dozen interviews and over 300 surveys from migrants of both types. The latter analyzed using the causal mediation framework of Imai, Tingley, Keele 2010.

The Effects of Pre-Migration Factors on Host Country Voting. Unpublished paper. Presented at the Midwest Political Science Association Conference 2011, Chicago, IL. Explored potential ways in which immigrant-citizens' pre-migration experiences affect their long-term political behavior. Based on hierarchical modeling of a composite dataset constructed from individual-level voter surveys and country-level datasets such as Polity IV, Correlates of War and Penn World Tables.

Choi B, Lim D, and Mitchell J. *Korean-American Attitudes toward and Impact on U.S.-Korea Policy.* Washington: CSIS, 2008.

CHAPTER 1

Introduction

The political economy of international migration (IM) has been studied extensively in political science and closely related disciplines. Consequently, there is a rich literature covering the many causes and outcomes of the phenomena. A downside to this bounty is that it is difficult to account for the myriad theories using just traditional quantitative methods, such as structural equation modeling (SEM) or regression analysis. The limitation of these traditional methods is that they require data (whether observational or experimental in origin) in sufficient quantity containing all the cogs that turn in the theory of interest. Simply put, there are too many theories and not enough data to subtantiate them all. A finding based on one dataset may be difficult to replicate simply because no other dataset captures all the requisite variables. Or a promising new theory may not be tested because gathering the required data is too costly.

Agent-based modeling (ABM) is a simulation based methodology that sidesteps these challenges by turning the logic of analysis on its head. Where traditional methods examine existing data and ask whether certain relationships exist, ABM assumes the existence of these relationships, generates data on that basis and sees whether the results are reasonable. To employ ABM, the practitioner begins by defining a model on the basis of theory. At a minimum, this involves deciding environmental parameters, starting values for agent characteristics, and behavioral rules defining how agents respond to environmental stimuli and interactions with other agents. Running the simulation allows the agents to behave according to their rulesets, and at the end of the day, the practitioner has complete information on all aspects of the model run, which he can analyze to assess how well the theories underlying the model performed.

What all this means is that ABM is an ideal fit for the study of IM for at least three major reasons. First, it avoids the data problem by generating its own data (using priors). Second, it combines micro-level actors and macro-level phenomena, which fits the theoretic scope of migration. Third, because theory is used to generate data (rather than data limiting theory), the existence of a rich literature ceases to be an analytic limitation. Indeed, the only limitation on the number of theories that may be incorporated into a model is whether the practitioner has the programming and analytic chops to manage them.

Given these advantages, I provide in this study a foundational ABM for the study of the causes and consequences of international migration, which I call the (M)igration (A)BM encompassing (P)olitical, (E)conomic, and (S)ocial Factors (MAPES). MAPES focuses on the interactions of individuals and legislators acting as rational agents in response to economic, social, and political stimuli. I use it to show how their interactions impact migration across a series of scenarios designed to validate the model and highlight the usefulness of ABM in the study of IM.

I validate the MAPES approach by showing that its predictions meet the expectations of existing IM theories, given reasonable parameters. And, while the model presented herein is introductory (i.e. it only uses a set of basic IM theories), I show that several non-trivial phenomena still emerge from the simulations, demonstrating the value of this approach. Specifically, I run the model on three scenarios representing different migration systems. The first of these is the bilateral economic gradient (BEG) scenario, which I use to show how rational decisionmaking by individuals gives rise to anti-immigrant sentiment and the flow patterns predicted by macro-economic migration theory. Next, I use the trilateral equal (TE) scenario to show how MAPES exhibits the theoretical phenomenon of path dependence, in which small random changes lead to large differences in migration system outcomes. Finally, I model the real-world (U)S-(M)exico (B)order in (C)alifornia (UMBC) to predict the impacts of raising the cost of border crossings from south to north. Through these efforts, I hope to both make actual substantive contributions to the corpus of IM studies, as well as how MAPES to be a valuable tool for studying international migration.

CHAPTER 2

Background

This study draws on two major bodies of literature. The first is the substantive topic of international migration, and the second is the methodology of agent based modeling. Both span a broad array of disciplines. I first review the substantive theories upon which MAPES is based, a body of literature which should be familiar to political scientists who study IM. The section on ABMs will be more wide ranging and explore its applications in different fields, which highlight their utility for our current purpose.

2.1 Theories of International Migration

International migration, the movement of people across one or more state borders, is difficult to explain with "one size fits all" theories because it can be precipitated and facilitated by many different factors. IM scholars must consequently distinguish between several distinct types of movement, differentiated along different causal axes. Further, because the causes are not often confined to one disciplinary domain, its scholars also hail from different disciplines, bringing with them the unique, not always compatible perspectives of their parent fields.

Much of the extant scholarship on IM is case-specific, focusing on unique or small N situations. Such studies are amenable to narrowly focused, well-tailored theories, but their findings may not be applicable outside those niche cases. A good example of this is Williams and Pradhan (2009), who examine the impact of gender in migration choices made in response to varying levels of political violence in Nepal. Their study is commendable for using longitudinal data on a case outside mainstream migration studies, the Maoist insurgency in Nepal, but obviously their theory holds little relevance outside a certain (not even all) class of refugee migration. Likewise, the canonical macroeconomic theory of international migration, Harris and Todaro (1970), is "probably the oldest and best-known theory of international migration" but is no less privy to the problem of domain specificity.[1] Certainly, it is broader in scope than Nepal but still only applies to labor migration, and is meaningful only insofar as there exists a difference in levels of capital between origin and destination.

This is not to say that all extant literature is narrowly focused; many scholars explicitly acknowledge the multifaceted nature of IM and take a more holistic view, examining the relationships between the different types of migration rather than fixating on one. Zolberg (1999) lambasts sociologists for ignoring the role of politics and the State in their migration theories. Mabogunje (1970) uses the so-called "Systems" approach to root migration in the interplay of economic, social, technological and environmental causes; Wallerstein (1974)'s world systems theory ties macroeconomic factors with history, culture, and geopolitics; and D. S. Massey et al. (1993) suggests a synthesis of different theories to better understand economic migration as a whole. Yet others, such as Richmond (1988), suggest movement away from typologies towards a unified mindset, since most individuals move for combinations of rather than singular reasons. Émigrés can (and quite often do) pursue economic opportunity while simultaneously fleeing political/ideological/religious persecution, so that "an absolutely clear distinction between the economic and the sociopolitical determinants of population movement is not appropriate," even though "there may be exceptional cases where... causes can be identified as 'purely' economic or political."[2]

Finally, some scholars even counsel outright against seeking unifying theories. Portes (1997) puts it quite bluntly (and in my belief, correctly): "There is no overall encompassing theory of immigration... The reason is that the different areas that compose this field are so disparate that they can only be unified at a ... vacuous level."[3] Castles (2010) echoes this view: "General theories are rarely cross-disciplinary; rather they tend to postulate logical structures that cover all conceivable forms of behaviour from the theoretical perspective of

[1] D. S. Massey et al. (1993, p.433)

[2] Richmond (1988, p.14)

[3] Portes (1997, p.810)

a single discipline... The complexity of migratory behaviour, that goes across all areas of human existence, cannot be readily accommodated within such elegant structures."[4] Instead, he advocates "theories of the middle-range," a term he borrows from Merton (1957) to describe properly executed small-N / case studies.

Each approach has its strengths: what the narrow-focus lacks in scope it makes up in parsimony; the systems approach offers explanatory power in exchange for elegance. What seems clear is that migration is such an immense and varied phenomenon that there is traction to be gained by parsing it up into more manageable categories, artificial though they may be. And it is important to remember that just because migration is multifaceted does not mean a theory is less valuable simply because it only focuses on one of those facets. Adopting this position, I now survey some of the more well-known theories of migration, along multiple typological lines.

2.1.1 Economic, Social, and Political Factors

One common categorization scheme is to distinguish between economic, social, and political factors affecting international migration.[5,6,7,8,9] Money and jobs are the most obvious reasons people leave hearth and home to move to strange lands. Pritchett (2006) posits that four forces in the world drive such non-political (i.e. economic) international migration: "wage gaps, demographics, 'everything but labor' globalization, and the service future of labor demand in industrial countries."[10] Many of the canonical theories of migration onset

[4] Castles (2010, p.1574)

[5] Cohen (1996)

[6] King (2002)

[7] King (2012a)

[8] Or as Makowsky, Tavares, Makany, and Meier (2006) puts it, "violence, economics and cultural networks."

[9] Cohen (1996), King (2002, pp.90–91) and King (2002, pp.136-138) use the terms migration binaries, dichotomies or dyads, because they combine social and political causes. I present them as separate categories because there are clearly enough differences to distinguish between the two.

[10] Pritchett (2006, p.13)

are based on one or combinations of these factors, each of which "seeks to explain the same thing... [employing] radically different concepts, assumptions, and frames of reference."[11]

2.1.1.1 Economic Factors

Economic factors fall into two broad subcategories. The first are those which operate at the macro-level, among systems, countries, and markets. The other subcategory pertains to the base units that comprise the aggregate categories: individuals and households. Here I briefly review some of the more seminal theories and schools of though pertaining to each subcategory.

Most well-known amongst the macro theories, the neoclassical macroeconomic approach explains migration when there exists a wage differential between countries with large (low wage) and small (high wage) endowments of labor. Workers wishing to maximize wages will migrate from low to high wage countries, and the change in labor supply in the countries will respectively raise and depress wages. The flow will stop when an equilibrium wage is reached between the source and destination countries (Harris & Todaro, 1970).

Dual labor market theory layers on a sociological aspect to the macroeconomic approach to argue that "the critical factors in understanding the migration process and its evolution over time lie in the developed region," which are specifically pull factors comprised of labor demands.[12] The core of the theory is that workers from low wage countries are funneled onto the bottom rung of a two tiered labor market in the receiving country. The labor market is bifurcated because of social/institutional pressures, but the secondary labor sector is, by design, not privy to those pressures. Thus, employers can maintain a rotating door in which they hire and fire migrants to suit their needs, with no constraints. This then maintains a variable but ever-present demand for migrant workers.

Finally, world systems theory considers the role of historical circumstance, how the histories of colonialism and capitalism have led to the establishment of a center and a periphery,

[11] D. S. Massey et al. (1993, p.432)

[12] Piore (1979, p.17)

each with its designated role in the global economy. While not strictly an economic theory, this categorization is appropriate for our present purpose given its inherent focus on aggregate economic patterns at the expense of individual agency (Wallerstein, 1974).

By contrast, microeconomic theories shrink the unit of analysis to the smallest units of economic decision making, which is the individual. As DeVanzo (1980, p.3) puts it, "an individual migrates in the expectation of being better off by doing so." Most of the microeconomic corpus centers on what respect the migrant does better off in. Most typically, laypersons tend to think of the unit-level migration decision as income maximizing. That is to say, an individual chooses to migrate if the utility gained from income in the prospective destination exceeds the sum of income utility in the present location and the cost of moving. If one wishes to get fancy, a discount term may be added to account for the time-value of money. The canonical equation for this is DeVanzo (1980, p.5):

$$PV_{ij} = \sum_{t=1}^{T} \frac{U_j^t - U_i^t - C_{ij}^t}{(1+r)^t} \tag{2.1}$$

PV is the present value of moving from location i to j, the U terms are the utilities derived from wages at those locations, and the C term is the cost of moving, which includes the costs incurred including transport, acculturation, seeking work. The summation and r and t terms reflect the fact that benefits accrue over time but future benefits should be discounted. A rational person should move if $PV > 0$ for a particular destination.

It is easy to see how equation 2.1 can be generalized to anything measurable as utility. And the usefulness of such an approach is borne out in many studies. For example, Gibson and McKenzie (2009) find in their examination of migrants from several Pacific Islands that the migration decision is most strongly influenced by "risk aversion, patience, and choice of subjects in secondary school... [and] family and lifestyle reasons..." rather than being "strongly linked to... the gain in income... from migrating (p.3)." More whimsically, Graves (1979) bases his study on the "obvious and casual empiricim [of]... retirees migrating to... Florida or Arizona" and finds that "migrants are puchasing a bundle of location-specific goods... [including] climatic amenities (p.145)" In short, warm weather has utility for some people.

Other refinements on the microeconomic approach consider information costs, uncertainty and risk, and migration as a form of insurance and asset diversification.[13,14] They may also shift the unit of decision making from the individual to the household, especially as it relates to risk and insurance. The bottom-line is that the micro-economic approach to theorizing migration is not just about economics in the lay sense; it can provide a useful foundation on which to theorize in general. This is readily seen in the subsequent bodies of theory which are explored: social and political factors.

2.1.1.2 Social Factors

Social factors impact migration by altering preferences and incentive structures for groups and individuals in both the sending and receiving locations. Network effects (through social networks) comprise one important mechanism that facilitates migration. This is the idea that migration begets more migration by altering the costs and benefits of migration through interpersonal relationships. These social networks begin to take hold when the initial one or few pioneer migrants establish footholds in so-called "daughter communities" (D. S. Massey & Espinosa, 1997). The pioneers act as conduits for further migration by bringing in members of their home country social networks, "a process Jones (1982b) has labeled 'channelization'."[15] These movements are motivated by a matching of labor demands in the daughter community and a labor surplus in the sending; all involved parties benefit. The pioneer migrant can expect advancement as he is placed in a supervisory role over those from his network. For employers, "social ties can be used to enforce obligations... and... cement implicit contracts regarding the rights and responsibilities of each party."[16] The migrants who are channeled may have an easier time with migration-related tasks, such as transport

[13] Stark and Lucas (1988)

[14] DeVanzo (1980)

[15] D. S. Massey (1987, p.163)

[16] Waldinger and Lichter (2003, p.83)

and the subsequent job search (Curran & Rivero-Fuentes, 2003).

To put it more generally, social networks directly affect the calculus of migration for individuals (Hatton, Williamson, et al., 2005). By being a member of a migrant social network, individuals can reduce the significant costs of moving arising from transportation, the job search and settlement. The magnitude of this effect is such that it significantly undercuts the effect of increasing border controls (Cornelius, 2001). Even as "the demand for smuggling services and the risks of crossing the border have gone up, [as well as] the price of being smuggled," help from family members who have previously migrated has, at least until now, made this not a serious barrier to migration.[17] Perversely, "the main impact of higher fees... has been to enrich... smuggling groups" rather than significantly decrease undesired immigration.[18] As migration continues, there results an accumulation of knowledge and experience about the migration process within migrant networks. This collective storehouse facilitates future migration among members through "the accumulation of social capital, by which members of a community gain migration-related knowledge and resources through family members and friends who have already traveled."[19] As a potent example of this, Singer and Massey (1998) find that as individuals become well-versed in the process of migration, they learn to play their role in "the social encounter between Border Patrol officers and undocumented... migrants [which] has become scripted and highly ritualized."[20]

Social networks also ensure that the effects of migration persist, by helping migrants survive in their new homes. For example, D. S. Massey (1987) finds that Mexican migrants to the US rely on connections from the homeland – kinship, friendship and *paisanaje* (i.e. coethnicity) – to deal with the exigencies of life in in a foreign place. In utilizing these connections, "they often produce new forms of association that not only promote the cohesion

[17] Andreas (2000, p.95)

[18] Andreas (2000, p.96)

[19] Fussell and Massey (2004, p.152)

[20] Singer and Massey (1998, p.565)

of migrants in the United States but also facilitate their reintegration into the community."[21]

In the sending country, a history of out-migration can alter social structures and the labor market, leading to more migration. It may give rise to a "culture of migration," in which "international migration becomes so deeply rooted that the prospect of transnational movement becomes mundane: young people... come to see migration as a normal part of the life course."[22] Even those who choose to stay may be impacted. Emigrants may return and buy up land with their savings, displacing agricultural workers who have no choice but to join the next wave of migrants.[23] Even if they do not return, the émigrés may send remittances which can affect the economy in their sending community. Fadloullah, Berrada, and Khachani (2000) (cited in De Haas (2007)) find that remittances are "used for daily expenses, conspicuous consumption and so-called 'non-productive' investments... which would spur inflation and not create employment... leading to a dangerous, passive dependency on remittance income." In so doing, remittances change consumptive expectations and introduce new cultural norms to sending communities.

In the receiving country, social norms may be affected by sustained in-migration, further facilitating migration. For example, jobs at the bottom of the job hierarchy often become associated with migrant workers, and consequently become difficult to fill them with anything other than migrant workers (Piore (1979), pp.34-35). Firms have no choice but to seek more migrant labor. Further, several studies note the disparity between the stringent enforcement of migration laws at the border versus relatively lax practices inside territorial borders (Andreas, 2000; Cornelius, 2001). This observation arises from the fact that "intimate contact with members of minority groups in the form of friendships can reduce levels of willingness to expel legal immigrants from the country."[24] More specifically, law enforcement can become selective in response to receiving country firms reliance on foreign labor. Wells (2004) finds that state institutions will be more inclusionary toward immigrants where they

[21]D. S. Massey (1987, p.113)

[22]Kandel and Massey (2002, p.981)

[23]Mines and Martin (1984) and Reichert (1981) as cited in Fussell and Massey (2004)

[24]McLaren (2003)

"comprise a valued part of the local economy, where they are well-connected to native-born residents and community institutions, and where the political culture is liberal and inclusive" (p.1338). In a similar vein, Ellermann (2005) finds that an immigrant's local context (i.e. social network) affects the degree to which restrictive immigration policies such as deportation are implementable because "the coercive uprooting of individuals... threatens to undermine the legitimacy of bureaucratic decisions."[25] Such observations arise because local governments have significant power "to facilitate, inhibit, or otherwise affect the quality of immigrant absorption."[26] In effect, by being a member of non-migrant social networks, migrants can significantly impact migration outcomes by restricting options for policymakers through pressures exerted on their behalf by natives. The social context of migration clearly matters.

2.1.1.3 Political Factors

Finally, we must consider the impact of political factors on international migration. Such factors may be structural in nature, arising from governments and their institutions, or events-based, such as specific conflicts and the attendant threat of violence. In the former case, governmental actors make and enforce migration policies in response to domestic politics, affecting the flow of migrants in and out of their country. Because migration significantly impacts so many different aspects of society, it "is a potent cross-cutting issue ... that defies the standard liberal-conservative divide and that produces fierce internal conflicts for the major parties," resulting in what Tichenor (2002) aptly describes as "uneasy coalitions of odd political bedfellows in which distrust and rival interests abound."[27] The net effect is that the state and its institutions are simultaneously incentivized to restrict certain types of flows, and facilitate others.

On the one hand, state agents may be incentivized to control and impede flows for eco-

[25] Ellermann (2005, p.3)

[26] Glazer (1998, p.65) cited in Wells (2004)

[27] Tichenor (2002, p.39)

nomic, socio-cultural, or national security concerns (Tichenor, 2002; Cornelius & Rosenblum, 2005; Teitelbaum, 1995; Rudolph, 2003). Native workers of the receiving country may pressure politicians to restrict flows out of fear of labor competition. Early studies into this issue found that immigrant labor does indeed compete with native workers, especially for low skill jobs. This depresses wages and increases anti-immigrant sentiment among pressured natives (Borjas, Freeman, Katz, DiNardo, & Abowd, 1997; Scheve & Slaughter, 2001; Malhotra, Margalit, & Mo, 2013). More recent studies, however, have shown that these effects are not as strong as previously thought, and that immigrants may actually increase low skill wages *in net* (Peri, 2011; Raphael & Ronconi, 2007; Hinojosa-Ojeda & Center, 2010; Hainmueller, Hiscox, & Margalit, 2015). Yet, what may matter politically is not reality but perceptions, and the evidence is strong that immigrants are often perceived, however unjustly, as an economic threat to natives (Esses, Dovidio, Jackson, & Armstrong, 2001; Blinder, 2011; Ekins, 2013; OECD, 2010).

Receiving country natives may also harbor so-called "compositional concerns" – a fear that immigrants from different cultures will alter the receiving country's social fabric in undesirable ways (Hainmueller & Hopkins, 2014; Dustmann, Preston, et al., 2000; Scheve & Slaughter, 2001; Zolberg & Woon, 1999; Layton-Henry, 1992). These fears may be conditioned by a variety of factors including income, education, ethnicity/race, contact, and cultural affinity/similarity (Frasure-Yokley & Greene, 2013; Hainmueller & Hopkins, 2015; Iyengar et al., 2013; Ramakrishnan, Esterling, & Neblo, 2014). And, depending on the population studied, such concerns may be more or less potent than economic concerns in affecting anti-immigrant sentiment (Malhotra et al., 2013). Regardless, politico-cultural arguments can motivate anti-immigrant sentiment among natives, which can influence their political representatives to act acccordingly.

Finally, national security concerns can be a powerful reason for restrictive immigration policy. In the American case, economic and cultural reasons have been used to justify restrictive immigration since the dawn of the 20th century, but it was not until the advent of the Cold War that national security began to play a role in such policies (Rudolph, 2006).[28] More recently, 9/11 and the rise of Islamic extremism have caused debates over immigration policy to increasingly be framed in terms of national security (Tumlin, 2004; Rottman, Fariss, & Poe, 2009; Koslowski, 2004; Givens, Freeman, Leal, et al., 2008). Europe has likewise experienced the securitization of its post-9/11 immigration policy due to a wave of terrorist incidents in countries with large immigrant populations from the Middle East – France, Germany, and the UK (Givens et al., 2008; Milton, Spencer, & Findley, 2013; Goldenziel, 2010). As of this writing, national security and immigration have taken center stage in political contests in all three countries, with potential to significantly influence policy in the years to come.

Given such justifications, the State and its agents are politically incentivized to regulate migration flows, and this may be accomplished in one of three places. First and most obviously, flows may be controlled at the border through effective policing. Second, unlawful entrants and visa violators/overstayers may be detained and deported from further inland. Finally, by changing the incentive structure for potential migrants (e.g. being very effective with border control and deportation), flows may be deterred from ever occurring (Cornelius, 2001; Cornelius & Tsuda, 2004; Ellermann, 2005; Willis, Predd, Davis, & Brown, 2010).

Additionally, it is worth noting that the state on the sending side may wish to control flows as well. Totalitarian and authoritarian regimes have long held this practice, the most egregious of these being preindustrial China, Korea (aka the Hermit Kingdom) and Japan (the Sakoku policy); Nazi Germany; the entire Communist Bloc during the Cold War; and Cuba and North Korea in the present day. Even relatively liberal states may wish to control the outflow of their citizens because they are literally losses of human capital. Whether losing raw labor, i.e. brawn, or suffering from so-called "brain-drain", unregulated outflows can have potentially disastrous consequences for a country's long-term prospects. Recognizing these problems, traditional sending countries have implemented a range of different mitigation strategies. For example, Fitzgerald (2008) describes how Mexico deals with the loss of talent and labor due to international forces outside its control (namely U.S. immi-

[28] Historically, national security concerns have long been a justification for restrictive entrance policies. Examples include Korea (aka the Hermit Kingdom), fedual Japan, Imperial China, to name a few

gration policy). Similarly, De Haas (2007) documents the strategies adopted by Morocco to mitigate emigration and actually harness the significant Moroccan diaspora to further internal development.

On the other hand, domestic pressures can also incentivize state actors to facilitate migration, in particular by targeting certain flows while excluding others. Domestic firms, for obvious reasons, have an interest in cheap migrant labor and may lobby policymakers to serve that interest.[29] Indeed, "many of today's strongest migratory systems were initiated through deliberate, government-sponsored recruitment of "guestworkers" during the 1940–1970 period."[30] And such governmental facilitation of migration continues today, throughout the developed world. The WWII era Bracero program, which brought Mexican agricultural laborers to the US, finds modern spiritual successors in the H1B (specialty) and H2A (temporary agricultural) visas in the US, the H2 (laborer) and F6 (marriage migrant) visas in South Korea, and the Schengen visa in the EU.

Outside of economic concerns, interest groups such as religious institutions and NGOs can restrain the ability of state actors to fully control the flow of migrants (Castles, 2004; Cornelius & Tsuda, 2004).[31] In particular, in liberal democracies, acceptable discourse concerning immigration policy is bounded so that some options simply become too taboo to even consider (Freeman, 1995). Interest groups may attenuate restrictions through lobbying or performing acts of activism on behalf of migrants. Ellermann (2005), for example, finds that even though migrants themselves may not have much political power, the communities into which they become embedded can exercise power on their behalf through advocates who use "case mobilization" to "appeal to bureaucratic actors to exempt particular individuals from implementation."[32] General civil rights movements, not specific to mgirants, and in-

[29] Tichenor (2002, pp.39-40): "During the Gilded Age, capitalists, such as Andrew Carnegie, described the flow of tractable immigrant workers into the country as a "golden stream." Today's capitalists and free market conservatives draw the same conclusion, arguing that the nation's economy benefits from foreign workers willing to do jobs and accept wages that U.S. citizens would not."

[30] (Cornelius & Rosenblum, 2005, p.102)

[31] Castles (2004) extensively covers why migration policies fail, incl role of interest groups

[32] Freeman (1995, p.7)

dependent judiciaries may also hinder restrictive policies (D. Massey, 1999). Pressure can also come from outside individual states. Hansen and Weil (2001) point out that there is a broad convergence among liberal democracies toward more inclusionary nationality policies because of the contradiction between liberal norms and large disempowered segments of the population (with countries facing international as well as domestic disapprobation). Sassen (1996) similarly argues that the state is losing its ability to restrict immigration due to "the growing accountability of the state to international human rights codes and institutions."[33]

On a longer time-scale, immigrants and their children may themselves become a potent political force, exercising political power as any other citizen – in office and at the ballot box (Togeby, 1999; Goldsmith & Holzner, 2015; Merolla, Pantoja, Cargile, & Mora, 2013; Levin, 2013). Not surprisingly, studies find that immigrant communities tend to be strongly pro-immigrant in their policy preferences (Stewart, Caren, Quinn, & Young, 2013; Pantoja, Ramirez, & Segura, 2001; Zepeda-Millán, 2014; Just & Anderson, 2014). As these communities become larger and gain more clout, elected officials must respond as they would any other constituency, by representing their interests and courting their vote (Jacobs, Martiniello, & Rea, 2002; Jiménez, 2008).

Given the worldwide record of less than complete success, it is tempting to dismiss, out of hand, the State as relatively ineffectual in controlling migrant flows. However, it is worthwhile to reconsider. On the one hand:

> The record is littered with the wreckage of government interventions that appeared to work reasonably well at first but had little staying power, or that had long-term consequences exactly the opposite of the initial, intended effects. These interventions rarely dry up "unwanted" migration flows or even significantly reduce them; more often, they simply rechannel the flows and create more opportunities for people smugglers to cash in on the traffic. (Cornelius & Tsuda, 2004, p.41).

[33]Sassen (1996, p.101)

And, as Soysal (1994) and Sassen (1996) point out, the international consensus on liberal norms severely constrains what states can do to flex their muscle in restricting migration.

On the other hand, even though there is a gap between what states intend and what they achieve, the number that actually matters is the size of flows pre- and post- enactment (Zolberg, 1999). In other words, the significance of restrictive power is not whether a state intended to completely stop immigration and only managed to stop 100,000, but that 100,000 were stopped rather than none. Viewed in this light, the state arguably does hold significant power, though its magnitude is hard to measure due to the nature of the counterfactual. The historical record supports this point: Tichenor (2002), in his expansive review of American immigration policy, shows that changes in government institutions, entities, parties etc. have significant impact on flows. And it is important to remember that even though international norms and supranational organizations do impose limitations, it is still the state that is "the immediate guarantor and provider"[34] of rights, and that absent the use of force, the only supranational enforcement mechanism is disapprobation. While states may be compelled to conform to liberal expectations in the long-term, they still have considerable leeway in the short-term to exercise sovereignty, subject only to domestic constraints.

Leaving behind structural considerations, the other class of political factors are what I call event-based. These political factors most typically involve outbreaks of violence and almost always increase the flow of migration.[35] The predominant theory describing this type of migration is the threat-based decision model. In it, potential migrants "migrate away from a conflict on the perceived threat to their personal security. When the perceived threat to their security increases beyond an acceptable level, they migrate away."[36,37] Note, this

[34]Soysal (1994, p.143)

[35]Williams and Pradhan (2009) find that this effect actually depends on the *level* of violence confronting individuals. In their case, they find people tend to stay home for low to moderate levels of violence, while high levels trigger the flight response.

[36]Williams and Pradhan (2009, p.3)

[37]See also: Shellman and Stewart (2007), Davenport, Moore, and Poe (2003), Moore and Shellman (2004), Melander and Öberg (2006), Moore and Shellman (2004), Davenport et al. (2003), Apodaca (1998), Schmeidl (1997), Gibney, Apodaca, and McCann (1996), Weiner (1996), Edmonston and Lee (1992), Clark (1989), Zolberg, Suhrke, and Aguayo (1992), Stanley (1987)

idea is not very different from the microeconomic equation 2.1.

Persons undergoing event-based political migration are typically called refugees. However, the terms "political migrant" and "refugee" are not necessarily synonymous. First, not all who migrate for political reasons are refugees in the common sense of the phrase. Modern history is replete with examples of political migrants who voluntarily leave their originating countries to pursue homeland political reform in more liberal settings. "Exile" is a more suitable name for such migrants. We may add to the obvious example of Vladimir Lenin cases such as Tomáš Masaryk of Czechoslovakia, numerous opposition politicians from Ba'athist Iraq, and Poland's post-WWII government in exile.[38] In many ways, such exiles often share less in common with the desperate family fleeing politically motivated death squads than with the masters who control them. It therefore seems wrong to lump them into the same lay category of "refugee," regardless the shared notion of politically motivated migration.

Another reason refugees and political migrants are not always the same thing is that not all refugees are motivated by political reasons. The person fleeing famine or epidemic is just as much a refugee as the one fleeing death squads. Though these causes are categorically distinct from each other, they are thematically conjoined by the issue of agency. That is to say, the political elite-activist-exile most likely had more choice in deciding to migrate than the typical refugee. It is therefore useful to discuss this issue of choice since it comprises the next major typology used in migration studies: voluntary versus involuntary migration.

2.1.2 Forced versus Voluntary

Individual agency has long been a separating criterion for migrant typologies (Petersen, 1958; Price, 1969; Kunz, 1981; Richmond, 1988). Even in these earliest studies, there is a clear recognition that politically motivated movement is a subset of the refugee phenomenon which itself is but one type of involuntary movement. Petersen (1958), for example, distinguishes between flight, displacement, the slave trade, and the "coolie" trade. While Petersen's latter

[38] Shain (2010, p.XIII-XXIX)

two categories may seem anachronistic, modern forms of both exist in phenomena such as human trafficking and contractual labor migration.

Later studies have examined agency in more granular forms. Kunz (1973, 1981) employs what he refers to as kinetic models to classify different levels of involuntariness resulting in categories such as anticipatory, acute, majority, alienated and reactive migration. Richmond (1988) goes further to argue that agency must be conceptualized as a continuum because "all human behavior is constrained... by the structuration process within which degrees of freedom of choice are limited."[39] In essence, all migration is involuntary to a certain degree because choices are always subject to constraints.

Richmond further argues that agency on its own is an insufficient categorizing criterion because though "there may be exceptional cases where... causes can be identified as 'purely' economic or political... the majority of population movements are a complex response to the reality of a global society in which ethnoreligious, social, economic and political determinants are inextricably bound together."[40] This idea is echoed by Kunz (1973): "It is not infrequent that the loss of liberty or danger to life is preceded by gradual economic restrictions on the whole society or some sections of it. In such cases the anticipatory refugee may be mistaken for a voluntary migrant seeking better opportunities."[41]

What Kunz and Richmond point to are the limitations of any single typology due to the complex interrelated causes of migration. Indeed, the blending of categories is not limited to cases of politically and economically persecuted refugees. As (King, 2012b) points out, "Intra-EU migration can be classified as both internal mobility... and as international... Temporary migration can morph into permanent settlement... irregular migrants can become legalised through special schemes."[42]

Finally, though not the focus of this research, so-called environmental refugees and move-

[39] Richmond (1988, p.14)

[40] Richmond (1988, p.14)

[41] Kunz (1973, p.132)

[42] King (2012b, p.8)

ments caused by health crises also fall into the gray area between forced and voluntary migration. Such non-traditional causes can directly impact migration; for example: "the potential for migration when linked to an increase in sea level is considerable.. hurricanes, rains and droughts ... over a long period of time... make migration the only possible option for the population affected."[43] They may also amplify and be amplified by more proximate causes such as war and poverty (Kalipeni & Oppong, 1998). And, like outbreaks of violence, they may inhibit migration under certain conditions, rather than encourage it.[44]

2.1.3 Initiation versus Perpetuation

Until now, I have been loosely using terms like "facilitate" and "impact" to describe the effect of factors on migration. It is worth adding some precision to these terms by distinguishing between factors that *initiate* migrations versus those that *perpetuate* it.[45] The distinction is easily seen in the example of macro-economic theory. To wit, if that theory were one that described initiation, then we would see migrant flows between all countries, indeed all regions between which there exists a wage/capital differential. Clearly, we do not. On the other hand, the evidence strongly suggests that given the existence of a flow, such a gradation plays a large role in maintaining it. More generally, macro-economic theory and several of the other economic factors, should be considered factors that perpetuate but do not necessarily initiate flows.

Initiating factors are most easily identified when there is a clear connection between intent and outcome. In the political domain, direct government intervention in the form of labor or marriage migrant programs are clearly initiating factors. A similar social factor

[43] Piguet (2008, p.7)

[44] Smith, Wood, and Kniveton (2010, p.1): "By contrast, studies of migration of agricultural populations in the Sahel have shown that rather than encouraging migration, decreases in rainfall (and the subsequent bad harvests) tend to limit the ability of households to invest in long-distance movement [6,7]. As a result it has been argued that there is considerable uncertainty in the prediction of climate change induced migration [8,9]."

[45] According to Willekens (2013), the distinctions between "theories explaining the perpetuation of migration fows and the emergence of migration corridors and migration systems was introduced by D. S. Massey et al. (1993)."

arises in the "channelization" story when a pioneer migrant becomes a recruiter for follow-on migrants from his home community. However, not all examples are so easily seen. World Systems theory has an initiation component because it points to countries with intertwined (colonial) histories as most likely to share migrant flows. Speaking broadly, any theory that can explain why certain flows occur while others do not is to some extent a theory that addresses initiation. This leads to another point: these two types of theory are not mutually exclusive, and many theories feature both components.

In particular, many theories which explain both onset and perpetuation are referred to as "path dependent" factors, which are processes that in their occurrence produce conditions that raise the odds of their perpetuation and lower the probability of their reversal. Hansen identifies two ways this may happen: through "lock-in," when certain options are rendered wholly unattainable by original choices, and... 'disincentive effects," when original choices make future options not impossible but deeply unattractive."[46] For example, in the economic domain, businesses may initially be attracted to cheap migrant labor, but over time develop a dependence on it, breaking away from which is extremely unpalatable. Or a sending community may become so reliant on remittances that it becomes unable to sustain itself in their absence.

With respect to politics, policies may unintentionally create feedback effects in the form of political/bureaucratic incentives for key state actors. For example, Andreas (2000) finds that the enactment of more stringent border controls is initially intended as a counter to migration flows but over time, becomes a self-perpetuating course of action. He finds that the net effect of an intensified border control campaign is to decrease illegal crossing by a negligible amount (since would-be migrants simply try until they succeed) while actually incentivizing bureaucrats to "catch and release" since such a strategy pumps up their apprehension numbers. In a truly perverse turn of events, the policy justifies its own ineffectual continuation which in effect, facilitates the persistence of migration, through what, in effect, amounts to inaction.

[46] Hansen (2002, p.271)

What is unique about path dependent factors is that though they are initially a consequence of or reaction to migration, "in the course of migration [they] come to function as independent causes themselves."[47] Consequently, path dependence is a phenomenon which is not restricted to some specific domain but occurs across social, economic and political processes. As path dependent factors take on lives of their own, their contribute to the persistence of international migration by facilitating future migration or hindering forces which would restrict it.

2.1.4 Level of Aggregation: Micro versus Macro

One final typological scheme worth considering is level of aggregation (LoA). Macro theories, as the name suggests, operate at the highest level, on and between states and state-level actors. At the opposite end, micro theories act on the fundamental units of migration, individuals. In between, mezzo theories act on intermediate units of aggregation such as social networks, subnational institutions and specific locales (D. S. Massey et al., 1993; King, 2012b).

As with the preceding categorization schemes, there are instances where LoA provides meaningful boundaries and others where it is less clear. For example, the LoAs for macroeconomic and microeconomic theories are obviously contained in their names. On the other hand, world systems and dual labor market theories interact states with cities, firms, and individuals, thereby spanning multiple levels.

Explicitly recognizing the LoA of a theory is crucial to understanding how a theory may interact with others, even if not so envisioned by its author. Capital flows may in fact push and pull migrants, but that doesn't mean that individuals don't have a high degree of agency in the migration decision. As D. S. Massey et al. (1993) puts it, "because theories... of international migration posit causal mechanisms at many levels of aggregation, the various explanations... are not neessarily contradictory unless one adopts the rigid position that

[47] D. S. Massey et al. (1993, p.448)

causes must operate at one level and one level only."[48]

Further, LoA analysis is useful because it helps the researcher delineate the assumptions, implications and limitations that arise from theory or a methodology. For example, the model of political refugees used in Williams and Pradhan (2009) grants micro actors (individuals) two choices (to hide or flee) in the face of potential violence from a monolithic macro-level actor, the Maoist insurgency. However, their model omits a third option, which Williams and Pradhan themselves point out, of joining the insurgency. LoA analysis suggests that this discrepancy arises from the following question: how would they model the "join" option, given that individuals and the insurgency are operationalized at two different levels of analysis? As a consequence, the implications of their study then seem limited only to cases where the "join" option is not available to the potential refugees.

Within the context of this discussion on LoA, clearly we need an approach to studying international migration that can span multiple levels of analysis. More broadly, IM research would benefit greatly from a methodology that allows for the interaction of theories from multiple disparate backgrounds, operating on and between different units and levels of analysis.

2.2 Agent Based Modeling

Agent Based Modeling (ABM) is an analytical method well suited to dealing with the myriad complexities of international migration. In simple terms, ABM is a computer simulation method in which numerous programs, the "agents," are designed to respond to stimuli from the environment and each other, then allowed to behave according to those rules to see what aggregate patterns arise. The agents model the most basic units of whatever phenomenon the researcher is studying and typically feature some non-deterministic decision making capability that can give rise to macro level variations, even among simulations with identical starting conditions (Takahashi & Goto, 2005; Barbosa Filho, Neto, & Fusco, 2013).

[48] D. S. Massey et al. (1993, p.454)

ABMs share the same end goal as other quantitative social science methodologies, which is to "[model] the mechanisms and causal pathways that generate the phenomena we want to explain or predict."[49] However, they are different in their point of departure since they conduct "'social science from the bottom up' as compared with the 'top-down' view of traditional social science."[50] Whereas equation based models (EBMs) such as regression analysis and structural equation modeling (SEM) focus primarily on relationships and causes, "ABM identifes key *actors* (individuals, institutions) in the population and uses hypotheses on their behaviour to infer the behaviour of the population," so that "attention is shifted from factors to actors."[51] In practice, what this means is that "ABM begins, not with equations that relate observables to one another, but with behaviors through which individuals interact with one another... direct relationships among the observables are an output of the process, not its input."[52] This approach is reasonable so long as the relationships and structures one is interested in are generated by the interactions of individuals rather than through some external force. In the case of migration, it seems pretty clear that the former is the case, rendering ABM an appropriate research methodology.

ABMs have been used in a wide range of disciplines including computer science, biology (Gardner, 1970), management science (Parunak et al., 1998; Prietula, Carley, & Gasser, 1998), and criminology (Patrick, Dorman, Marsh, et al., 1999), not to mention many of the social sciences. The latter includes economics (Topa, 2001); sociology (Schelling, 1971); demography (Billari, Prskawetz, Diaz, & Fent, 2008; Zinn et al., 2012); and political science (Cederman, 1997; Kiel, 2005). Though varied in specific topic, researchers from each of these fields have studied migration through the prism of ABMs (Smith et al., 2010). Economists, for example, have used ABMs to examine migration vis-a-vis labor (Espíndola, Silveira, & Penna, 2006; D. S. Massey & Zenteno, 1999) and housing markets (Sun & Manson, 2010). In the field of demographics and human geography, residential mobility (Heppenstall, Crooks,

[49]Willekens (2013, p.1)

[50] (Epstein and Axtell (1996) as quoted in Kiel (2005, p.273)

[51]Kiel (2005, p.273)

[52]Parunak, Savit, and Riolo (1998, p.10)

See, & Batty, 2011) and climate change (Smith et al., 2010) are some of the topics that have been studied. Within political science, ABMs have been used to study the impact of political violence on migration (Makowsky, 2006).

Regardless the specific topic, ABMs typically follow certain guidelines that separate them from other simulation methods. Takahashi and Goto (2005) summarize some foundational guidelines given by Langton et al. (1989) and Epstein (1999).

Langton et al. (1989)

- The model consists of a population of simple agents.
- There is no single agent that directs all of the other agents.
- Each agent details the way in which a simple entity reacts to local situations in its environment, including encounters with other agents.
- There is no rule in the system that dictates global behaviors
- Any behavior at levels higher than the individual agents is therefore emergent.

Epstein (1999)

- heterogeneity: agents differ in some defined preference set
- autonomy: behave without a system of top down control
- explicit space: agents interact on some explicit space such as a landscape or N-dimensional lattice
- local interaction: agents interact with near neighbors but not with distant neighbors on the landscape
- bounded rationality: "bounded rationality" assumed of human actors. This limited rationality is enacted through simple rules and the limitations inherent in agents using only local information and engaging in only local interactions

MAPES adheres to the Langton and Epstein rulesets with two reasonable modifications to accomodate characteristics of international migration. First, I relax the dictum that agents may only interact locally and with local environments. This restriction makes sense for agents modeling bacteria or geese. However, it is less sensible for a topic involving human communications, such as migration. In our case, there clearly can and should be communication between non-local actors. Whether friends from the same hometown, or an employer and prospective foreign hire, persons in different locales communicate and share information pertinent to the migration process. To disallow nonlocal communications, therefore, would be to render the model unrepresentative of the phenomenon.

One could react to this seeming incompatibility between the ABM patriarchs and our particular problem domain by throwing out that particular dictum. However, this would be overly reactionary because there are still interactions that occur in migration that should obey that rule (viz. physical and many social interactions). The more reasonable approach is to expand the definition of "local" to include communications media, in addition to the physical/geographical domain.

A related issue is Epstein's rule that agents should use only local information in their decision making. This too should be relaxed given the topic of study. For example, a person contemplating a move will probably consider employment prospects in potential destinations before making their decision. This is not local information in the geographic sense, but it is local in the context of the topic of study. After all, in the modern world, it is trivial to do an Internet search to search for jobs on the opposite side of the world. As with interactions, this study takes the stance that permitting such a relaxation of "local" still satisfies the spirit of Langston and Epstein's dicta.

This does not mean that we are violating the stricture against global behaviors and autonomy. Agents may have access to information in other locales but only to the extent that they would actually do so in the real world. For example, a prospective migrant may learn from his buddy abroad that there is an opening in his construction crew, and it would be fair to model this. What the proper ABM does not do is give the agent perfect information so that he makes decisions knowing that the true unemployment rate in his destination is 95.49% and his probability of gaining employment is 32.33% (repeating, of course). Allowing adaptations such as the former but not the latter, MAPES adheres to the overall spirit of

Langston and Epstein's approach.

So, what does an ABM look like? Figures 2.1-2.2 depict the seminal Sugarscape ABM (Epstein & Axtell, 1996) in its simplest form. The model description is remarkably succinct. There is one class of agent, which moves around a 2-dimensional grid (aka the Sugarscape) and consumes sugar. The logic for these actions is "from all lattice positions within one's vision, find the nearest unoccupied position of maximum sugar, go there and collect the sugar."[53] The only required environmental rule is that sugar regenerates according to one of several possible parameterizations (e.g. immediately, gradually, dependent on sugar levels in surrounding grid), but the basic idea is simply that it regenerates. As their book progresses, Epstein and Axtell layer on additional levels of complexity to model reproduction, disease, and trade. However, at its core, the simplest rules are all that is required for Sugarscape to work. What makes ABMs beautiful is that even such a simple model is "sufficient to generate macrostructures of interest... [that] can lead to hypotheses of social concern."[54]

[53] Epstein and Axtell (1996, p.25)

[54] Epstein and Axtell (1996, p.53)

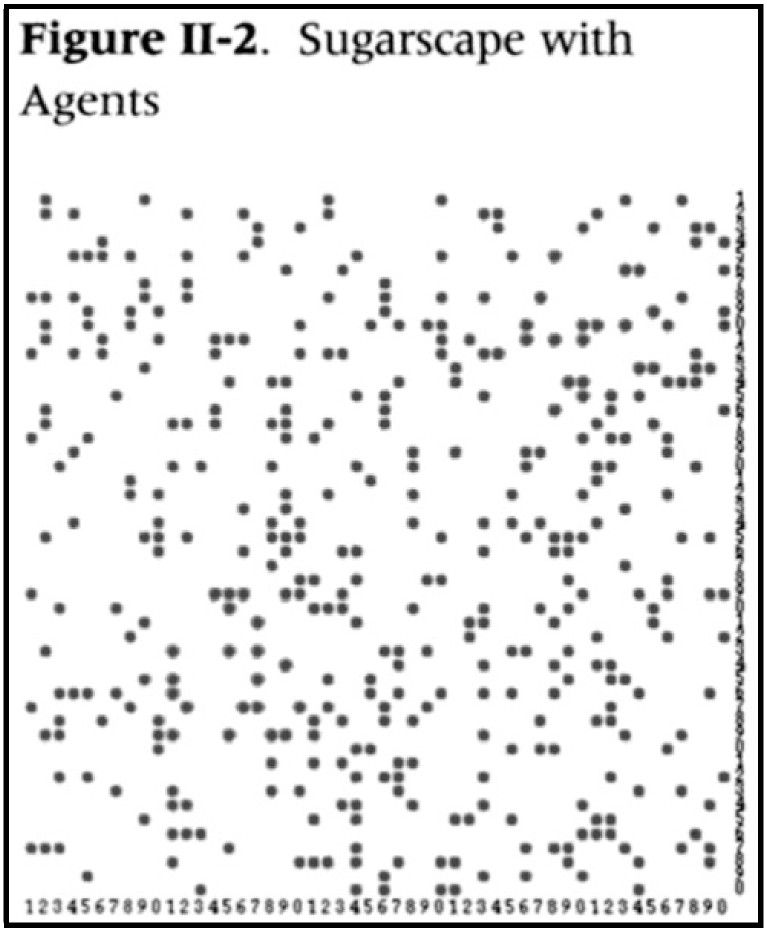


Figure 2.1: The original Sugarscape diagram from Epstein and Axtell (1996), p.22

<u>Agent movement rule **M**</u>:

- Look out as far as vision permits in the four principal lattice directions and identify the unoccupied site(s) having the most sugar;[8]
- If the greatest sugar value appears on multiple sites then select the nearest one;[9]
- Move to this site;[10]
- Collect all the sugar at this new position.

Figure 2.2: The agent movement rule for Sugarscape Epstein and Axtell (1996), p.25

Obviously, the agents in MAPES are quite a bit more complex than the Sugarscape agents. Despite the complexity, the underlying principles of heterogeneity, autonomy, explicit space, local interaction and bounded rationality are all the same. Before considering the details of the agents used in MAPES, let us consider the strengths and weaknesses of the ABM methodology, especially as it pertains to the social, political and economic aspects of migration.

2.2.1 Strengths

2.2.1.1 Representative of the phenomenon

One of the primary strengths of ABMs is that "the agents in an ABM correspond one-to-one with the individuals... in the system being modeled, and their behaviors are analogs of the real behaviors."[55] In other words, ABMs are ontologically representative of the phenomenon of migration (Bankes, 2002). This is crucial because the aggregate migration flows that are observed are the result of a collection of individual migration decisions. Each migrant weighs the factors that are important to him and makes his decision in ways that may be idiosyncratic or independent of what other migrants are doing. Just because one person heavily weighs factor X in their decision to migrate does not mean that another will do the same.[56] By design, ABMs mimic this exact process. Each agent collects the data pertinent to his decision making, then follows his programming to make decisions and act in discrete ways much as his real-world counterpart would. Outcomes are the result of these decisions and actions, not the calculated result of some "god" equation that defines what the outcome should be given the inputs.

EBMs (equation-based models), on the other hand, do not concern themselves with process. Rather, they focus on the ultimate relationships between inputs and outputs. They may use data at the individual level but rather than model the intermediate steps that lead

[55] Parunak et al. (1998, pp.11-12)

[56] To put it more eloquently, "one agent's behavior may depend on observables generated by other individuals, but does not directly access the representation of those individuals' behaviors, so the natural modularization follows boundaries among individuals."(Bankes, 2002)

to some outcome, they can only surmise how much an input matters in achieving an output by looking at overall patterns across all individuals. At its most literal, the approach taken by EBMs implies that the potential migrant examines some formula and decides that he will migrate with 28% probability since he is 23, male, high school educated and unemployed, and each of those factors should be weighed in such and such a way. EBM approaches focus entirely on how *much* things matter rather than *how* they matter, and therefore miss the opportunity to "provide a place to express the enormous amount of data and knowledge about the behavior, motivations, and relationships of social agents, [whether] individuals or institutions" allowing them to "[exploit] exactly that category of information which is the focus of many social sciences."[57] They may be accurate at prediction, but there is a certain dissatisfaction that results from simply plugging data into a mathematical black box without consideration for the individual level process.

But beyond that emotional satisfaction, there is a deeper analytical advantage resulting from the process-focused approach of ABMs. Most non-trivial phenomena, whether from physics or political science, tend to be non-linear in nature. A multi-step process like the decision to move may be the result of a series of decisions which are simple to represent individually, but taken in sum, lead to outcomes that have complex, non-linear relationships with inputs of interest. The ABM has no trouble here since it works off the simple individual steps. The EBM, however, works off the final observables and will have a much more difficult time applying its analytic process to the situation. Further, even for relationships that can be intuited and represented, the involvement of a differential equation or two can easily render the relationship mathematically intractable (Kiel, 2005, p.285).

Working with EBMs, much of the toil of the analyst occurs in picking the correct functional form or set of covariates on which to run their equations. This is because EBMs must make a large number of assumptions about a whole host of factors that are more of mathematical rather than substantive relevance. What is the functional form of the relationship? Are all the inputs strictly additive? What should the distribution of errors be? As a result,

[57] Bankes (2002)

the analyst often becomes bogged down in massaging the analytical method rather than modeling the substance. ABMs do not suffer this problem because no such assumptions are made. Agents behave according to simple rules which either are sufficient in representing the behaviors they model, or are not. Yes, the analyst will still spend a great deal of time fixing their ABM, but the time will be spent focusing on the substantive processes involved rather than the mathematical minutiae of the method.

A further advantage of ABMs is that they are better able than EBMs to model complex taxonomies (i.e. hierarchies of relationships). This is important because migration affects and is affected by multiple levels of aggregation. Individual migrants live in communities with natives as neighbors, and their social, political and economic interactions constitute their local community. Several of these communities comprise larger geographies such as cities and countries, each of which is possessed of a government and economy, among other characteristics. And, each of these aggregate entities is itself composed of individuals, all leading to multiple overlapping levels of aggregation. That such a phenomenon would be difficult to represent using traditional EBMs is a serious understatement. An equation-based method such as hierarchical regression would, at the very least, require assumptions about how each individual characteristic of interest is distributed across each of the relevant levels of aggregation. Then parameters for those distributions would need to be estimated, as well as hyperparameters for the parameters themselves. Each distributional assumption that must be made and value that must be estimated represents a potential point of failure. This is not to even mention that it is difficult to conduct time series analysis in hierarchical contexts, and not clear how one would model dynamic associations (e.g. a person migrating).

ABMs, on the other hand, do not face these difficulties because taxonomic relationships do not have to be fit to some complicated equation, merely depicted in a sensible fashion. In an ABM, everything can be an agent whether the individual or the government of the city in which he lives or the legislator governing the country. Each agent can have as many associations as required, then simply go down the list during each period of the simulation and follow whatever programming logic derives from each association. The individual agent does not require some overarching picture of the entire simulation, just the characteristics of each agent it interacts with. Once the ABM has been run, analysis only needs to decide the extent to which the outcomes of interest meet expectations. These analyses can range from dead-simple (e.g. contingency tables and simple graphs) to various types of regression, but the key is that they can be agnostic with respect to the complex relationships that generated the data; they only care about the gap between what is and what is desired. Finally, the analyst may adjust the ABM to better meet expectations by simply changing an association or an action, rather than fiddling with complex equations.[58]

None of this is to downplay the power of the traditional methods; they have many merits and appropriate use cases. As Parunak et al. (1998) put it, "EBMs may be better suited to domains where the natural unit of decomposition is the observable or equation rather than the individual" whereas "ABM's are better suited to domains where the natural unit of decomposition is the individual rather than the observable or the equation."[59] Indeed, this study uses EBMs to analyze the output of its ABMs, when it becomes necessary to shift focus from process to relationships. The overall point here is that in the case of migration, the preponderance of research has been done using EBMs, which have brought to light many interesting relationships. However, ABMs are more suitable to taking the next step, to validate those previous findings and shed further light on how they work.

2.2.1.2 Causality and scenario analysis

As with most simulation methods, establishing causality is relatively trivial in ABMs because the practitioner has 100% control over all aspects of the simulation. If a change in input or the model itself leads to a change in output, then one knows exactly what initiates the

[58]Parunak et al. (1998) provide an excellent example of this point regarding taxonomy. They model a supply chain involving firms and suppliers and discuss how ABMs can accurately represent boundaries between tiers that EBMs cannot: "Each firm has its own agent or agents. An agent's internal behaviors are not required to be visible to the rest of the system, so firms can maintain proprietary information about their internal operations. Groups of firms can conduct joint modeling exercises while keeping their individual agents on their own computers, maintaining whatever controls are needed. ... It is much less likely that such a firm would submit aspects of its operation to an external "equation manager" that maintains specified relationships among observables from several firms."

[59]Parunak et al. (1998, pp.11-12)

causal chain. This is not to say that the full causal mechanism is understood, because there may be intermediate processes that lead to the ultimate effect. However, identifying these is a matter of iteration (i.e. observe what else changes, manipulate and repeat) rather than supposition or statistical sleight of hand.

A more practical consequence of this characteristic is that using ABMs it is possible to conduct so-called "scenario analyses" which Takahashi and Goto (2005) define as "possible outcomes by a policy alternative in a given situation and... a mechanism that results in a specific outcome."[60] In other words, insofar as the analyst is able to build an accurate ABM of a situation of policy interest, he is also able to assess the impact of various policy alternatives by precisely identifying their effects and the mechanisms through which these effects occur.

As a germane example, one could study the effect of a physical wall at the US-Mexico border using a sufficiently well built ABM. In the simplest case, one might model the economies of bordering geographies and compare economic outputs for the status quo versus a completely impermeable wall. One could then introduce gradations of effectiveness to see how the porousness of the wall affects outcomes: Linearly? Non-linearly? In undesired ways? A next step could be to build out the ABM to include other factors such as local politics, social networks, education, and cost of healthcare, then examine the impacts of walls of varying effectiveness. The varying levels of wall effectiveness would each constitute a different policy scenario, and the analyst may study as many outcomes and intermediate steps as he wants, provided only that he is willing to build an ABM of sufficient complexity. Constants defining the various factors at play would be drawn from the actual reality on the ground, increasing external validity.

It would be far more difficult to accomplish the same tasks using an EBM approach. An EBM would need to first draw data from analogous real-world situations so that there is something to test the equations on. This of course assumes that there are a sufficient number of analogous situations to be found. And even if there are, most likely there they will not be

[60]Takahashi and Goto (2005, p.3)

perfectly analogous to the policy in question (aside from Israel and Cold War Berlin, how many borders are there in the world with significant physical walls?), or the variables for the various cases will not line up; there will inevitably be major constraints due to lack of data.

This point regarding variables is a significant one. Policies can affect outcomes through complex processes, often unexpected, so that without enough data, the analyst faces a difficult challenge in convincingly linking intervention with effect. EBMs have been employed for so long that there have been developed many sophisticated approaches to establishing causality, but they are nowhere as simple as for ABMs. As it pertains to the specific topic of migration, "While in-depth survey-based approaches have been developed that work to disentangle the multiple factors influencing migration... they do not allow predictions ... under different conditions from those under which the original surveys were performed. However, dynamic approaches such as agent-based modelling provide a means to adjust various parameters to further investigate situational changes and future scenarios."[61]

2.2.1.3 Emergence

If one of the goals of social scientific research is to discover new patterns, perhaps the most compelling argument for using ABMs is that they can exhibit emergence, which are system-level behaviors that arise due to the interactions of many independent constituent parts (i.e. agents) comprising that system. The classic example is the movement of a flock of birds – there is no overarching equation governing (distinct from describing) their movement, and yet through the simple rules followed by each individual bird, we see that there are aggregate patterns in movement. One can imagine, and in fact it has been shown, that ABMs can exhibit emergence in social science contexts, including with respect to migration (Parunak et al., 1998; Bankes, 2002; Bonabeau, 2002).

For example, Espíndola et al. (2006) study the micro economics of migration and report that the macro-economic Harris-Todaro model is emergent from their ABM, as is the phenomenon of reverse migration. Likewise, Zhang and Jager (2011) use an ABM to study

[61] Smith et al. (2010)

population dynamics in rural Britain, and their model exhibits the emergent phenomenon where "the migration of a single family may cause a cascade effect, causing more people to migrate and having a negative effect on the vitality of this municipality and the quality of life of the remaining residents" and that these "emergent effects sets the stage for the other residents in deciding to stay or to go."[62] Their model also exhibits path dependence in that "the context of a declining population the presence of a social need causes the population to initially decrease at a slower rate, until a critical population size is reached, after which the population decrease accelerates."[63] Emergence means that ABMs can be used as a powerful tool for both validating extant system-level theories from the ground up as well as discovering new directions for research.

2.2.2 Limitations

2.2.2.1 Emergence (again)

While ABMs are able to exhibit emergence, there is often a lack of rigor in defining what counts as such. Of the previously mentioned papers, Espíndola et al. (2006) addresses this issue better. They present a series of graphs showing (fairly convincingly) that net migration stops when the wage gap disappears. However, beyond this "eyeball test," no metrics are presented to support their declaration of mission success. Zhang and Jager also make similar claims on the basis of visual impression alone, stating that "the results demonstrate clear spatial patterns"[64] and "[this is] visually very clear."[65] To be fair, their paper deals with a far more complicated phenomenon than Espíndola et al. (2006) so the looseness of the analysis may be a direct consequence of that fact. But these papers are exemplars of the statement by Bankes (2002) that "most papers... rely on human observers to declare emergence to have occurred... Formal definition of what is meant by emergence is the exception rather than

[62] Zhang and Jager (2011, p.6)

[63] Zhang and Jager (2011, p.45)

[64] Zhang and Jager (2011, p.31)

[65] Zhang and Jager (2011, p.36)

the rule, and quantitative tests that a given model achieves the sort of emergence advertised are rare."[66]

It is obviously difficult to formulate a universal test of emergence because what counts as such depends on the phenomenon being studied, so will vary from case to case. On the other hand, if one is able to provide a convincing quantitative definition of emergence for one's specific project, there is no reason standard hypothesis testing tools cannot be used. This dissertation *applies* ABMs but does not *innovate* in their application, which is to say that I do not claim to present some novel way to quantify emergence. Still, I minimize "hand waving" arguments to the extent possible and where I can, I quantify what I believe to be emergent patterns using traditional hypothesis testing.

2.2.2.2 External Validity

Another critical challenge in using ABMs is the difficulty faced in establishing external validity: whether an ABM is actually representative of the real-world phenomenon it claims to model. More specifically, are (1) "the theories and assumptions underlying the conceptual model... correct and the model representation... reasonable for its intended purpose" (conceptual validity); and (2) Is "the behavior of [the] model... sufficiently accurate for its intended purpose" (operational validity) (Takahashi & Goto, 2005, p.10). Of course, establishing validity is not a challenge isolated to ABMs – any model that grants significant freedom to the builder in choosing specification, whether the simplest linear regression model or a complex model of fluid dynamics, faces this problem.

Unfortunately, the problem is especially acute for ABMs because fundamentally, an ABM is something of a conjuration, a set of rules authored by the researcher to represent the phenomenon he is studying. There are no hard constraints from existing data, so as long as the rules somewhat resemble reality, a plausible argument can be made that the model is "valid." This means that it is quite possible for the researcher to "twist and tweak [his]

[66] Bankes (2002, p.7200)

behavioural assumptions until the model yields the desired output."[67]

Given this high degree of freedom, how do we define "reasonable?" For example, in this project, person-agents endeavor to maximize utility drawn from political, social, and economic factors. Is this a reasonable portrayal of the real-world human beings these agents are supposed to represent, and if so, who makes that determination and on what basis? Further, once the author and reader/arbiter determine that a model is theoretically valid, how do we decide what are reasonable values for the associated weights and constants that ultimately specify the ABM? Unfortunately, there are no easy answers to these questions because "there is no perfect means of validating ABMs [as] validation is a type of a *social process*."[68]

Klabunde and Willekens (2016) offer the sage if somewhat obvious advice that "if modellers follow theory more closely they do themselves a favor: Their models become less vulnerable to the criticism of being arbitrary... The more closely behavioural rules are in line with theory, the less valid is this critics' argument."[69] This advice may be more fully embraced by relying on existing data as well as theory: the actual weights, constants, and other numerical devices that implement theory should use real-world data where possible. An example of doing this is seen in Hao and Mitchell (2013), who study "rural-to-urban interprovincial migration in China [using] empirical data of the 2000 census as well as provincial statistics." They validate their model by comparing its output with known "trends and patterns from the census... [because they] are similar, the ABM provides strong evidence to support [their] theoretical framework."[70] In short, the practitioner should use existing theory and data to predict known outcomes so that predicting hypothetical outcomes is more believable.

Willekens and Klabunde also (rightly) place great importance on the social aspects of establishing validity; ultimately, "validation" is a matter of whether persons outside the

[67] Klabunde and Willekens (2016, p.15)

[68] Klabunde and Willekens (2016, p.11)

[69] Klabunde and Willekens (2016, p.15)

[70] Hao and Mitchell (2013, p.2)

research team are willing to believe the results. They reasonably point out that "if the rules in agent-based models are close to the way rules are implemented or described in other types of non-ABM research, it becomes easier for non-ABM researchers from that field to understand and use the results from an ABM."[71] Further, they emphasize that ABMs, like all research to a certain extent, are about dialogue. First, stakeholders/research consumers provide the researcher with overall direction and researchers provide results to satisfy that appetite. The former party then gives feedback on their belief/skepticism, to which the researcher responds by explaining why theoretical and model specifications were chosen.

Given that this dissertation is the first iteration of this particular line of research, I do not yet have any feedback to satisfy. My intent, however, is to make the model and analyses underlying the dissertation as accessible as possible so that it is relatively easy to replicate and validate the findings I present. In so doing, my hope is to garner useful feedback and respond to it in a constructive manner, thereby improving the model and its acceptance within the research community.

[71] Klabunde and Willekens (2016, p.15)

CHAPTER 3

Project Design

The main objectives of this dissertation are to propose and validate a baseline ABM of migration that incorporates political, economic, and social factors and to demonstrate how it may be used for future research on international migration. I describe and implement this model, then run it with different sets of input parameters comprising different scenarios to validate model mechanics and support or undermine existing IM theories. I make no grand innovations in theory; rather, the innovation of this dissertation lies in the laying of a foundation on which theories from the many different subfields of migration studies can all interact. To the extent that it is successful, this model can serve as a testbed enabling us to know which combinations of "behavioural models [do] not work to create a particular aggregate outcome pattern... [and] get closer to finding the ones that do."[1]

3.1 Objectives

This project has two overarching criteria for success. The first is that given the agent (micro) level rulesets, the model should behave consistently with its real-world analogues at the system (macro) level. For example, (1) in most cases persons should flow from locations of low to high economic opportunity, and (2) larger social networks at a destination from a particular source should correlate with more in-migration from that source. There is clear documentation of these phenomena in the IM literature (Harris and Todaro (1970) and D. S. Massey (1987), respectively) so the criterion would be satisfied to the extent that what is observed in MAPES qualitatively matches what is seen in the literature. Given the

[1] Klabunde and Willekens (2016, p.13)

breadth of possible behaviors that may be observed from these systems, a full list of possible behaviors will not be enumerated ex ante; it will, however, be made clear in the analysis what qualitative behavior a particular set of model output is being compared to.

If an observation from the model does not obviously comply with observations from the real world, it will still be considered a success if the deviation from expected behavior is consistent with the substantive logic of the model (i.e. is not an implementation error) and therefore can be manipulated in predictable ways. As a simplistic example, MAPES may indicate that persons flow from a location of high to low expected utility under certain conditions. While it may be tempting to attribute such behavior to a flaw in the model, it is also possible that the unexpected outcome is actually some substantively interesting edge case that holds theoretic value. It could be the case that imperfect information (which is designed into the model) is causing persons to make decisions that would be rational were the information timely.[2] Or, the difference in expected utilities might be dominated by some secondary factor such as an unusually high cost of moving or time-value discount. In either case, once I understand what the mechanical cause is, I should be able to manipulate and/or mitigate the unexpected behavior. In such a case, I would consider the criterion of consistency to have been met.

In a reductionist sense, this first criterion simply asks whether the model works as intended. This goal of achieving reasonable system-level behavior, however, is not trivial because the agents in an ABM, by definition, act without a system of top down control (i.e. they exhibit autonomy). The modeler can only manipulate micro-level inputs to indirectly affect the system-level outcomes of interest. Further, an agent cannot be said to be modeled properly if it cannot interact satisfactorily with other agents, no matter how well-behaved it may seem in isolation. While each agent is simple on its own, it and all the other agents and environmental factors it interacts with must be tuned properly such that they are also well behaved when operating in concert. The challenge lies in this high requisite level of coordination.

[2] Timeliness in this context refers to the individual receiving information that is *still* accurate. A piece of data may have been accurate at some point in the past but it may longer be timely.

The second major success criterion is that MAPES provide sufficient instrumentation to answer questions about the modeled systems at a low level of granularity. While implementation is obviously important, equally important is whether the model encompasses enough of the factors that matter to allow meaningful study of the substance. If the model is specified with enough breadth, the hope is that MAPES will allow us to answer questions such as the following:

Movement

- What are the primary determinants of the size and direction of flows?
- How do flows change with time?
- Are there subflows of interest or is movement essentially monolithic?
- Who is migrating, and why?

Political

- What is the effect of migration on the ideological make-up of the population?
- Ditto legislatures?
- How does the ideological composition of the legislature change in response to migration?

Economic

- How does migration alter the competitive landscape for jobs?
- How are people's economic outcomes altered due to migration?
- Does production overall increase or decrease?

Social

- How does migration alter social networks?
- How do social networks affect the flow of persons and why?

In order to answer such questions, the typical workflow will be to generate sets of simulation output using MAPES, and then use a variety of standard statistical tools to analyze the data. Obviously, graphical and regression analysis are employed extensively. Also, because all the data is generated via simulation, we can employ essentially experimental methods to tweak parameters of interest, rerun the model as many times as required and directly establish causality. Finally, because everything is an agent in MAPES, we can use what Takahashi and Goto (2005) refer to as "micro-dynamics analysis [which] attempts to explain why a particular outcome is generated from the viewpoint of the dynamics of micro-level parameters. Every macro-level behavior is formed by micro-level agents' behaviors that change depending on their internal decision making models" (p.9). In essence, we can trace agents' decision making and behaviors step by step to see exactly how they impact macro patterns.

3.2 Scenarios

I present the results of running MAPES on three different scenarios. A scenario is defined as a "set of critical experimental parameters that may have an effect on future system behaviors" Takahashi and Goto (2005, p.6). Within MAPES, parameters are defined at the simulation level and for collections of agents. Simulation-level parameters deal with characteristics of the simulation itself, and details of *classes* of agents (as opposed to *instances*). There are 22 such parameters that are tuned in this initial version of MAPES, but there are numerous places even in this early implementation of the model where more variables may be added. Obviously, as the model gets more elaborate in future iterations, the number of parameters can grow ad infinitum. The 22 parameters are listed in tables 3.1-3.2.

Simulation characteristics are simple things like how long to run the simulation, and run-time details that are not relevant to the substance of the model (e.g. filesystem paths for output, logs, etc.). Agent-class details are parameters that are shared across all instances of an agent (i.e. for a particular class of agent), regardless of affiliation (e.g. nationality, what city they are from, where they currently reside). These range from a seemingly small detail like the probability with which a person undergoes a social interaction, $Pr(interact)$,

Class	Variable Name	Type	Range	Default	Description
City	COST_PER_DIST	Double	0 – Inf	2	Cost (in units of simulation currency) per unit of movement
Country	ELECTION_TICKS	Integer	1 – STOP_AT	6	Number of ticks between legislative elections
Country	ENTRY_COST_MAX	Integer	0 – Inf	1000	Maximum value of entry cost into a different country
Country	PR_DEPORT_MAX	Double	0 – 1	0.1	Maximum per-tick probability of deportation of non-citizens
Economy	JOB_FILL_PCT	Double	0 – 1	0.75	Proportion of jobs to fill at tick 0
Job	TICKS_BEFORE_HIRE	Integer	0 – Inf	2	Number of ticks to collect applications before hiring if job is unoccupied
Person	ATN_MAX_D_IDEO	Integer	0 – 100	50	Max ideology gap between persons forming social tie
Person	D_IDEO_INT_MIG	Integer	-100 – 100	-30	Ideology change when migrating
Person	D_IDEO_NEG_FOR	Integer	-100 – 100	10	Ideology change for negative interaction with foreigner
Person	D_IDEO_POS_FOR	Integer	-100 – 100	5	Ideology change for positive interaction with foreigner
Person	D_IDEO_STP_SZ_DN	Integer	1 – Inf	4	Per tick change in individual's ideology towards mean of his social network, as 1/percent difference.

Table 3.1: Simulation Parameters

Class	Variable Name	Type	Range	Default	Description
Person	FFN_LOCAL_MAX	Integer	0 – Inf	20	Max number of social ties in a city
Person	FFN_LOCAL_MFMU	Integer	0 – Inf	10	Number of social ties for max social utility in a city
Person	FFN_PR_ADD	Integer	0 – 100	20	Probability of adding a local social tie per tick
Person	JOB_EDU_BANDWIDTH	Integer	0 – Inf	3	Bandwidth of acceptable education level for applying to a job
Person	MIGRATION_ENABLED	Boolean	0, 1	1	Is migration enabled?
Person	MIGRATION_TIMEOUT	Integer	0 – Inf	10	Minimum number of ticks to wait after migrating before migrating again
Person	MOVE_COST_DISCOUNT	Double	0 – 1	0.25	Proportion of move cost person must have in order to move
Person	PR_NEG_FOR	Integer	0 – 100	50	Probability of having a negative interaction, given a stranger is a foreigner
Person	PR_SOCIAL_INTERACTION	Integer	0 – 100	50	Probability of social interaction with a stranger
Person	UINC_MEAN_COEFF	Double	0 – Inf	1.5	Proportion of mean local income a person must have to have max income utility
Global	STOP_AT	Integer	1 – Inf	200	Number of ticks to run model

Table 3.2: Simulation Parameters (cont'd)

to obviously important ones like whether migration is allowed at all. For the purpose of this study, these parameters are kept constant at the values noted in tables 3.1-3.2 except where they are varied for comparative purposes (as noted throughout the presentation of results).

Aggregate-agent parameters are those that are set among collections of agents (typically grouped by geography) and determine characteristics for specific instances of agents. To clarify, by characteristics I mean the values each *instance* of an agent has for the variables discussed in chapter 4. In the case of a person-agent, these would be age, education, sex, etc. Aggregate-agent parameters are the parameters that describe how a collection of agents are distributed in those characteristics. In the case of person-agents, they are generated by originating city, with their education distributed normal with mean μ_{edu} and σ_{edu}; age distributed Bernoulli with p; etc. So for each city in a scenario, there are aggregate-agent parameters for all the relevant characteristics. There are different aggregate-agent parameters for all the different agent types. These are elaborated in chapter 4.

Within this dissertation, what differentiates a scenario from another is the combination of agent-class and aggregate-agent parameters that are used for a particular set of simulation runs. Across scenarios, only parameters relevant to the topic at hand are varied. The remainder are kept constant across scenarios. In brief, these scenarios are:

Bilateral Economic Gradient 2 countries, 1 city each. Relative to each other, one is a typical sending country, and the other is a typical receiving country.

Trilateral Equal 3 identical countries, 1 city each.

U.S.-Mexico Border – California A simple implementation of the Southern California - Mexico border. 3 cities on the US side, 2 on the Mexico side. Parameters drawn for 2014-2017 from governmental bodies (e.g. Office of Immigration Statistics, U.S. Department of Homeland Security; Instituto Nacional de Estadística y geografía) and 3rd party sources (viz. Numbeo.com).

3.2.1 Scenario 1: Bilateral Economic Gradient

In the bilateral economic gradient (BEG) scenario, there are two countries with one city each. Country A is a stylized receiving country – it has a relatively small workforce, and a relatively right-shifted distribution of jobs (i.e. high mean education level for the job distribution). Country B is a stylized economic migrant sending country, with low job opportunity, and a surplus of labor. Aside from the differences in job and labor endowments, the two countries are identical: respective legislatures are the same in size and political distribution, as is the political distribution of each starting population.

The primary expectation of this model is that there will be a net movement of persons from country B to A, caused by a gradient in average expected utility (which is a function of wage and other factors) between the two countries in accordance with macroeconomic theory. It is expected that net movement will stop when an equilibrium state is reached, removing the imbalance. There are likely to be two mechanisms leading to the equilibrium state, and even once that is reached, cessation will be of *net* movement, not *all* movement.

The first equalizing force will likely be the change in average per-country expected utility resulting from wage gains. Person-agents constantly seek to find better jobs than the ones they are currently in. If the occasion arises, they may change jobs in their current country, or move to the other country if the expected utility gain is great enough. This unit-level optimization will contribute toward reaching a global optimum where *average* expected utility is equal across the two countries. This means that the average person is indifferent to living in either country. When this happens, migration should still occur in a Brownian fashion, but there should be no net change in populations. Again, this is distinct from no migration occurring at all. In order for that to occur, it must not make sense for *any* individual p to make a cross border move (i.e. $EU_p^n < EU_p^r$ where $p \in$ all persons P, r is person p's country of residence, and $n \in$ all non-resident countries N). That requirement is clearly an edge case of (i.e. is more stringent than) $E\bar{U}^A = E\bar{U}^B$, which essentially says that it makes sense for the same number of people in country A to move to B as vice versa.

Simultaneous with the wage-related changes in $E\bar{U}$, there will also be a force exerted by socio-political factors that reduces the utility gradient between countries A and B. This will arise from immigration policy in the receiving country becoming restrictive to the point where the resulting reduction in inflow (from rising entry costs) and increase in outflow (from deportation) reduces net inflow to zero. In other words, this mechanism will alter other components of the utility calculation for each individual since his expected utility for a non-resident location n is calculated as $EU^n = (1 - Pr(deport))(U_{income} + U_{social})$ if his savings are greater than the cost of moving, and $EU^n = 0$ if savings are less.

The change in policy (which causes the change in utility calculation) will arise because the initial inflow caused by an open policy increases the number of foreigners in country A, increasing the probability of negative social interactions with foreigners (since Pr(negative interaction) is a function of, among other terms, the probability of having an interaction with a foreigner). An increasingly xenophobic native population will vote in more conservative legislators, who enact more restrictive policy. The "socio" part of the nomenclature arises from the fact that the rightward political shift is fundamentally caused by negative social interactions.

3.2.2 Scenario 2: Trilateral Equal

The trilateral equal (TE) scenario is one composed of three identical countries. Each has a single identical city, and each city has the same identical sets of citizens and jobs. In other words, all agents are identical across the three countries. The only differences from run to run are the initial social networks and job assignments. This is because agent-to-agent associations are randomly created at run time, rather than assigned as parameters (with the exception of location).[3]

My main expectation of this model is that there will initially be random migration between all three countries, with no net change in populations, similar to BEG after it reaches steady state. Then, if path dependence is a phenomenon that MAPES can actually simulate,

[3]This is for the sake of scenario parsimony. It is certainly possible to specify every single agent-to-agent association but that was found to be extremely tedious and ultimately irrelevant to meaningful results.

one of the countries will become a destination country for the other two, as some chance policy liberalization snowballs into a difficult to reverse pattern. Which country this occurs in should be random from run to run.

Once this occurs, I propose that there may also be contagion effects as liberalized individuals from the new destination move to the remaining 2 countries and spread their ideology through social networks. Whether and how quickly that happens seems likely to be completely random. On the one hand, a new liberal arrival may form his first social tie with someone very conservative and end up having his impact greatly attenuated (recall that members of social networks regress toward the group's ideological median over time). On the other, if coupled with a likeminded person, they can start spreading their ideology far more effectively. Be that as it may, enough sparks will light a fire, and I expect to see that truism in play here as a secondary effect.

3.2.3 Scenario 3: U.S.-Mexico Border – California

The final scenario to be explored is a MAPES representation of a real-world case: the Southern California-Mexico border (UMBC). The main purpose of this simulation is to show that MAPES can be applied to real cases and show results likewise reflecting reality. Like in BEG, there exists a clear economic gradient between the two countries but UMBC has additional complexities. There is a political gradient between the US and Mexico, with more restrictive immigration policy in place in the US, resulting in higher cost of entry into and higher probability of deportation out of the US.

There are multiple cities in each country, meaning intra-country migration is possible. In fact, we should observe a fair amount of intra-country migration since there are differences in market conditions between cities in the same country. This scenario also highlights the impact of geography. That is, given the elongated, vertical shape of the region, we should see a transit corridor running South to North, with closer (i.e. cheaper) moves, being more common than farther moves.

In terms of politics, there is a complication in that the "countries" in this simulation are actually just tiny regions of larger countries, not even fully constituting states in the real world. This means that the legislators representing this regions are also only tiny fractions of their respective countries' legislatures. One way to deal with this situation is to model the full legislatures of the respective countries behaving exogenously. This has the downside that analyzing the impact of legislators from the region under study becomes difficult as the signal would probably get swamped by the full legislatures. On the other hand, running the MAPES model as-is with the local legislators constituting whole legislatures overblows their importance in the big picture. Nonetheless, the latter course is what I follow because it facilitates analysis and seems a reasonable caveat for the purpose of this dissertation. Finally, there is an implicit assumption that national policy is enforced uniformly, and that there are no local effects on that policy. We know from the extant literature that this is not always true (see Glazer (1998), Wells (2004)), but it too seems a reasonable simplification for this model.

The overarching expectation for this scenario is that it will produce results that mirror reality or are clearly attributable to limitations of the model. This is obviously a very broad criterion, but I give it in lieu of enumerating an exhaustingly long list of irrelevant characteristics. Suffice it to say, model success will depend on not having nonsensical results such as more net migration to Mexico, or a complete closing of the border, or a city losing its entire population.

CHAPTER 4

The Model

MAPES is built as a composite of existing theories (where possible) and common sense assumptions in domains (1) not explicitly covered by migration theory; or (2) which may be covered by theories from fields too far afoot. An example of the latter might be the way I model the job hiring process: firms advertise openings, individuals apply if the job characteristics match their search criteria, and once a threshold number of applications is received, the firm hires a random applicant. Certainly there are theories from business / managerial science that cover this process in detail, but this common sense implementation seems adequate without consulting an entire separate body of literature.

The point of departure for this model is that migration and its consequences are fundamentally the result of individual conditions, actions and interactions. Individuals decide why, where and whether to migrate. Residents of the destination interact with the newcomers and compete with them for jobs. Persons from each group form political opinions and perform political actions on the basis of interactions with each other. Individual legislators respond to resulting electoral incentives to enact policies impacting longtime residents and newcomers alike. Each of these are therefore represented as agents in the model.

While migration is the sum of individual actions, much of what we observe and care about is at higher levels of agglomeration. We observe flows of people and demographic shifts in the populations of cities and countries. We measure economic changes in terms of unemployment statistics and GDP. We see changes in legislatures and policies enacted as a result of the political process. Each of these may be expressed as a combination of a collection of individuals and other characteristics inherent to the institution. For example, a city is primarily a collection of people, but it also has an economy which in turn is comprised of jobs, each of which may be associated with an individual (depending on whether it is filled). Likewise, a collection of cities comprises a country which has the additional characteristics of a legislature, economy, borders, etc. While such entities are not agents in the sense of having independent decision making ability (i.e. agency), their properties are what we often care about. They are therefore tracked as a sort of secondary, passive agent in this model.

Tables 4.1–4.2 catalogue the various IM theories reviewed in section 2.1. A subset of these are built into the ABM used for this project (as indicated in the final column of tables 4.1–4.2), but there is no reason why some or all the rest could not be incorporated in future iterations. Detailed specifications of the current implementation follow.

4.1 Agents

I define two primary classes of agent: individual persons, who may or may not become migrants, and legislators, who enact immigration policy reflecting the interests of their constituencies. Both of these agent types are multiply associated with and through several secondary agents: locations (viz. cities and countries) and jobs. All of these agent classes interact with each other in every time period (i.e. tick). One way to run MAPES is to randomly generate all agents at the outset of each run based on prespecified aggregate-agent parameter values, as discussed in section 3.2. I also maintain the ability to precisely specify individual agents in a given model run. A useful technique is to generate a random population using the aggregate-agent parameters, then feed those specific agents into all subsequent runs of the model so that initial conditions are identical. This technique is useful for bootstrapping confidence intervals for outcomes, and I use it extensively throughout this study. The model may be run with as little as a single individual and a single location, or as many of each primary and secondary agent type as possible, constrained only by the computing resources available.

Category	Factor / Theory	Driver	Flow Direction / Effect	Included
Economic	Macro-Economic	Labor demand	High to low	No
Economic	Micro-Economic	Wages	Low to high	Yes
Economic	New econ. of migration	Risk	High to low	No
Economic	Dual labor market	Secondary labor demand	Rotating door hiring in receiving locale	No
Economic	World systems	History & Inter-state relations	periphery to center	No
Social	Channelization	Social ties	Sending to daughter community	No
Social	Calculus of migration	Migration cost	Reduce cost of migration	Yes
Social	”Paisanaje”	Kinship, coethnicity	Increase social capital & utility in immigrant communities	Yes
Social	Cultures of migration	Longterm reliance on diaspora	Increase sending locale dependence on migration	No

Table 4.1: Factors affecting and affected by migrant flows

Category	Factor / Theory	Driver	Flow Direction / Effect	Included
Social	Local context	Perceived legitimacy of policy	Attenuation of policy enforcement	No
Political	Economic threat	Wage pressure	Anti-immigrant sentiment	Yes
Political	Compositional concerns	Foreign culture of immigrants	Anti-immigrant sentiment	Yes
Political	National security	National security threats	Anti-immigrant sentiment	No
Political	Sending side concerns	Loss of human capital	Controls on emigration	No
Political	Demand for migrant labor	Labor demand	Selective entrance policy	No
Political	Liberalism	Human rights, liberal ideology	Attenuation of policy enforcement; implementation of open policies	No
Political	Naturalization	Naturalized / incorporated immigrants	Immigrants directly affect receiving locale politics	Yes
Involuntary	Physical conflict	Physical threats	Away from sources of threat	No
Involuntary	Environmental	Resource shortages, physical threats	Away from sources of threat	No

Table 4.2: Factors affecting and affected by migrant flows (cont'd)

4.1.1 Individuals

Each individual is characterized by the following traits: sex, age, education, ideology, and financial endowment; the first is static and the rest can change as the simulation progresses. Ideology indicates an ideal point on the single dimension of immigration policy, with lower values (i.e. scale left) indicating preference for more liberal policy (e.g. lower cost, less enforcement, greater quotas). *SAVINGS* is simply the amount of money possessed by the person. Each person may be employed, in which case they earn income. If they are not employed, they may apply to any open jobs. An employed person may apply to other jobs as the occasion warrants (discussed below).

Variable	**Type**	**Range**	**Description**
ID	String	-	Unique ID
AGE	Integer	0 - Inf	Age
EDUCATION	Integer	0 - Inf	Education
SEX	Integer	0, 1	Sex
CITY_FROM	String	-	Unique ID of City From
CITY_IN	String	-	Unique ID of City In
SAVINGS	Double	-Inf - Inf	Cumulative Income
TICKS_SINCE_MIGRATION	Integer	0 - Inf	Ticks Since Last Migration
IDEOLOGY	Integer	0 - 100	Ideology

Table 4.3: Person Characteristics

Each person has several locational affiliations. First, they have a static affiliation to their origin, the city in which they start the simulation. They also have an initial nationality, which can change if they spend enough time in a foreign country to satisfy its naturalization requirement. Finally, at any given time, they resides in one of the cities (and by extension, countries) in the simulation. If they are a citizen of the country in which they are resident, they may vote; else they may not. Dual citizenship is not permitted.

The individuals in this ABM are motivated by political, economic, and social concerns. Therefore, their happiness (i.e. utility, U_{total}) and actions (e.g. migrating, voting) are functions of factors from each of those domains. Each person's political utility derives from their ideology and context (e.g. ideology of others, policy environment). People vote for politicians who will enact policies in alignment with their preferences regardless their electoral prospects (i.e. they vote sincerely). In order to vote, each individual evaluates all the legislators in their locality for proximity to their ideology, and votes for that with the closest ideal point. In the current version of MAPES, political utility is not included in the calculation of U_{total} (which affects the migration decision) because it was unclear whether and how other aspects of the political environment (e.g. democratic-ness of the country) would be included. Therefore the direct effect of politics on an individual's migration decision is felt through the economic utility and movement cost terms (discussed below), not a separate $U_{politics}$.

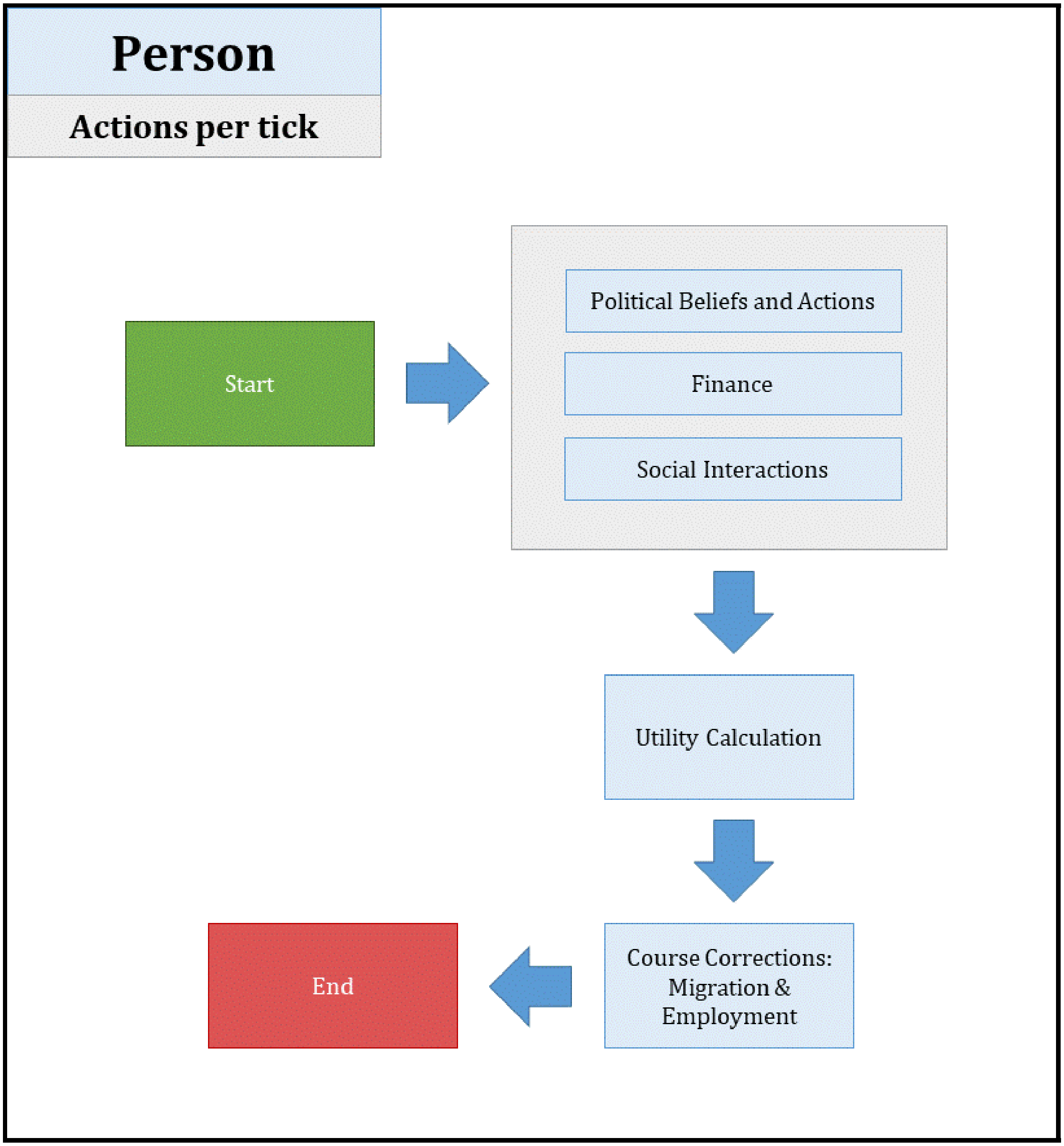


Figure 4.1: Person: Actions per Tick. Each person undertakes these action each tick.

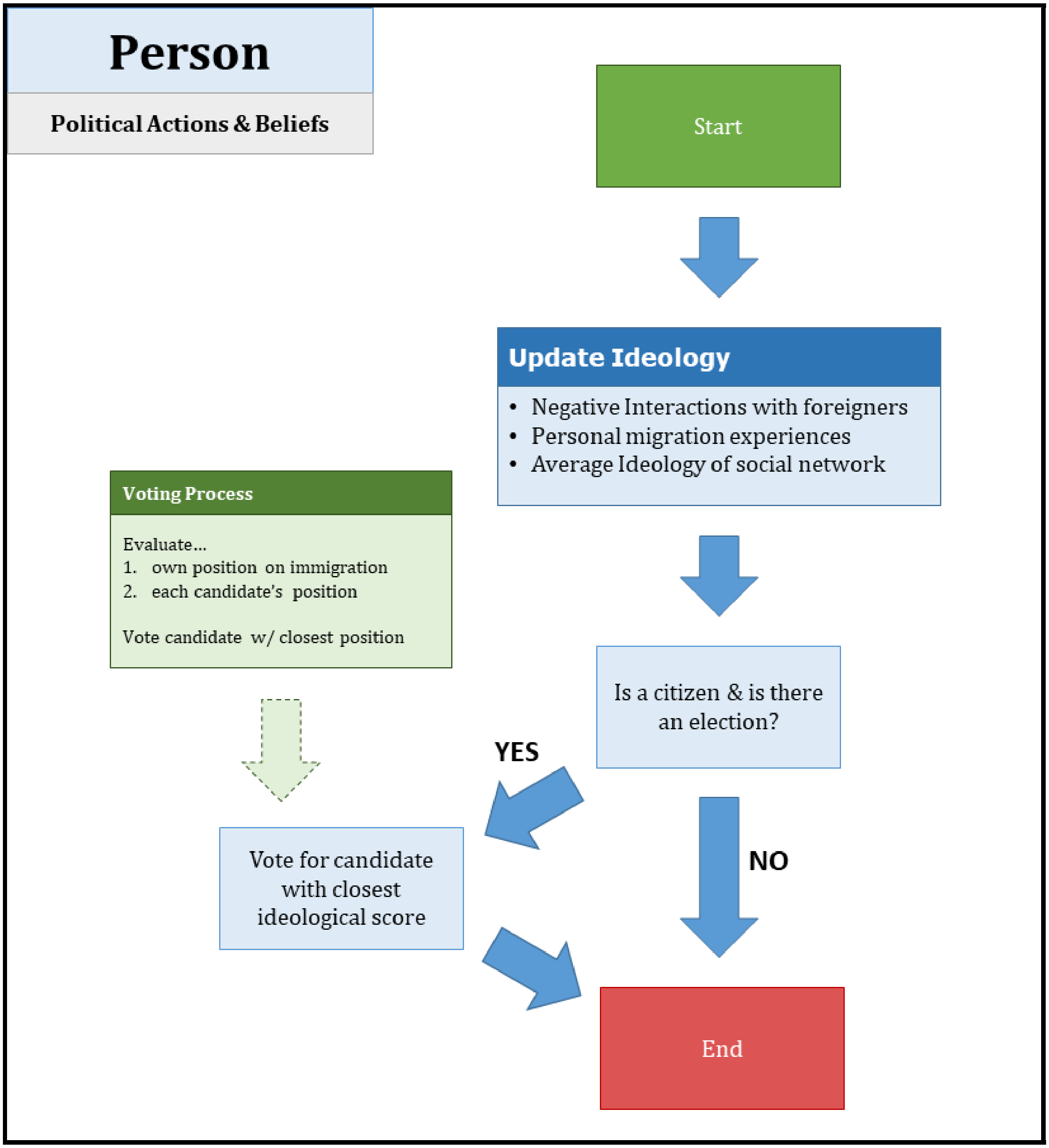


Figure 4.2: Person: Political Actions & Beliefs. Persons and the legislators they vote for each have an ideal point on a single ideological spectrum on which immigration policy is the sole issue. Elections are SNTV (single non-transferable vote).

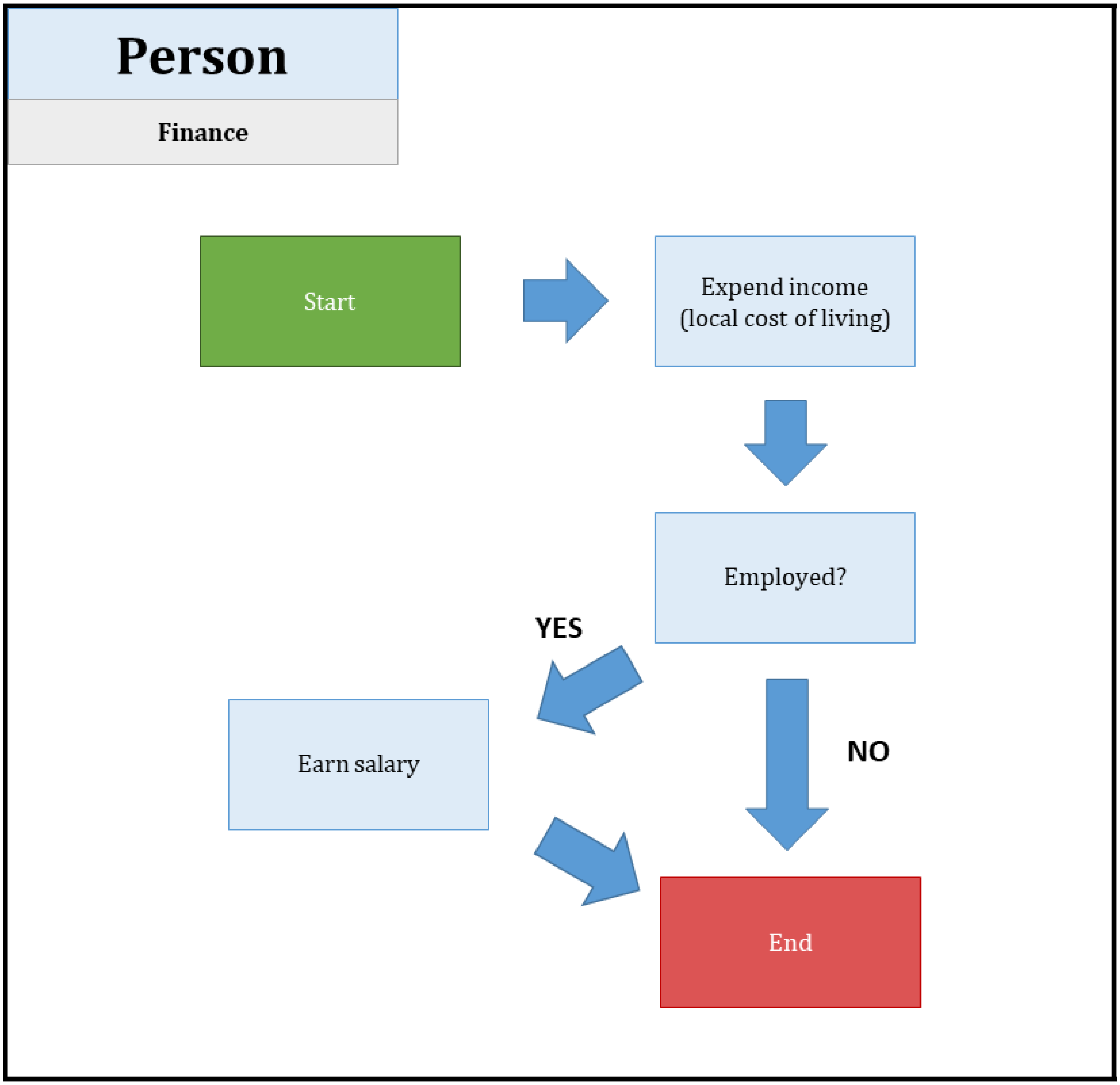


Figure 4.3: Person: finance. Persons work to earn money and spend that money commensurate with the local cost of living. If they are unemployed, they apply to local jobs.

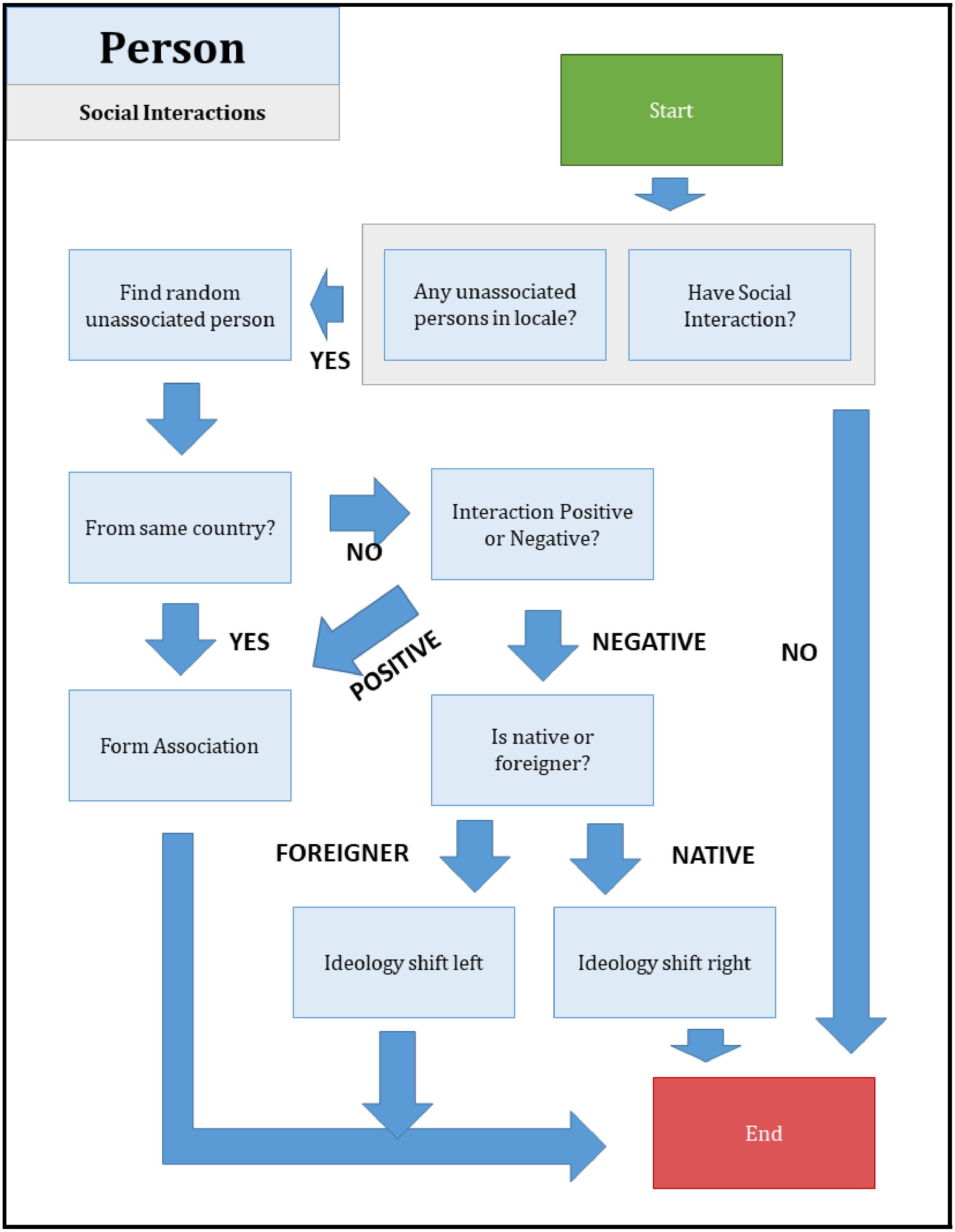


Figure 4.4: Person: Social Interactions. The social interactions a person experiences colors their ideological preferences.

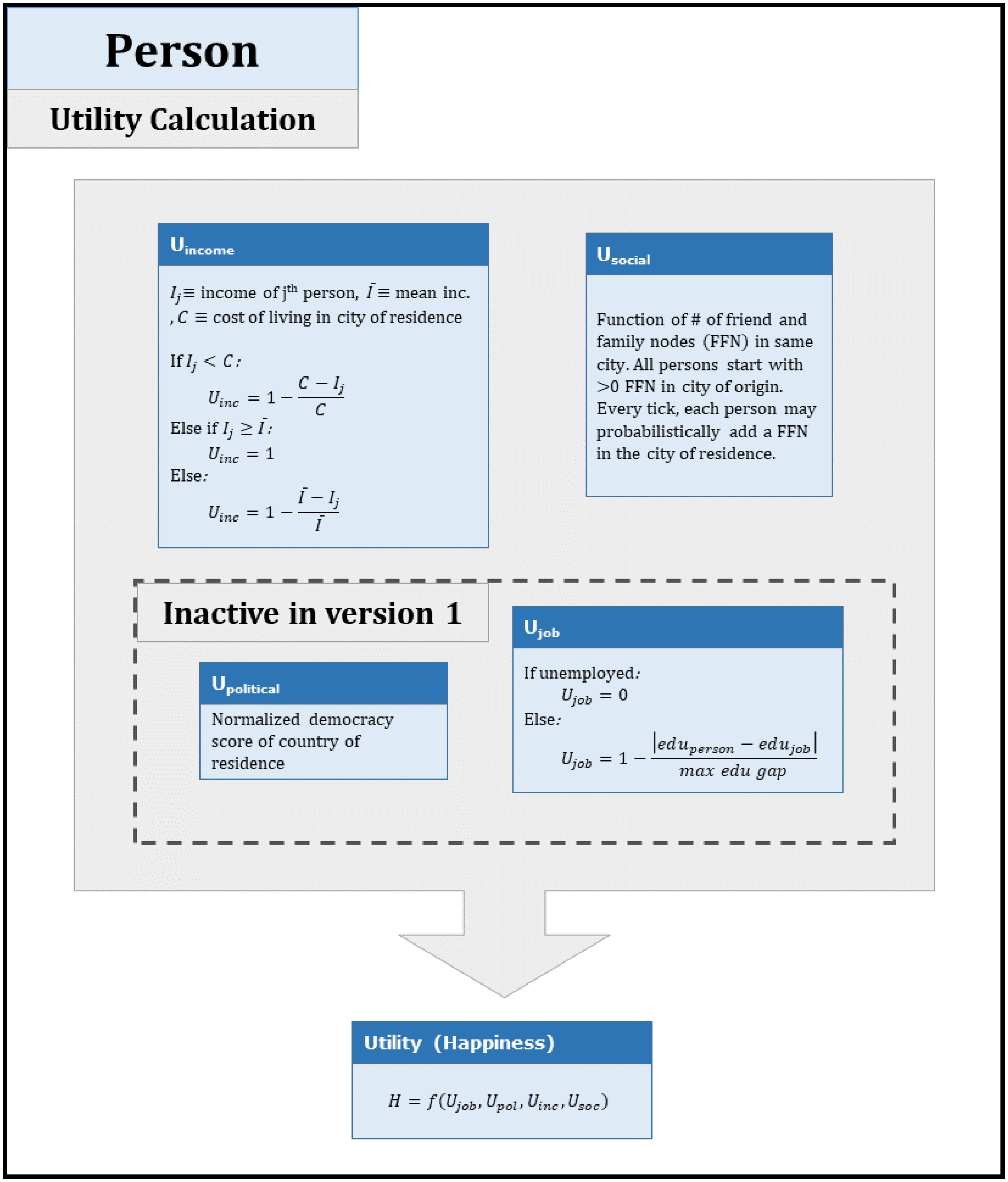


Figure 4.5: Person: Utility Calculations. Persons calculate their overall happiness (i.e. utility) each tick based on income and social factors. In future versions of MAPES, they may also incorporate factors such as political factors and job satisfaction.

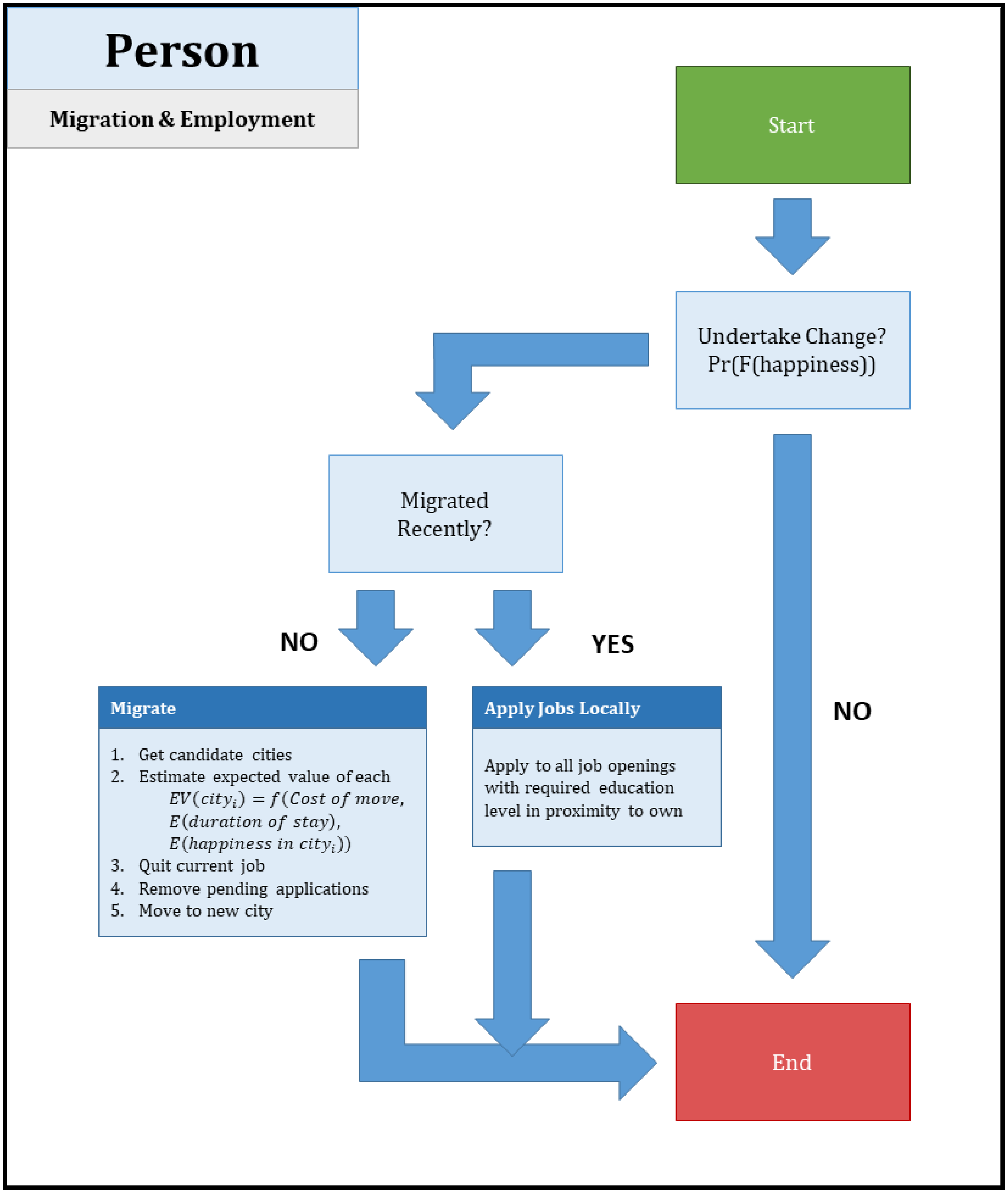


Figure 4.6: Person: Migration. Each tick, persons decide whether they are content based on their current happiness. If not, they may migrate and/or apply to new jobs to rectify the situation.

On the economic front, individuals prefer more of everything - wealth, security, opportunity. All else equal, they are happier in and migrate to locations that offer more of each. U^P_{income}, happiness deriving from person P's economic situation, is calculated as follows: if a person is evaluating their current U_{income} (i.e. in their current city), one of three calculations is made. First, if unemployed, no happiness is gained from economic factors. If their income is greater than some threshold income (mean income by default) in their city of residence, $U_{income} = 1$. Finally, if their income i is less than $\bar{i}$, local mean income, but greater than 0, $U_{income} = i/\bar{i}$; their happiness is proportional to how far below mean income they are. If the person is estimating prospective happiness in a city other than that in which they currently reside, they make one of two possible calculations. First, if their current income is greater than the expected income in the target city, the expected utility from income is the ratio of expected income to current income.[1] If current income is less than expected income, the expected utility from income is equal to unity.

Social factors are captured through means of a social network (i.e. graph). All individuals in the ABM may form social ties to other individuals, and the combination of all the individuals (i.e. nodes) and social ties (i.e. edges) forms the social network of the simulation. Each individual's happiness from social factors is directly proportional to the size of their local social network. Specifically, for person P, U^P_{social} = local network size^P/max network size, where max network size is a simulation-level parameter defining the number of local social ties an individual needs to be totally satisfied with their social life.[2] As mentioned earlier, each tick of the simulation, each person may form another social tie in their current city up to the max allowed network size. If a person has associates in a foreign city, their cost of migrating to that city is dicounted in proportion to their personal network size.

Every tick, each person performs several actions. First, they spends money equal to the cost of living associated with each city, and if they are employed, earn income from their job. Second, they probabilistically form a social relationship with another resident of their city of residence. Third, they updates their ideology to converge toward the average of those in their

[1] Expected income is calculated as the mean income across all open jobs in the target city

[2] "Simulation-level parameters" are discussed in section 3.2

social network. Fourth, they evaluate their current happiness (i.e. utility) and undertake a life change with probability equal to the ratio of actual happiness to the maximum possible happiness. Finally, if they decide to undertake a life change, they (1) migrate to a new city if they have not migrated recently; and (2) apply to new jobs.

The process of forming social ties works as follows. Each person P has a social interaction with someone outside their social network with probability $Pr(\text{interaction})$, which is a simulation-level parameter. If they are to have an interaction, a random person S is chosen from all strangers in their current city. If P is a native, one of two events occur. If S is also a native, a social tie is formed, meaning that they add each other to their respective social networks. If S is a foreigner (i.e. born in another country), the interaction is a negative one with probability $Pr(\text{neg. interaction})$, which is also a simulation-level parameter. In the case of a negative interaction, no social tie is formed and instead, person P becomes more xenophobic, increasing their ideology score (i.e. favoring more restrictive immigration policy) by a simulation-level parameter. With probability $1 - Pr(\text{neg. interaction})$, the interaction goes well and a social tie is formed. If P is a foreigner relative to their current city, they form a social tie as normal. Finally, there is a hard limit on how many social ties any individual can have at any given time. This is done for two reasons. The first is that in my opinion, this mirrors real life – given a large enough population, one simply cannot maintain meaningful ties with literally everyone, which is what would happen in a simulation run long enough without such a bound. Second, this limit also facilitates the use of this model for larger populations. Without it, calculations performed for all individuals based on their respective social networks would be of order N^2 per tick, which would get very computationally expensive, very quickly.

As to the migration decision, it is a deceptively complex process. Klabunde and Willekens (2016) provide an excellent overview of why this is the case. First, is deciding to migrate separate from or integral to determining a destination? Second, what is the order of operations given a model's answer to the first question? Klabunde and Willekens survey numerous modeling strategies to answer these questions, ranging from eliminating or severely reducing the location decision (D. S. Massey & Zenteno, 1999; Cai, Oppenheimer, et al., 2013;

Berman, Nicolson, Kofinas, Tetlichi, & Martin, 2004; Biondo, Pluchino, & Rapisarda, 2012; Barbosa Filho, de Lima Neto, & Fusco, 2011; Silveira, Espíndola, & Penna, 2006; Espíndola et al., 2006) at one extreme, to elaborate decision processes that require calculations for all possible destinations and attribute a great deal of strategic capability to each individual (Klabunde & Willekens, 2016; Hassani-Mahmooei & Parris, 2012; Schweitzer, 1998). As implied by the large range of approaches, there is no consensus model, leaving matters largely up to the discretion of the modeler. Klabunde and Willekens counsel that "it helps if there is some direct empirical justification for the factors included in the decision rule, but still the functional form of it remains arbitrary... We suggest to decide on one decision theory - e.g. utility maximization, the Theory of Planned Behaviour, etc. - and then to incorporate ideas from other disciplines as elements of the decision theory chosen."[3]

As suggested by Willekens and Klabunde, the primary decision theory used here is utility maximization. In order to migrate, the person evaluates all possible destination cities for expected utility. As earlier discussed, this overall utility is comprised of economic, social and (indirect) political components. The EU_{total} of each candidate destination is weighted by $1 - Pr(deport)$ where $Pr(deport)$ is the probability of being deported, determined by the immigration policy set by each country' legislature. The person migrates to the city with highest EU_{total}, with no migration occurring if that city is the city of residence.

There is no birth, death, marriage, or any other type of event which would impact the total number of persons in the simulation. The framework is amenable to such additions but they are not implemented here for the sake of keeping this initial model as simple as possible. I discuss the consequences of this choice in the discussion section.

4.1.2 Politicians, Legislators & Legislatures

Each politician (i.e. potential legislator) has an ideal point on the single dimension of immigration policy, which is set at the beginning of the simulation and does not change. In a typical simulation, there are several politicians associated with each city, and the number of

[3]Klabunde and Willekens (2016, p.13)

politicians in the city exceeds the number of legislative seats available so that elections are competitive. The legislature of each country is comprised of legislators from all of its constituent cities. For all the countries, elections are held periodically and the electoral system is single non-transferable vote (SNTV). SNTV was chosen because it achieves the desired electoral outcome (a legislative body comprised of more than one member per constituency) using a method which is easy to model: simply count votes and take the desired number of candidates with the highest votes. By contrast, a system like first-past-the-post would not have resulted in the desired multi-member constitutencies, and any system requiring more than one round of voting would have been more computationally expensive. None of the real-world problems associated with SNTV are pertinent to the model since voters have perfect information about their own ideologies and that of the candidates, and there is no strategic voting (i.e. voting to win rather than based on beliefs). After each election, the legislature sets its country's immigration policy at the chamber's median ideal point, ranging from completely open, to some upper (restrictive) threshold set as a simulation parameter.

Variable	**Type**	**Range**	**Description**
ID	String	-	Unique ID
CITY	String	-	Unique ID of City Located in
IDEOLOGY	Integer	0 - 100	Ideology

Table 4.4: Legislator Characteristics

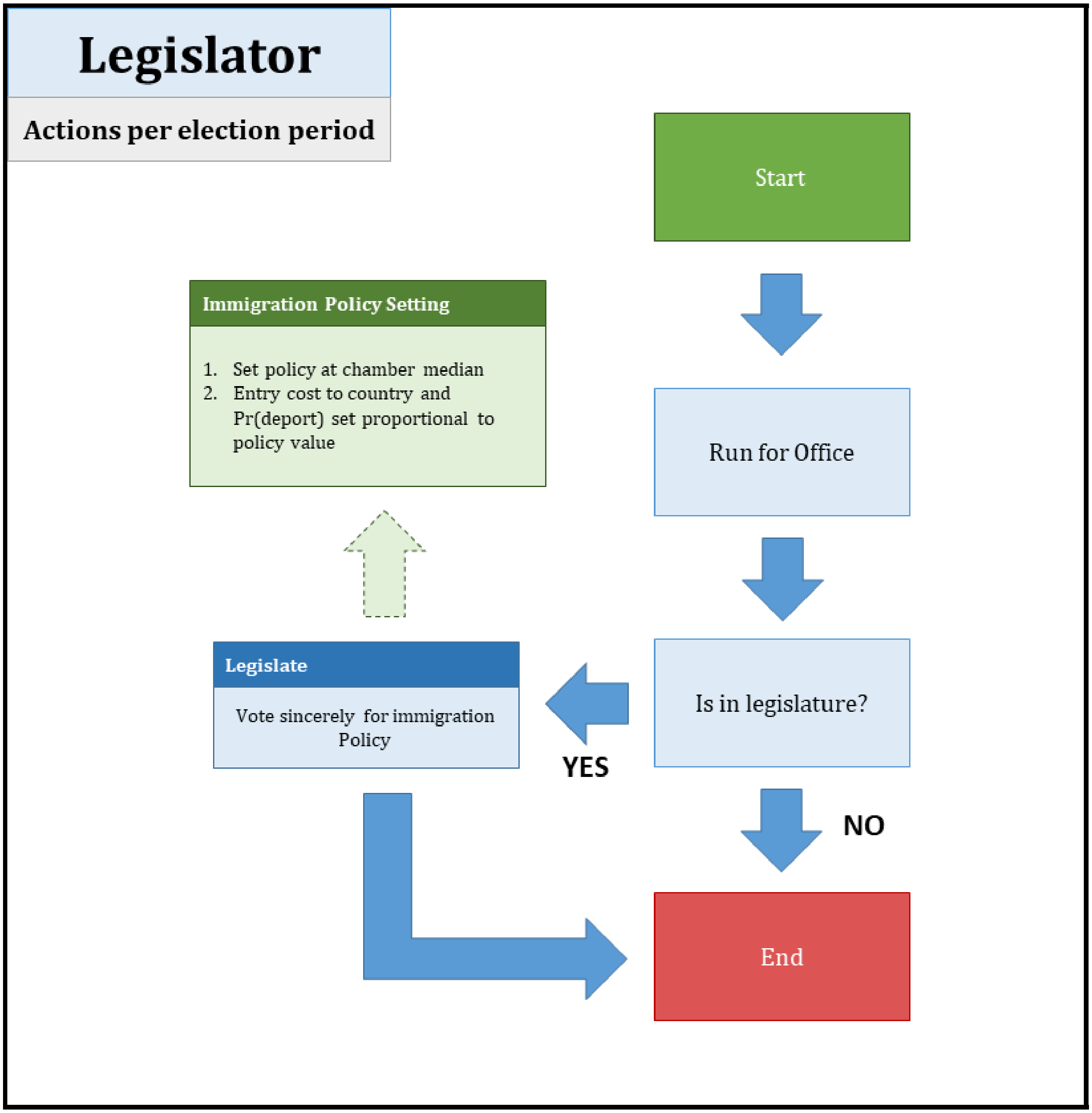


Figure 4.7: Legislator: Periodic Actions. Periodically (ticks > 1), citizens elect a new legislature in a single non-transferable vote (SNTV) election. After each election, the elected legislators set immigration policy as the chamber median of their ideal points. This policy is used to set migration related variables.

4.1.3 Jobs & Economies

Jobs have the following properties: product generated, wage and ideal education level. When generated randomly, jobs will show a correlation between wage and ideal education level, specified as an aggregate-agent parameter. A job that is filled generates output and pays its wage to the person filling it. In order to fill a vacant job, the job collects applications from among the local population. Job seekers only apply to jobs with ideal education levels within a certain threshold of their own education level, and once there are enough applicants, a winner is selected at random. Every tick, the employee is evaluated for job performance and can be fired with probability proportional to the amount by which their education level falls below the ideal education level of the job. In essence, any individual can theoretically fill any job, but they will not remain long in that position if their education falls below the requisite education level of the job.

Variable	**Type**	**Range**	**Description**
ID	String	-	Unique ID
WAGE	Double	0 - Inf	Wage
EDUCATION	Integer	0 - Inf	Ideal Education level
CITY	String	-	Unique ID of City Located in

Table 4.5: Job Characteristics

A collection of jobs comprises an economy, and standard economic measures are calculated using the jobs and employees associated with the economy. This includes metrics such as: GDP, (un)employment rate, expected wage, and mean wage. Additionally, each of these measures may be calculated for subpopulations, such as: (un)employment rate among natives, mean wage among immigrants, expected wage for persons with education level X, to name a few.

As with the choice of not having persons created or destroyed, this model of economy makes obvious and important omissions for the sake of analytical simplicity. Two in particular: there are no firms, and jobs may not be created or destroyed. The implications of these choices are discussed in chapter 8.

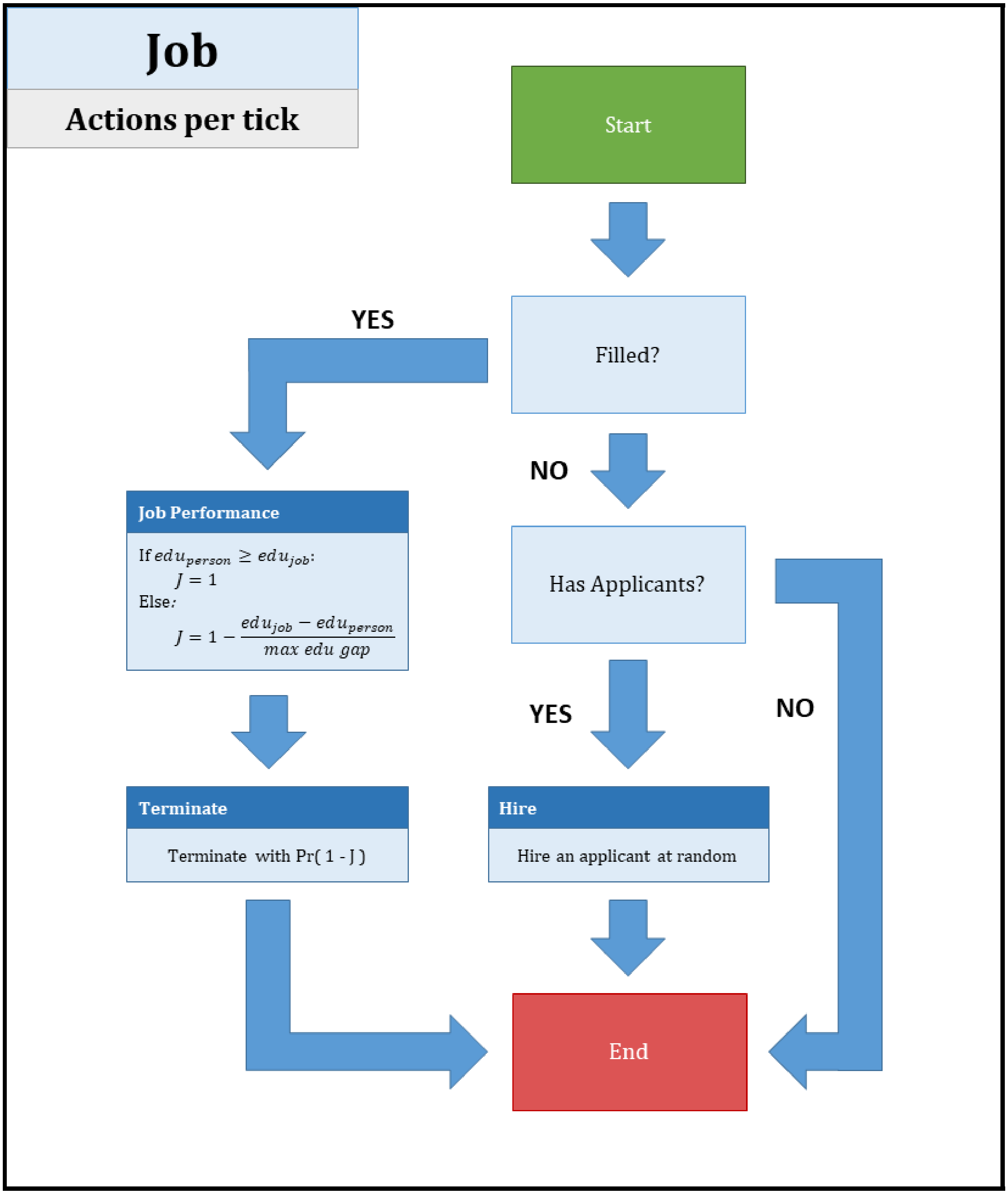


Figure 4.8: Job: Actions per Tick. Each tick, every job pays its employee if filled, and gathers applications / hires if not.

4.1.4 Cities & Countries

Cities are the basic locational unit in this ABM. Each city has a population, discrete location (i.e. latitute and longitude), number of legislative seats, cost of living, and local economy. The base cost of migrating from one city to the next is defined as $D * C_{travel} + C_{entry}$, where D is the Euclidean distance between the two cities, C_{travel} is the cost per unit distance traveled, and C_{entry} is the cost of entry if migration occurs into another country. In some versions of this ABM, this cost may be discounted if the migrant is plugged into a social network in the destination city.

Variable	**Type**	**Range**	**Description**
ID	String	-	Unique ID
CITY_NAME	String	-	Common Name
COUNTRY	String	-	Unique ID of Associated Country
LAT	Double	-90 - 90	Latitude
LONG	Double	-180 - 180	Longitude
POP_COUNT	Integer	0 - Inf	Population
JOB_COUNT	Integer	0 - Inf	Number of Jobs
COST_OF_LIVING	Double	0 - Inf	Mean Cost of Living
EDU_MVN_MU	Double	0 - Inf	Mean of Population's Education
EDU_MVN_SIG	Double	0 - Inf	SD of Population's Education
WAGE_MVN_MU	Double	0 - Inf	Mean of Population's Wage
WAGE_MVN_SIG	Double	0 - Inf	SD of Population's Wage
EDU_WAGE_MVN_RHO	Double	0 - Inf	Correlation of Pop. Edu & Wage
NUM_SEATS	Integer	0 - Inf	Number of Legislative Seats
NUM_LEGISLATORS	Integer	0 - Inf	Number of Politicians
IDEOLOGY_MU	Double	0 - Inf	Mean of Population Ideology
IDEOLOGY_SD	Double	0 - Inf	SD of Population Ideology

Table 4.6: City Characteristics

Each city is primarily an **aggregation of persons** and has a/an...
• associated economy • discrete location, enabling calculation of distance to other cities • cost of living which impacts residents' happiness and sustainment in the city • pool of legislators, in and out of office, and seats in the legislature

Table 4.7: City Summary

Countries are collections of cities and therefore derive several properties from their constituent cities. The country's economy is simply the agglomeration of all of its city's economies. Therefore, GDP is the sum of all the cities' GDPs, employment and unemployment rates are the weighted averages of the cities' corresponding figures, and so on. Likewise, the country's legislature is just the collection of legislators elected in all its constituent cities. The most important property of each country is its immigration policy, set by its legislature.

Variable	**Type**	**Range**	**Description**
ID	String	-	Unique ID
COUNTRY_NAME	String	-	Common Name
POLITICAL	Integer	0 - 100	Democracy Score
PR_DEPORT	Double	0 - 1	Probability of Deportation
ENTRY_COST	Double	0 - Inf	Cost to Enter Country
NATURALIZATION_TIME	Integer	0 - Inf	Number of ticks to naturalize

Table 4.8: Country Characteristics

Immigration policy is composed of the cost of entry, which is the amount of money a non-citizen must pay in order to migrate into the country, and probability of deportation, which is the probability that a non-citizen will be forcibly returned to their country of citizenship in any given tick. Cost of entry ranges between 0 and C_{entry}^{max} and probability of deportation ranges between 0 and $Pr(deport)_{max} < 1$, where the two upper thresholds are simulation

Each country is primarily an **aggregation of cities** and has a/an...
• associated economy, which is an aggregate of the economies of its component cities • legislature comprised of legislators from member cities • immigration policy comprised of... - probability of deportation: probability that a non-citizen will be involuntarily returned to their country of origin - cost of entry: amount added to a person's cost of moving when migrating in from another country

Table 4.9: Country Summary

parameters.

4.2 Model Limitations

The model presented above has, what I believe to be, the minimum components required to represent the political, economic, and social factors that comprise an IM system. I made the decision to work with a minimal model because this is an introductory study and the goal of validating the approach and the foundational model take precedence over building a model that is more comprehensive and therefore more difficult to work with.

One way to categorize these limitations is by (1) agent types that are not included, and (2) characteristics and mechanisms that are not modeled for the agent types that are included. The former includes things like: firms, state actors (e.g. border patrol, bureaucrats), and non-state actors (e.g. smugglers, international job recruiters, matchmaking agencies). For agent types that are included, characteristics that are of substantive interest but are not modeled include:

Person sex/gender, age, birth, death, marriage

Jobs industry type, job creation and destruction

Politicians more than one issue area, debut/retirement

At this time, I both acknowledge and accept these limitations for the sake of model tractability. Rectifying these shortcomings may have interesting consequences for the model and I discuss these in chapter 8.

CHAPTER 5

Scenario 1: Bilateral Economic Gradient

This chapter presents the results of running the MAPES model for the Bilateral Economic Gradient (BEG) scenario. As described in section 3.2, this scenario is comprised of two countries representing the classic economic migration dyad. Country A is the traditional receiving country, which has higher wages and more jobs, but a smaller initial endowment of labor. Country B is the traditional sending country which has a large initial labor pool but low wages and job availability. The parameters defining this scenario are summarized in table 5.1.

Both countries start each simulation with 100 persons whose education is distributed normally with means of 14 years for A and 10 for B, and standard deviations (SD) of 3 and 2.5, respectively. Ideology for all persons is distributed univariate normal with a mean of 50 ideology points and an SD of 15. Country A has 110 jobs and B has 70, and these are distributed multivariate normal in the fields of required education and provided compensation (i.e. wage) with means of 14 and 10 years for jobs, and 4000 and 1000 monetary units (MU) for wage, respectively. The corresponding SDs are 3, 2.5, 500 and 200. These wages are used to pay for the cost of living, which is 1600 MU in A and and 400 MU in B. Each country has 20 politicians (i.e. legislative candidates) who compete for 5 legislative seats. Their ideologies are all distributed normal with mean 50 ideology points, and SD 15. Some of these distributions of agent characteristics are graphically shown in figure 5.1. Finally, the countries are located 50 units apart. The unit for distance has no subtantive meaning but is used to calculate the cost of moving, which is 2 MU per unit distance (for all scenarios including those covered in subsequent chapters), or 100 MU per move for this scenario.[1]

City	City A	City B
Country	Receiving	Sending
Latitude	50	0
Longitude	0	0
Living Cost (Monthly)	1600	400
Population	100	100
Pop. Education Mean	14	10
Pop. Education Std. Dev.	3	2.5
Pop. Ideology Mean	50	50
Pop. Ideology Std. Dev.	15	15
# Jobs	110	70
Job Education Mean	14	10
Job Education Std. Dev.	3	2.5
Wage Mean	4000	1000
Wage Std. Dev.	500	200
Corr. of Edu. & Wage	0.8	0.8
# Leg. Seats	5	5
# Leg. Candidates	20	20
Leg. Ideology Mean	50	50
Leg. Ideology Std. Dev.	15	15

Table 5.1: BEG: Locale Parameters

This is clearly the least complicated of the three scenarios investigated in this dissertation, so was the most likely to satisfy theoretical expectations if MAPES performed as expected, but also the most catastrophic if it failed. Thankfully, this litmus test did not disappoint.

[1]Cost of moving was modeled in this slightly roundabout fashion to ease parameterization for scenarios with more locations. To put it another way, it is easier to have a constant cost per unit of movement and record coordinate locations than to record differing costs of movement between multiple dyads.

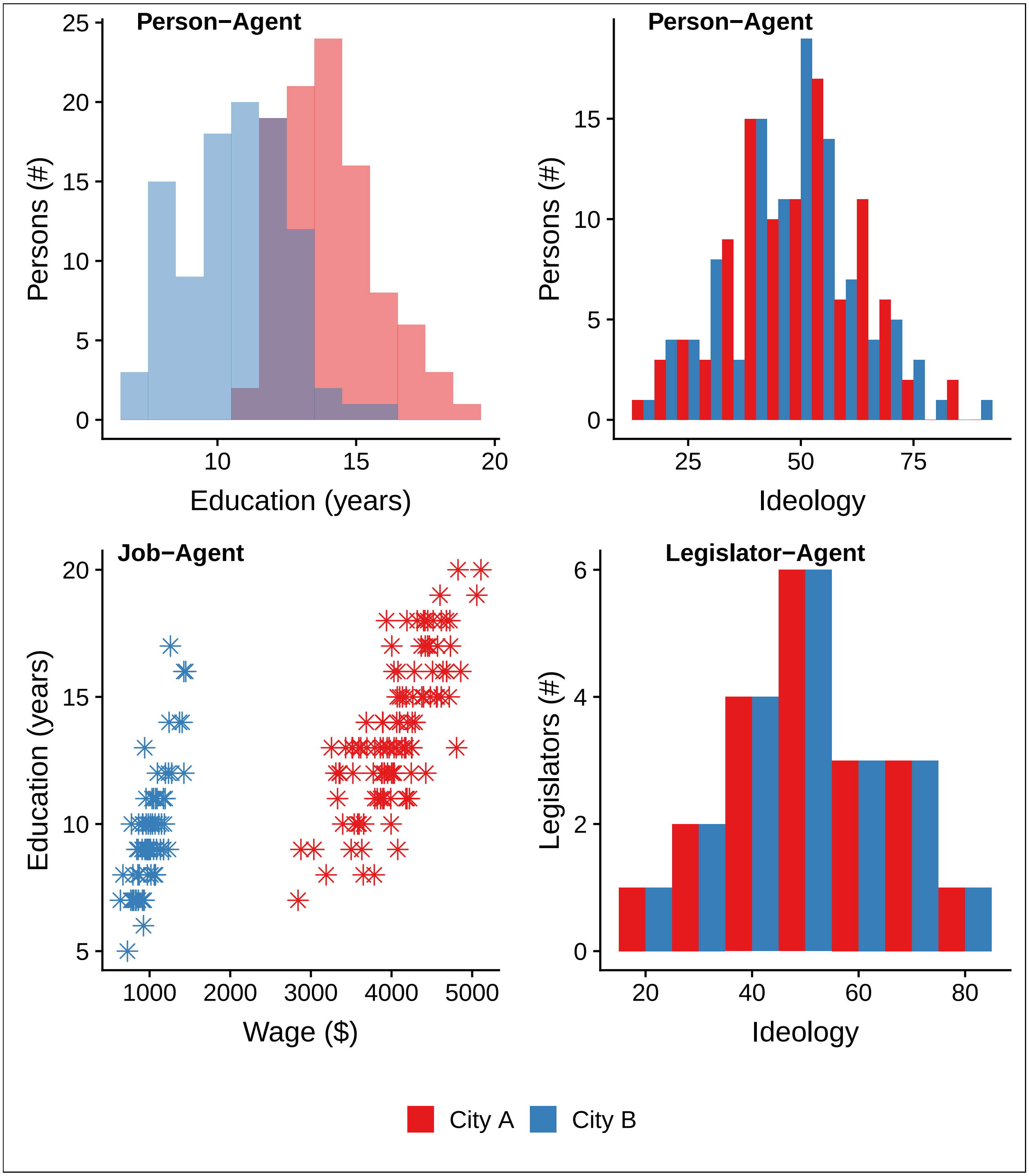


Figure 5.1: BEG Agent Characteristics

Bottom-line up front, net migration flows do indeed occur from the traditional sending country to the receiving. There is also some migration from the receiving to the sending country, resulting in the Brownian type motion we expect to see given microeonomic-type theories, but the net is still in the expected direction. And, as predicted, there is political conservatization in the receiving country in response to the influx of migrants. These results and more follow, organized as discussions on (1) general population trends; (2) politics; (3) economics; and (4) the experience of individual migrants.

5.1 Population

Figure 5.2 shows the populations of the two countries in BEG, A and B, over the lifetime of the simulation (100 ticks). For the BEG scenario, unless specifically noted, city and country A, and city and country B are referred to interchageably since each country only contains one city.

Each of the plots in figure 5.2 shows two trendlines: number of persons; and number of natives (persons originating from that country). The shaded areas (ribbons) around the central lines show the full range of responses observed across 10 runs of the simulation.[2] The gap between the two trendlines is the number of foreigners in the country. Foremost, as we would expect from macroeconomic theory, there is a net flow of persons from country B (labor surplus & low expected wage) to country A (labor shortage & high expected wage). We can see that there is a net flow because the line for number of persons increases to a point above the starting value (100) in A, and decreases below that for B.

Given that we know how person-agents work, this net flow must somehow be attributable to differences in expected utility between the two countries. However, this gap does not appear to derive from a difference in average expected wage. Figure 5.3 shows the difference

[2] That MAPES is fundamentally a simulation means that examining individual agents is slightly more complicated than working with real world data. The main complication results from the fact that multiple simulations must be run for a given set of input parameters in order to estimate confidence intervals. So every individual actually ends up having a dataset for each run in the run-set. Analysis / figures of individual agents or groups of individuals, therefore, must take the full set of runs into account.

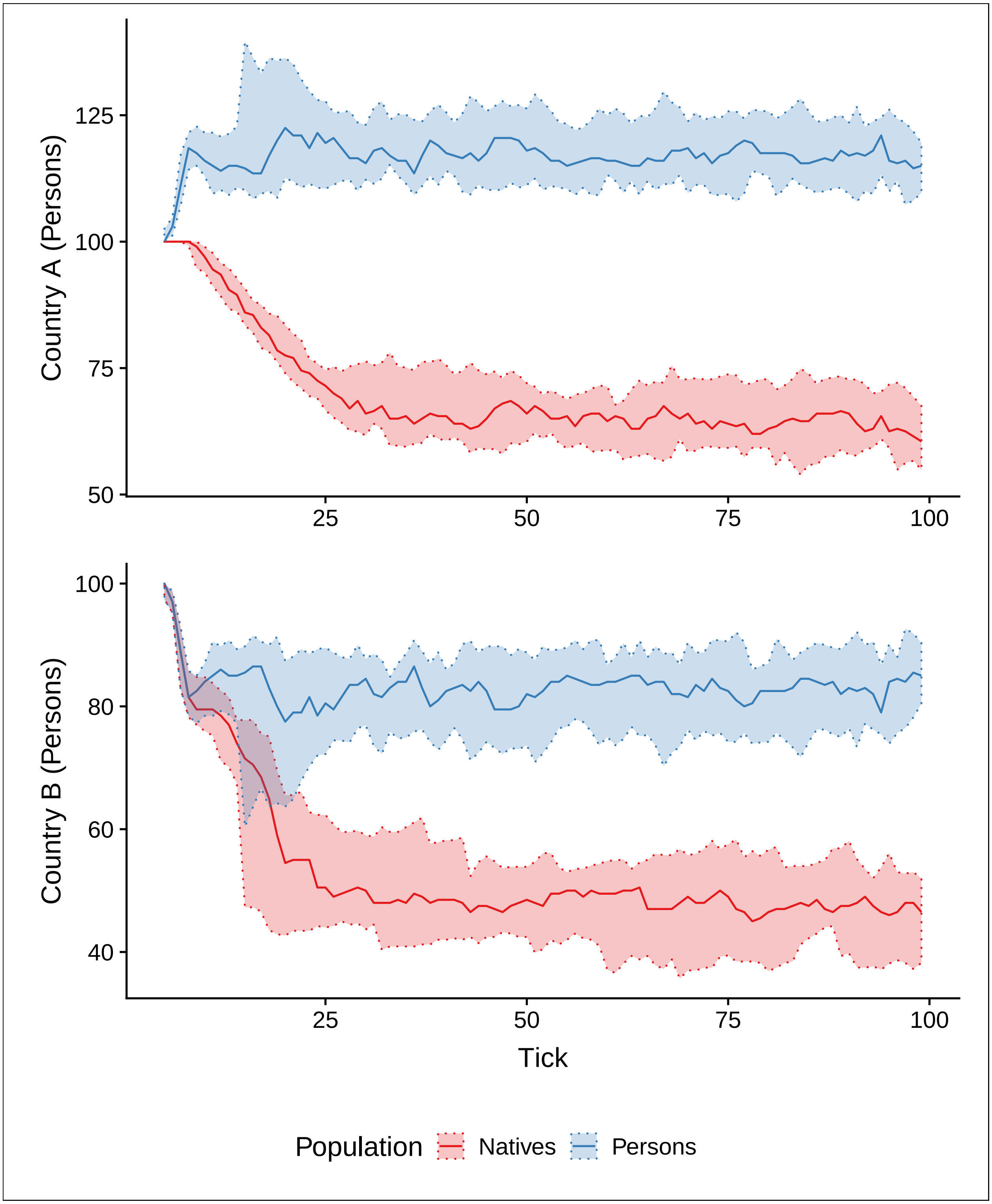


Figure 5.2: BEG: Populations over time

in expected wage between the two countries across all runs. Expected wage is calculated per tick as the mean wage of all unoccupied jobs in a particular locale.

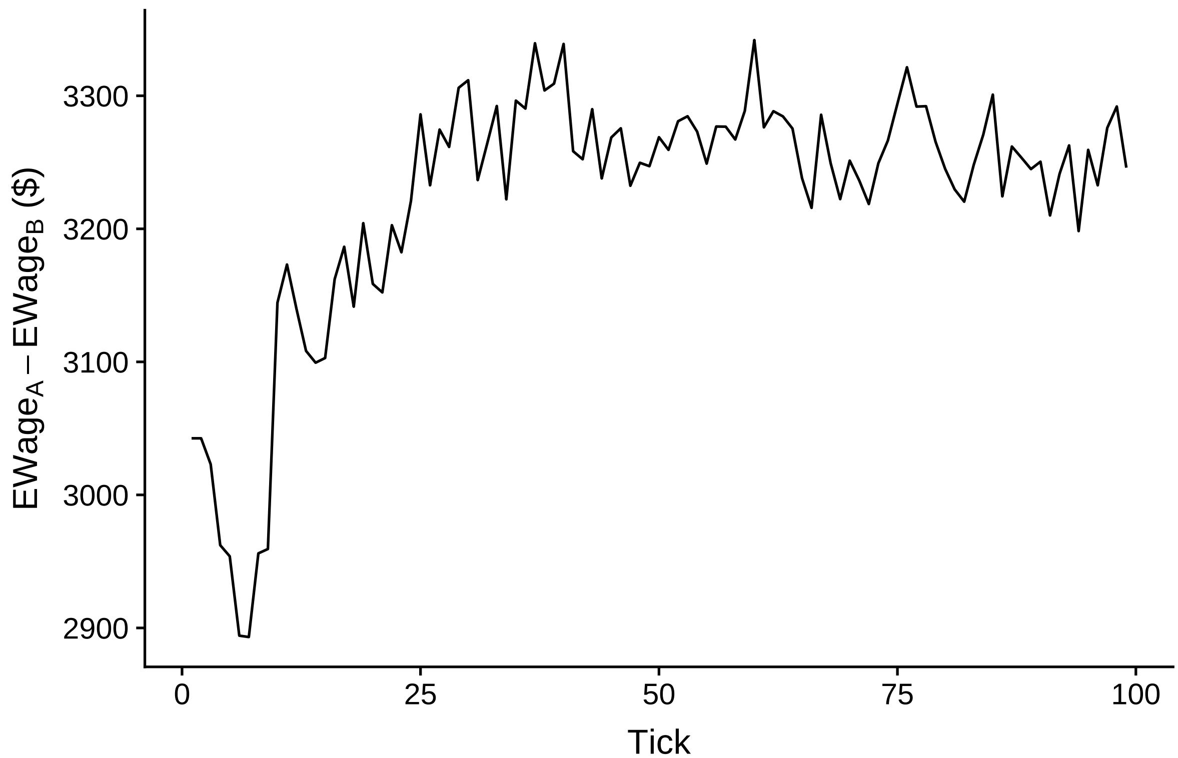


Figure 5.3: BEG: Comparative Expected Wage

Contrary to expectations, the gap actually increases over time instead of decreasing! We know *why* this is the case: as further elaborated in section 5.3, country A suffers an increasing shortage of labor for its highest education jobs. Therefore, since there is a correlation between high education requirements and high wages for jobs, the average wage of unfilled jobs in A gets pulled upward, increasing the gap. The real question then is what motivates the flow pattern we see. If the story were entirely about income gradients, we should not see the plateauing pattern we see in figure 5.2. Answering this question is one of the primary goals of the remainder of this chapter.

Continuing, the median population of country A increases to approximately 117 and maintains within 4 persons of that level for the duration of the simulation.[3] This makes

[3]By default, point figures refer to bootstrap medians. Minima, maxima, and other points within a

perfect sense since 110 is the total number of jobs available in country A. That the level fluctuates slightly above that value, rather than maintaining at precisely that level, is because of the time lag between the decision making process of individuals and the job hiring process. This requires some explanation.

Population size rises above the total number of jobs when more people enter than there are jobs available. This occurs because each individual bases his migration decision only on the expected wage in the destination at time of departure. They don't know how many natives they will be competing against, how many of their compatriots are also moving based on the same information, or whether they will actually be hired upon arrival. Hence the upward gap. Population size can fall below number of jobs for a variety of reasons but in this context, the main culprit is that jobs can't be filled immediately upon being vacated. At the very least, potential candidates must be made aware of the opening and the the hiring process must occur, resulting in short but unavoidable vacancies.

The converse of country A's population increase is that country B experiences a population decline from its starting point at 100 to approximately 82 – the total world population (200) minus the steady-state population of A (117). As with country A, there is noise around the steady-state population size, but the main culprit here is the population / labor market situation in A, rather than conditions local to B. We know this is the case because of the information presented in figure 5.4.

The first panel of figure 5.4 shows new hires per tick; a value of X% on tick Y for country Z means that of all the jobs present in Z on tick Y, X% hired a new employee to fill the position. The second panel does the same for firings. The final panel shows the ratio of the median (across simulations) number of firings to hirings. In essence, the third panel is an inverse measure of how many persons left their jobs voluntarily to seek better opportunities; lower values mean more voluntary departures. So the third panel shows us that persons in B tend overwhelmingly to leave only when there are better prospects in A. In other words, B is basically a bench where labor waits to enter the game in A. This renders the labor market

particular CI will be labeled so explicitly.

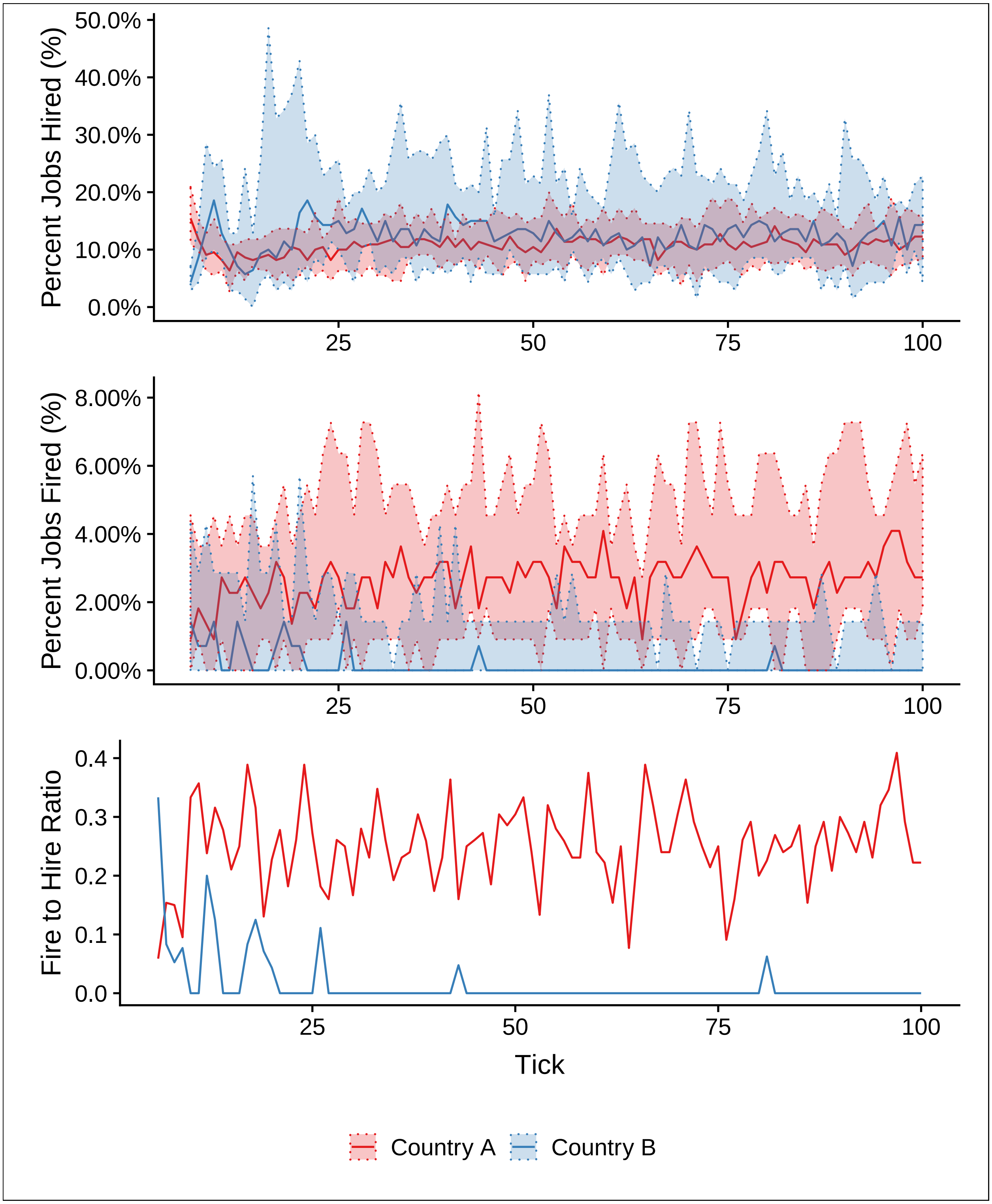


Figure 5.4: BEG: Job Churn

in B largely subservient to what is going on in A, outside of its control.

The next noteworthy observation may be made from figure 5.5, which shows the percentage of each country's population that is immigrant. As we expect, immigration into country A starts immediately, but what is surprising is that movement from A back into B starts soon thereafter and at a similar scale. In both cases, the growth of the immigrant population is roughly logarithmic, with the majority of growth having occurred by tick 25. The steady state immigrant percentage for country A is around 44, and that for B is approximately 43; not much less. The difference in immigrant percentage between the two countries is shown as the gray line in figure 5.5 (using the secondary axis). Clearly, the gap in immigrant composition starts high but appears to decay exponentially. While B's percentage never exceeds that of A, by simulation's end, the confidence intervals basically overlap so that they have similar immigrant compositions.

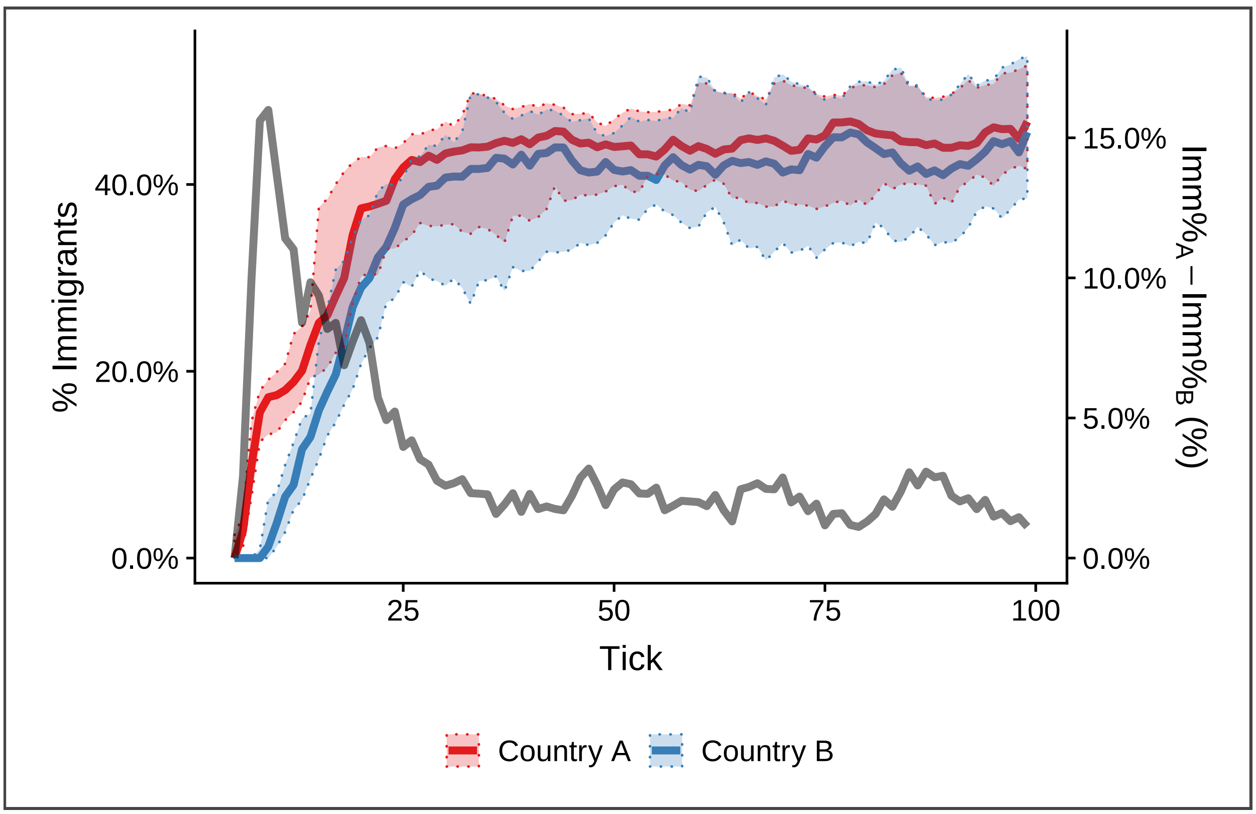


Figure 5.5: BEG: Immigrant populations over time

Substantively, what the flow into B represents is low-utility (i.e. unhappy) persons in country A eyeing, competing for, and sometimes taking opportunities in B. To put it another way, even though net flow is unidirectional, from country B to A, individuals move in both directions and labor conditions in both countries are affected. The end result is that in BEG, immigrant populations are *not* stable, changing in a Brownian fashion throughout the duration of the simulation. What is unclear at this point is who exactly is moving. It is obvious that the persons moving must have low utility but it is unclear why they do, and whether they come from a specific subpopulation or are broadly distributed across the entire population. These questions are addressed in section 5.4.

5.2 Politics

In this section, I consider some of the political characteristics of BEG and the overall MAPES model. Specifically, I look at the composition of the legislature over time, and the changing preferences of individual voters. In the process of doing so, I show how a seemingly small detail of MAPES mechanics actually has huge consequences for system-level political outcomes. The issues brought to light in this section actually inform updates to model mechanics that are reflected in scenarios covered in chapters 6 and 7.

As discussed in chapter 4, both normal persons and legislators have ideological ideal points with respect to the issue of immigration. Their respective distributions of preferences is shown in figure 5.1. Even though elected legislators represent the preferences of voters, they must be examined separately because different ideological patterns at the individual level can be observationally similar at the legislative level. Additionally, examining each in turn allows us to see how well voters' preferences are represented or distorted by the electoral processes modeled.

Figure 5.6 shows the median ideology of the legislatures (i.e. chamber median) of countries A and B for the lifetime of the simulation. The legislatures of both countries become more restrictive over time, with median ideology for A increasing from an initial value of 53 to 66, and 48 to 70 for B. For both countries, the path of this increase appears to basically follow the same three stage process. In the first stage, the countries peak out at 66, a little over a third of the way into the simulation with A reaching the peak shortly before B. In the second stage, both legislatures enter a brief steady state maintaining the aforementioned peak value until about halfway through the simulation. Finally, both countries leave this steady state to enter an area of volatility centered around chamber medians somewhat lower than the peak value reached in the previous stage, respectively 66 and 70.

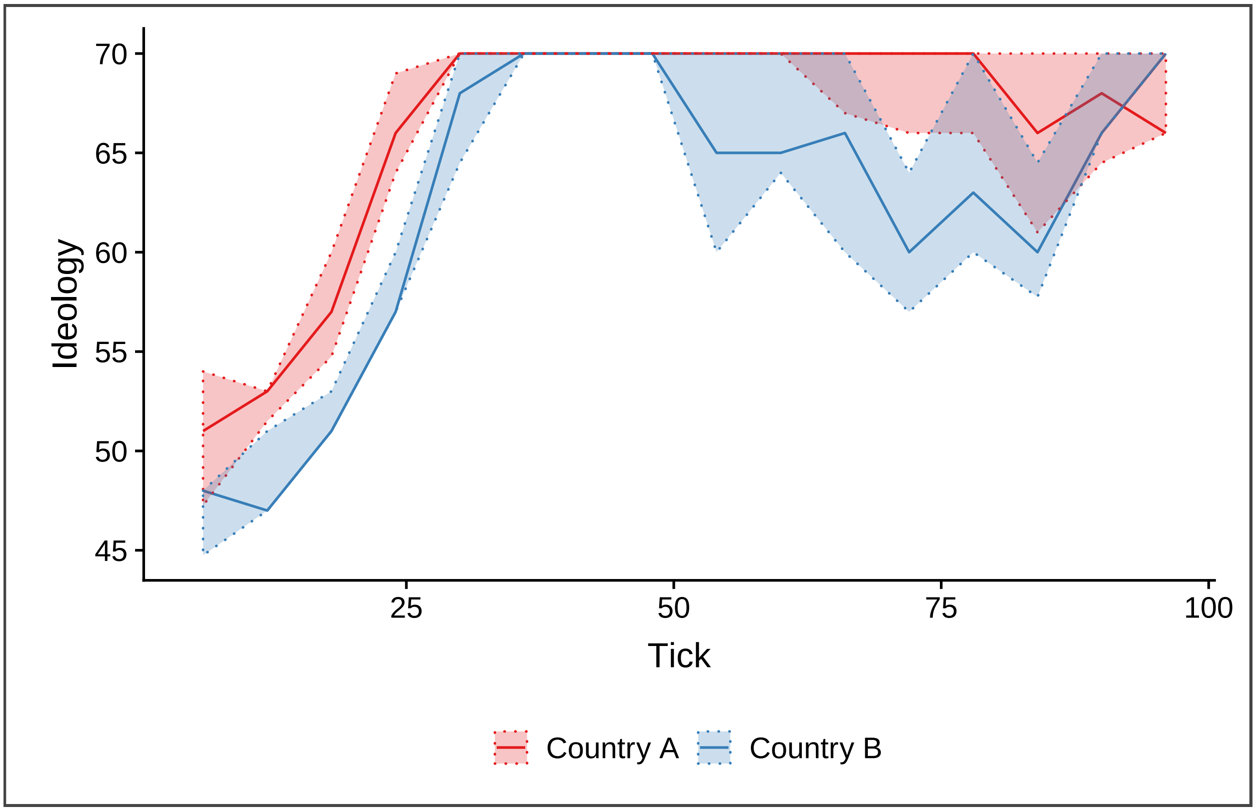


Figure 5.6: BEG: Legislature Median Ideology Over Time

There are at least three interesting aspects of figure 5.6. The first is that it basically mirrors figure 5.5 in the first half of the simulation, which is to say that the legislature's sentiments about immigration appear directly related to the proportion of immigrants comprising the population. This makes sense given the way MAPES models ideology and changes thereof. Recall that a native's ideology is updated based on his interactions with foreigners, which probabilistically go well or poorly, with the former leading to more open preferences and the latter to more closed. So, of course, having more immigrants will have an impact on voters' perception of them since there will just be more foreigner-native interactions. Larger numbers of immigrants will also affect natives' ideology through their social networks since there is the mechanism of each person's ideology regressing ever closer to the average of their social network. A native with an immigrant in their social network will feel a liberalizing effect since the immigrant will most likely have a preference for more open policies, and fellow natives whose attitudes have been colored by foreigner interactions will also pull them in the direction of their own preferences. During model construction, it was not clear at all what the net effect of these competing factors should be, and I will show in chapter 6 that the mechanisms through which these factors are felt are more complex than they seem. Regardless, from figure 5.6, we are seeing indications that, at least for the parameters used for BEG, the total amount of negative sentiment stemming from negative foreigner-native interactions and the proliferation of negative sentiments through social networks outweighs the liberalizing sentiment generated by positive interactions and their proliferation.

The second point of interest about figure 5.6 concerns the relative timings of the different stages of the graphs. To be clear, it is not terribly important whether ideology shifts to a steady state in tick 20 or 30; the absolute value is ultimately just a function of model parameters. What is of substantive import is the timing of certain events relative to others; why does B mirror A with just a few ticks lag, and why does A transition more slowly out of steady state than B? As mentioned above, the first question is really a question about the timing of immigrant populations. Whether the same holds true for the second depends on how well one believes figure 5.6 to mirror figure 5.5. Based on just a visual inspection, it is not completely clear that legislative medians are following immigrant proportions as closely in the second half of the simulation as in the first. If that is the case, then what causes this transition and what becomes the driving force? A potential contender for the latter is the population of returnees – persons who leave their native countries but return at a later date. This possibility is thoroughly investigated in section 6.2.

A final point worth considering is the fact that neither country's legislature has any variation during the steady-state (middle) stage, as seen in the lack of an interval around the point estimates in this period. This stands in stark contrast to the other stages of figure 5.6 where there are always intervals, however narrow. From a purely technical standpoint, this suggests that the peak value in the steady state is actually right censored (orienting higher values to the right). Speaking substantively, this means that the available pool of politicians doesn't have candidates conservative enough to adequately represent the rightward preferences of the electorate. From a different perspective, this means the legislature is having a moderating effect on the most extreme inclinations of the electorate. Whichever interpretation one subscribes to, it's also important to realize that this phenomenon arises in its present form only because the pool of politicians is static in MAPES. In the real world, one could imagine either new politicians emerging or existing ones shifting their own preferences, just as the electors do.

We can see whether and how much the actual electorate departs from its representatives by directly examining their preferences. Unlike the legislators, the ideological preferences of normal persons change over time, so it is useful to view individual trajectories, not just range intervals around group medians. Additionally, we can also examine how many votes are actually garnered by each individual legislator. In combination, these two views of the electorate will allow us to see whether any multi-modality emerges over time. This is important because the legislature's tendency towards restrictive policies we see in figure 5.6 can reflect at least two different situations in the underlying constituency. The first is that the population as a whole shifts rightward. The other is that some subset of the voting population moves rightward while the remainder remains fixed, or liberalizes to a lesser extent.

Figure 5.7 shows the actual vote counts for each candidate over time. Each dot is the vote count for a particular candidate (which one does not matter) and the colors represent different elections, by tick. For both countries, the distribution of votes received across the ideological spectrum is basically unimodal (i.e. has one hump) and shifts rightward over time without the emergence of any other modes. This indicates a general populational shift. The voters in this simulation do not exhibit significant ideological clustering, which would be indicated by the emergence of multiple modes i.e. several humps, instead of the single hump per country-plot. By clustering, I mean a grouping of individuals with similar, potentially identical ideology values. On the surface, this seems surprising; would it not

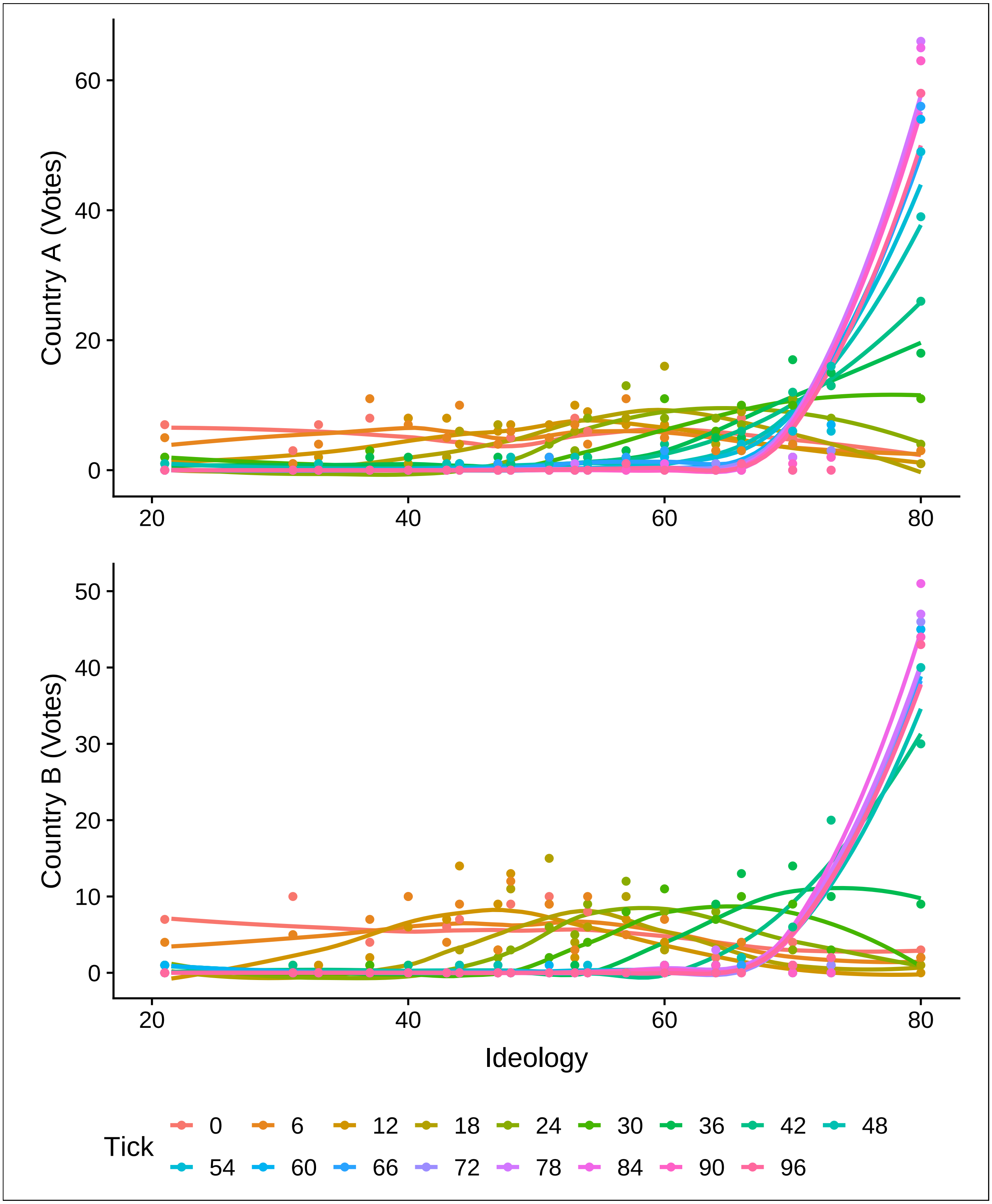


Figure 5.7: BEG: Legislator Vote Share by Ideology Over Time

be more reasonable to see the emergence of multiple modes analogous to the formation of political parties? The short answer is no, but to see why, we need to understand what a cluster is from the perspective of MAPES model mechanics.

Toward this end, figure 5.8 shows the progression of ideology for all individuals by current city, across several of the simulation runs. Each thin line represents an individual in the simulation and his ideology for the duration of the simulation. The thick lines are loess fits across all individuals in each city. As led to expect by the legislature plots, the ideological preferences of the underlying populations tend to shift right (restrictive). Lines which change color mid-simulation indicate moves from one city to the other.

There are two types of clustering occurring in figure 5.8. At the macro-level, it is evident that each country has a single large cluster. Once one zooms in, there are smaller mezzo-level clusters as well. An obvious example of this is the small tendrils (i.e. groups of persons) significantly diverging from the main trunk for each country. Even beyond these obvious examples, however, nearly all individuals do eventually end up in a cluster, which is basically their social network. This is seen in the quantization of ideology values: as one progresses forward in time, what were individual lines at the outset converge into groups of lines such that each "line" towards the end of the simulation is actually several persons who have converged through the mechanism of persons regressing toward the mean ideology of their social networks.

Why does this clustering not progress further, leading to (1) more convergence between the mezzo level clusters and (2) the emergence of what we would think of as parties (i.e. macro-level clusters of mezzo clusters)? The answers to these two questions lies in the minutiae of the MAPES model. Mezzo clusters do not converge more due to *max_num_nodes*, the simulation parameter limiting the maximum number of social ties a person may have; beyond this number, no individual may add any more ties. Once individuals' social networks become saturated, their networks will reach a steady state in ideology that may only be perturbed by shocks that (a) change the average ideology of the network through the entry of new or exit of current members, or (b) exogenously change the ideology of current members. In other words, mezzo clusters are limited in size and fixed in place by *max_num_nodes* so

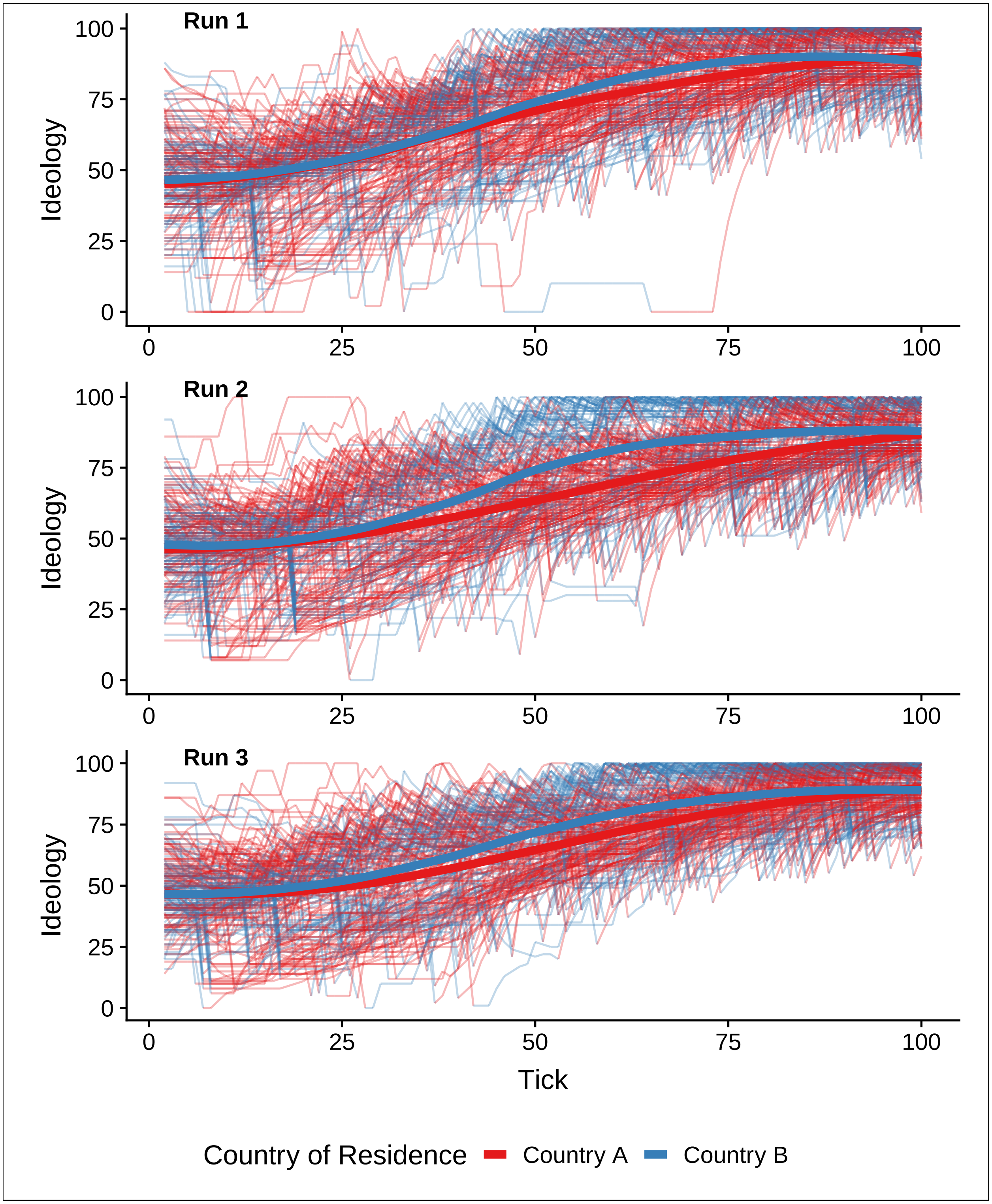


Figure 5.8: BEG: Population ideology over time

they cannot converge more beyond a certain point. Is this model feature realistic, and if not, how may it be improved? It is realistic in the sense that nobody knows everybody in the real world. Without the hard upper limit, every individual in the model would asymptotically gain a network tie to every other. This would then allow ideology to converge to a single, hivemind like entity; clearly unrealistic. On the other hand, this parameter presents a modeling quandary. Not having it is clearly problematic but having it also inhibits the formation of clusters larger than *max_num_nodes*, so whether a "party" forms or not becomes a parameter tuning exercise rather than something that organically emerges. Another mechanism is required to achieve the desired effect in clustering.

Another reason why no parties form is due to the way that social ties are formed in MAPES. A cluster is a group of persons who have similar ideology values either through simple chance, or (of more relevance) because they share more social ties between each other than they do with others. Therefore if there *were* distinct macro clusters, it would mean that there are population-level groups not just separated in ideology but also in network connections as well. In order for such groups to arise, there must exist some discriminating mechanism that leads to the rise of separated networks. However, in the default BEG, there is no such mechanism because network connections are formable between all nodes with equal probability, with the exception of between natives and non-natives, who have lower but still significant probability of forming ties. The lack of a discriminator means that MAPES allows persons from extreme ends of the ideological spectrum to form ties with each other, then moderate each other. In my opinion, this is unrealistic.

Another way we know that the lack of a social-tie discriminator is a primary cause of the single mass is that across all the plots in figure 5.8, the ideological spectrum for individuals begins wide but narrows dramatically very early in each run. Close examination of individual trajectories shows that the extreme ideologies that would probably lead to separate clusters in the real world here form (i.e. are canceled out by) improbable liaisons with moderates or extremists on the other end of the spectrum. This means that completely random social network formation causes everyone to regress to the mean ideology. To make MAPES more realistic, the mechanism for forming social ties must weight potential partners in a non-uniform way (e.g. on the basis of ideology), which while costly from an implementation point of view, is not conceptually difficult. Such an enhancement is made for model runs of the remaining two scenarios of this dissertation, and is detailed in section 6.2.

Yet another question raised by figure 5.8 is why many people exhibit a high amplitude, (sometimes) high frequency fluctuation in ideology. This question can be answered purely through introspection about model mechanics as well. Recall that a person's ideology change occurs due to three model events. First, persons who migrate experience a leftward (liberalising) ideological shift, of a magnitude defined as a scenario parameter (30 ideological points, see table 3.1). Second, natives who experience a negative interaction with a foreigner are given a similar rightward shift (-10 points). These account for the movements out to the extremes. These are rapidly followed by movements back toward the mean as a result of the third model event, which is the proliferation of ideology through social networks. Obviously, this dynamic does not seem realistic – real people probably do not experience such rapid swings in ideology. Does the artificiality of the underlying mechanics undermine our faith in the behavior of the overall system? How can the underlying mechanics be improved?

It depends on what we think we are looking at. On the one hand, our credulity is strained if asked to believe that each discrete data point on these charts represents the respondent's reasoned and all encompassing opinion at that instant time. Surely, real persons are not so fickle. On the other hand, if we think about these charts in terms of periodic averages or as integrals (i.e. what is a person's average ideology over periods of 5, 10, etc. ticks), the plots make more sense. For example, a person may start at an ideology of 50, and have event A in tick 10 cause a shift of -20, event B in tick 12 a shift of +10 and event C in tick 13, +8. Does that necessarily mean that the person truly went from centrist to very liberal (30) and back (48) in the short span of 3 ticks? Probably not. But it seems reasonable to imagine a person experiencing events, emotions, etc. of sufficient import to cause an *average*, slight liberalisation of -2 points over the same period. This is a much more reasonable proposition, which obviates the need for "fixing" the mechanics. Like the points preceding it, this shows but one of the complexities involved in interpreting MAPES results.

5.3 Economics

MAPES provides perfect knowledge about persons and jobs, so we can easily enumerate in-depth information about any aspect of the economy that is built into the model. For example, finding an arbitrary statistic such as unemployment figures among 1st and 5th time migrants to country A with education less than 12 is a simple one line query. This means we can easily determine whether there are nationality or migrant status based patterns in employment. In this section, I use this high degree of insight to answer two high level questions about the economics of BEG: (1) How does migration alter the competitive landscape for jobs? (2) How are people's economic outcomes altered due to migration?

It is worth noting that while we have full access to any economic aspect of the ABM that is actually modeled, obviously we cannot speak to things that are not modeled. This means there are myriad questions that MAPES in its current state simply cannot address. For example, we do not not know whether particular industries are impacted differently from others (since industries are not modeled). We also do not know how migration interacts with job creation (since jobs are static). Nor can we determine the impact on output-based aggregate figures like GDP (because jobs associate wage with education but do not have associated output values). We proceed to examine the economics of BEG keeping these limitations in mind.

So how does migration alter the job marketplace? One place to start is by examining unemployment figures. Figure 5.9 shows the employment for cities A and B over the duration of the simulation. For both cities, there is an initial state that lasts until around tick 40, after which employment numbers enter a steady state. For city A, which has 110 jobs, the number of natives employed falls from a median of 85 to around 55 in steady state. The jobs lost by A natives is taken up by immigrants from B, who start at 0 (since BEG starts with no migrants in either city), and stabilize around 36.

In the case of city B, of 70 total jobs, the number of natives employed starts at around 63 and stabilizes around 42). Immigrants from A occupy around 14 jobs on average. Interestingly, neither city ever saturates all its jobs; there seems to be a roughly constant structural

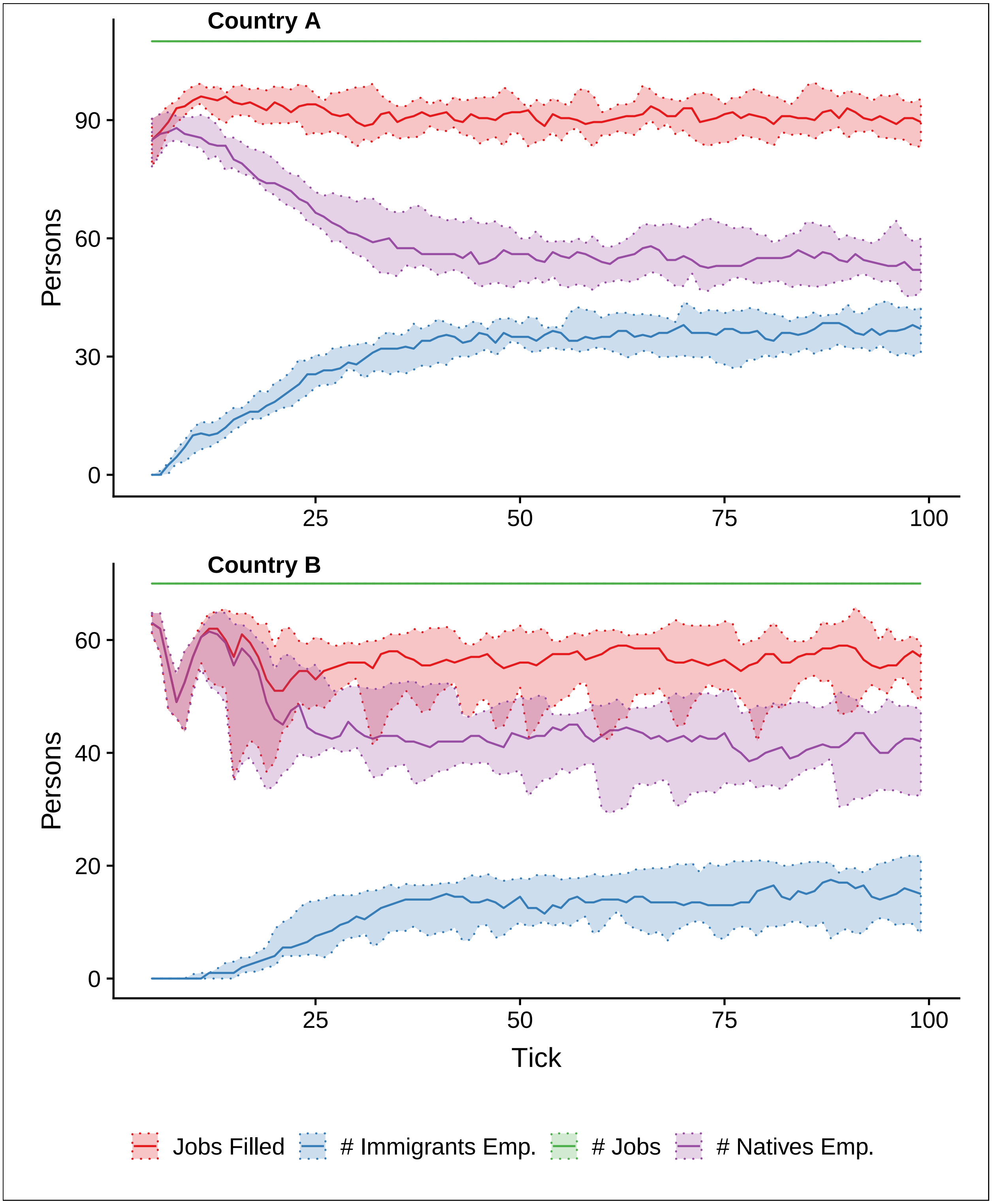


Figure 5.9: BEG: Employment

number of job openings at all times. It is unclear whether this arises from particular jobs always being unoccupied more than others (perhaps because there aren't enough skilled workers) or just the lag built into the hiring cycle.

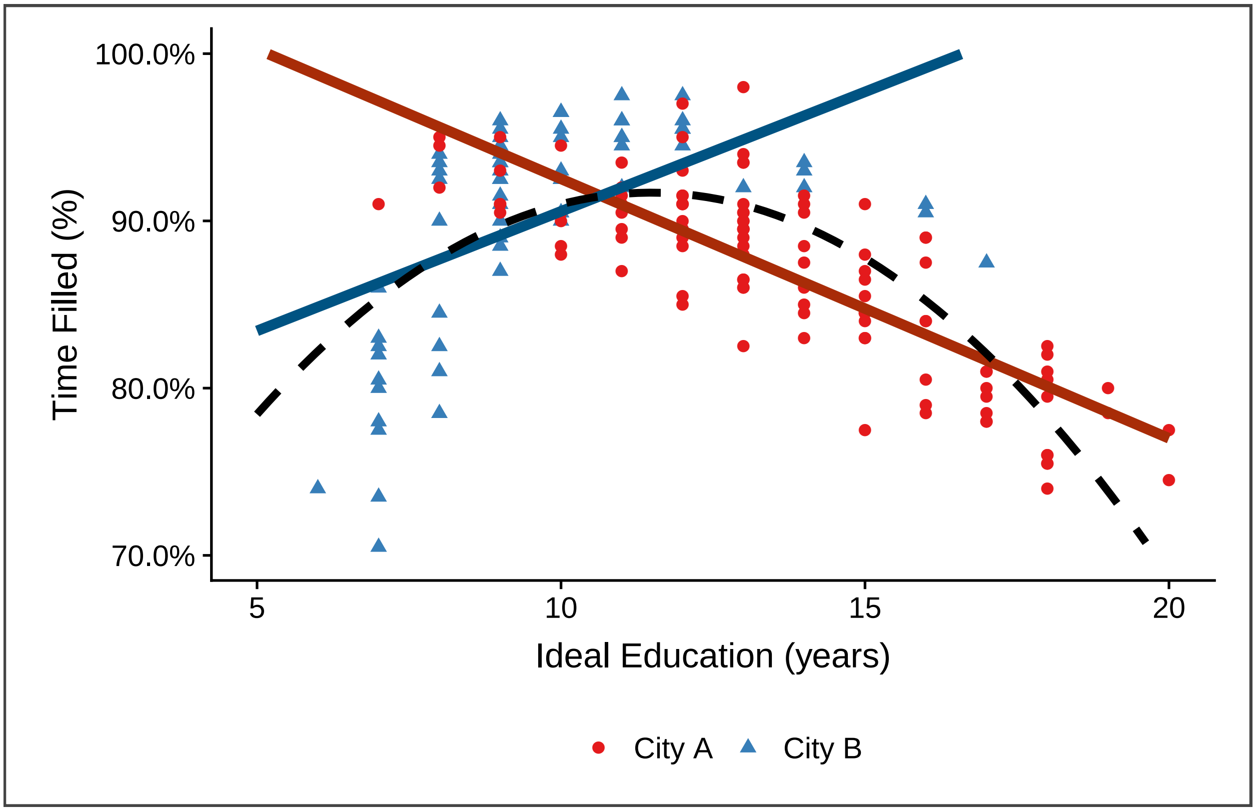


Figure 5.10: BEG: Percent Job Filled by Ideal Education

The numbers used to generate plots 5.9 - 5.10 provide an excellent opportunity to demonstrate how MAPES may be used to test and test itself against existing theories of migration. In this case, the theory that can be tested is the Harris-Todaro macro-economic theory of migration, which in its basic form provides a basic relationship between wages, workers and movement between two locations (refer to section 2.1). There are numerous refinements and complications which can be layered onto the canonical model but the basic relationship is that migration occurs so long as the following relationship holds:

$$\frac{w_L^C L^C}{N - L^R} > w_L^R \tag{5.1}$$

where w_L^C is the expected wage in the receiving location; w_L^R is the expected wage in the sending location; L^C is the number of jobs available in the receiving location; L^R is the number of employed persons in the sending location; and N is the total size of the workforce across the two locations.[4] Migration stops when:

$$\frac{w_L^C L^C}{N - L^R} = w_L^R \tag{5.2}$$

All of the components of equation 5.1 are directly measured in MAPES or calculable from data therein so we are able to directly test this proposition. Specifically, we should examine the relationship between the ratio of the LHS to the RHS of 5.1 on the one hand, and the timing of flows as shown in figure 5.2 on the other. If the Harris-Todaro model is theoretically able to interoperate with the other theories incorporated in MAPES and they have all been implemented correctly, in BEG we should see the LHS-to-RHS ratio (H/T ratio hereon) decline towards a threshold value (to be specified momentarily), with net movement occurring during the decline, and stopping once the threshold value is met. In the pure Harris-Todaro model, that threshold value would be one (by definition) but given that the individuals in BEG/MAPES base their decisions on factors other than just wages, that will likely not be the case, even if the overall functional form renders as expected.

Figure 5.11 plots the H/T ratio over time, alongside the number of immigrants in City A (the receiving location) for context. Contrary to theoretical expectations, the H-T ratio does not decline over time but instead, increases towards an asymptote at around 1.26, reached around the time marked as "Inflection Tick". The value of the asymptote is not important since, as stated earlier, MAPES clearly has more moving parts than the base Harris-Todaro model, so constants are liable to change. What is important is that an increase in the H/T ratio means that the expected wage in the receiving location is increasing, even as immigration decreases (indicated by the plateauing number of immigrants in city A). Astute readers will observe that this finding should not be a surprise at this point since we noted this finding (in passing at least) in section 5.1. Nonetheless, the cause of the finding is a mystery: why does the number of persons moving decrease even as their wage incentives to

[4]Zenou (2006, p.3)

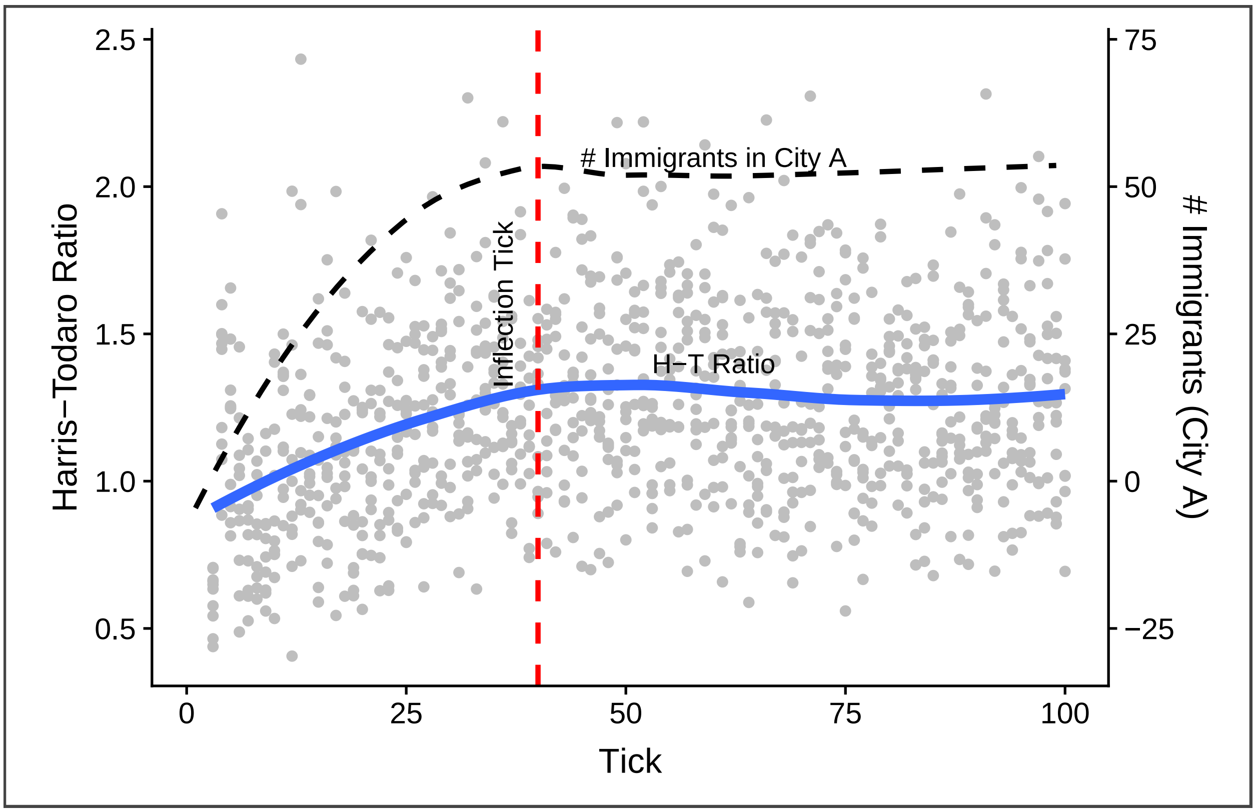


Figure 5.11: BEG: Harris-Todaro Equilibrium

move become stronger?

There are several possible explanations for this unexpected behavior. For one thing, the base Harris-Todaro model uses a uniform, exogenously defined wage for both the sending and receiving locations, so that the expected wage is only a function of employment probability, and not the composition of jobs that are available. In BEG, wages are distributed multivariate normal as a function of education, and the number of persons decreases with education level (across both locations). So, if there is some mechanism that causes the distribution of job openings to skew right on the axis of wage and required education, it is very possible for expected wages in open jobs to go up even as net movement slows, if there simply aren't enough qualified bodies to fill all the slots. Another important deviation is that the base Harris-Todaro model assumes constant, full employment in the sending location (Zenou (2006, p.2)), while the sending location in BEG (City B) has its own dynamic job market and permits migration from the receiving to the sending location to fill that market, which the basic H/T model does not. It seems likely that allowing bilateral movement permits more optimal market outcomes (since there is more freedom for agents to achieve optimal fit to employment conditions) but it may also break fundamental assumptions of the basic H/T model along the lines of the preceding hypothesis.

To be clear, I am not suggesting some great discovery or refinement of the classic theory. Since its original formulation, there have accumulated almost 5 decades of refinements and modifications for specific scenarios among which that in MAPES may already be covered. Rather, the pertinent point for this discussion is that by investigating in this manner, we are able to see how such extant theories interact with the assumptions and other models incorporated into MAPES. Further, this provides motivation for the next perspective we should take on the BEG data, which is to see how the job market, in terms of wages and required education, changes as a result of immigration. This may, in turn, help resolve our previous conundrum.

Figure 5.10 shows the relationship between education and the percent of time per simulation that particular jobs are filled. For example, a point at education 10 and $pct\_filled = 90\%$ is a job that has an ideal education level of 10, and is filled 90 ticks out of 100 per simulation.

Figure 5.10 shows that there is indeed a clear relationship between ideal education and job occupancy. The dotted black line shows an non-linear least squares (NLS) regression fit to a parabola - the fit appears to be quite good. Jobs on both the low and high ends of the education scale are occupied far less than those in the middle. This makes sense for several reasons.

First, we know that the populations generated were normally distributed in education, so figure 5.10 represents that to a certain extent. Further, on the low end of the distribution, even if a person has an education level that is appropriately low, he will aspire to move to a job where he can earn more income (since income is one of the factors that persons maximize). This means he will not stay in that job long, hence the steep drop-off on the left side. On the high end, there are simply not enough persons of the required education level to fill all positions at all times. This is because a person with education less than the ideal point has probability of getting fired greater than 0.

One caveat about figure 5.10 is that while it is tempting to view all jobs as part of a unified job market, that is obviously not so since they are divided among separate cities and countries. While the two markets combined make for a fairly convincing continuity, taken separately, this is not quite the case. At the very least, jobs requiring higher education in city B (viz. education > 12) do not take quite the downward trajectory that those in city A do. And by the same token, low education jobs in city A do not take the steep downward trajectory as those in city B.

More obvious are the patterns highlighted by the solid red and blue lines. These represent linear regression fits to data when separated by city. Naïvely interpreted, they seem to suggest that education has different effects in the different cities. However, common-sense-wise, this does not seem likely and is probably an artifact of the fact that the domain of education is quite different between the two cities. More likely, the parabolic relationship earlier discussed is closer to the truth.[5]

Moving on, we can examine the relationship between job occupancy and wages using

[5]Unfortunately, even though these data are generated by a simulation, we cannot just add more jobs at the sparse educations levels since that would change the fundamental premise of the BEG scenario.

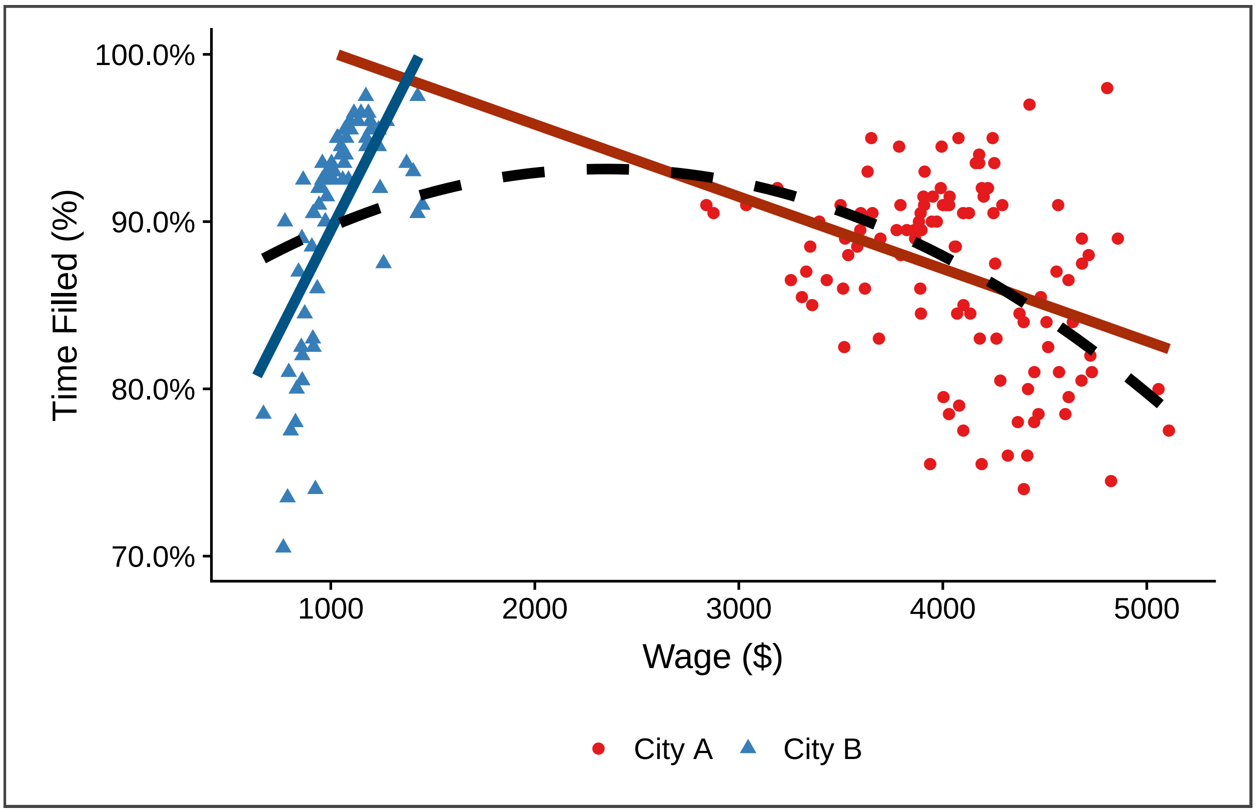


Figure 5.12: BEG: Wage and Percent Time Job is Occupied

figure 5.12. Similar to the relationship between education and job occupancy, there appears to be an inverse U-shaped relationship between the two indicated by the solid black line. Here, the disjoint between the labor markets in the two countries is made much clearer than in figure 5.10: there is a clear dictontinuity in wages between A and B. Again, the red and blue lines show the misleading effect of the discontinuity. That figures 5.10 and 5.12 look so similar should come as no surprise. After all, jobs are generated to have a high correlation ($\rho = 0.8$ per table 5.1) between wage and ideal education. This leads us to the question of whether income and education then are basically representing the same substantive effect, or if there is something different between them.

Table 5.2 shows several different regression models to help us understand the impact of education and wage on job occupancy. Columns 1 and 2 represent the parabolic fits shown in figures 5.10 and 5.12. Column 3 shows the collinearity discussed above: the coefficient for the wage and city interaction term (labeled "Wage/1k x In City B") loses statistical significance due to the high correlation between wage and ideal education. Things get interesting, however, with columns 4 and 5. Both models provide excellent fits of all the terms (indicated by the high statistical signficance of all the terms) and relatively high R^2. Neither of these models has a squared term for wage; instead both feature an interaction term between linear wage and being in city B, and an additional interaction term between education and city B in column 5. Substantively, these estimates indicate that there is a positive relationship between linear wage and occupancy, and that this is even stronger in city B.

The models presented in columns 4 and 5 are compelling because one can give a sensible explanation for what is observed. The core question is this: why does interacting wage with city make the wage variable in models 4 and 5 perform so much better than that in model 3? Recall from section 4.1.1 that (1) individuals consider moving with probability inversely proportional to average utility (including income utility); and that income maximization only comes into play once they start thinking about actually moving. Until that point, what matters in individual decision making is income utility, not absolute income. Further, income utility is calculated locally, and is only positively associated with income until mean income for the city is reached. Then, it maximizes and no further utility is gained by increasing

	Percent Time Job is Occupied				
	(1)	(2)	(3)	(4)	(5)
(Intercept)	0.339***	0.787***	0.307***	0.314***	0.371***
	(0.050)	(0.024)	(0.052)	(0.069)	(0.071)
Education	0.097***		0.107***	0.072***	0.058***
	(0.008)		(0.010)	(0.008)	(0.010)
Education^2	−0.004***		−0.004***	−0.003***	−0.003***
	(0.0003)		(0.0004)	(0.0003)	(0.0003)
In City B x Education					0.012***
					(0.004)
Wage x In City B				0.280***	0.201***
				(0.035)	(0.045)
In City B				−0.131**	−0.140***
				(0.054)	(0.053)
Wage		0.121***	−0.025	0.056***	0.069***
		(0.025)	(0.024)	(0.013)	(0.013)
Wage^2		−0.024***	0.003		
		(0.005)	(0.005)		
N	180	180	180	180	180
Adjusted R^2	0.481	0.135	0.491	0.672	0.684

*p < .1; **p < .05; ***p < .01

Table 5.2: BEG: Wage and Percent Time Job is Occupied. Wage variables are measured in 1,000 MU increments.

absolute income.

Because of this last point, jobs in both cities providing below local mean income will have higher turnover as people in those positions will more frequently consider moving. All jobs in A, even the low paying ones, will be attractive to the surplus labor in B since they will be at the stage of their decision making where they are maximizing absolute income, not just utility. They will therefore migrate to fill even the lowest waged positions in A, keeping them occupied. However, the reverse will not be true since wages are so much lower in B than A. Therefore, the low paying jobs in A will tend to have higher occupancy than those in B, which then weakens the positive relationship between wage and occupancy in A.

By contrast, education should not experience a similar interaction effect because in the current implementation of MAPES, there is no mechanism by which persons maximize some sort of education based utility. The main role played by education in the job market is by determining probability of getting fired, with persons having less than the ideal education level of a job getting fired with $p > 0$. Since this mechanism is the same between the two countries, from a model interpretation standpoint, there shouldn't be a significant difference observed in the impact of education between the two.

This explanation is tested in model 5 by including interaction terms for both education and wage. The coefficient for the education interaction term in model 5, with a magnitude of 0.012, only increases the effect of linear education (magnitude 0.058) by 20%. That for wage, on the other hand increases the base effect of 0.069 by 0.201, a substantial 289% increase. Whether the 20% is something else substantive or merely an artifact, the difference in degrees is stark enough to lend credence to the above explanation. Further, beyond the fact that this is all interesting from a substantive point of view, it is also an elegant demonstration of the way in which MAPES's known model specification can be verified through analysis of the output data.

We now return to the original question posed at the start of the section: how does immigration affect the job market? Based on job occupancy alone, we can make a couple of suppositions. First, there are definitely sectors of the job market that are not getting saturated, even with migrants in open competition with natives. Specifically, both the low and high ends of the job market have a significant amount of slack. For the reasons discussed earlier, about this not being a fully unified job market, there may be differences in the education levels at which there is low occupancy but the principle appears similar for both cities. Second, there is a sector of the job market, between education levels 9 and 14, that appears to be more or less saturated in both cities. Then, assuming that migration occurs at all education levels (shown to be true in section 5.4), there is certainly some degree of competition between natives and non-natives at those levels. The question becomes, what is the extent of this competition? Are we talking about 1% foreign applicants or 51%? To answer this question, we turn to figures 5.13 - 5.14.

Figure 5.13 plots by city the relationship between the ideal education level for each job against the percent of applicants that were foreign anytime that job had an opening. Each point represents an opening, so some jobs may be represented by more than one point, while others may show no points in these plots. The story presented by these plots is multipartite but clear nonetheless. In city A, migrants from city B dominate the applicant pool for jobs at the low end of the education spectrum, comprising between 75% and 100% of all job applicants for jobs requiring education level 9 or lower. As the ideal education level rises, they come into direct competition with natives, going 3 to 1 (i.e. 75% of applicants) for jobs requiring education level 11, 1 to 1 (50%) for level 13, falling to less than 1 to 3 (25%) on average by level 16. Finally, for jobs at the far right of the spectrum, they offer little to no competition. In city B, the opposite trend is true, with migrants from A offering no competition on the low end of the spectrum, rising steadily to comprise the majority of applicants on the far right. One important thing to note about these plots is that migrants dominating a particular job segment does not mean that they are pushing out natives; rather, these are jobs that natives are not willing or qualified to compete for, else the ratio of foreign to native applicants would not be high (more on this below). This applies to the low end of the spectrum in city A, and the high end in city B.

Another important aspect to consider is not just the percent of applicants that are foreign, but the actual number. While it would be frustrating to be the native 50% of only two

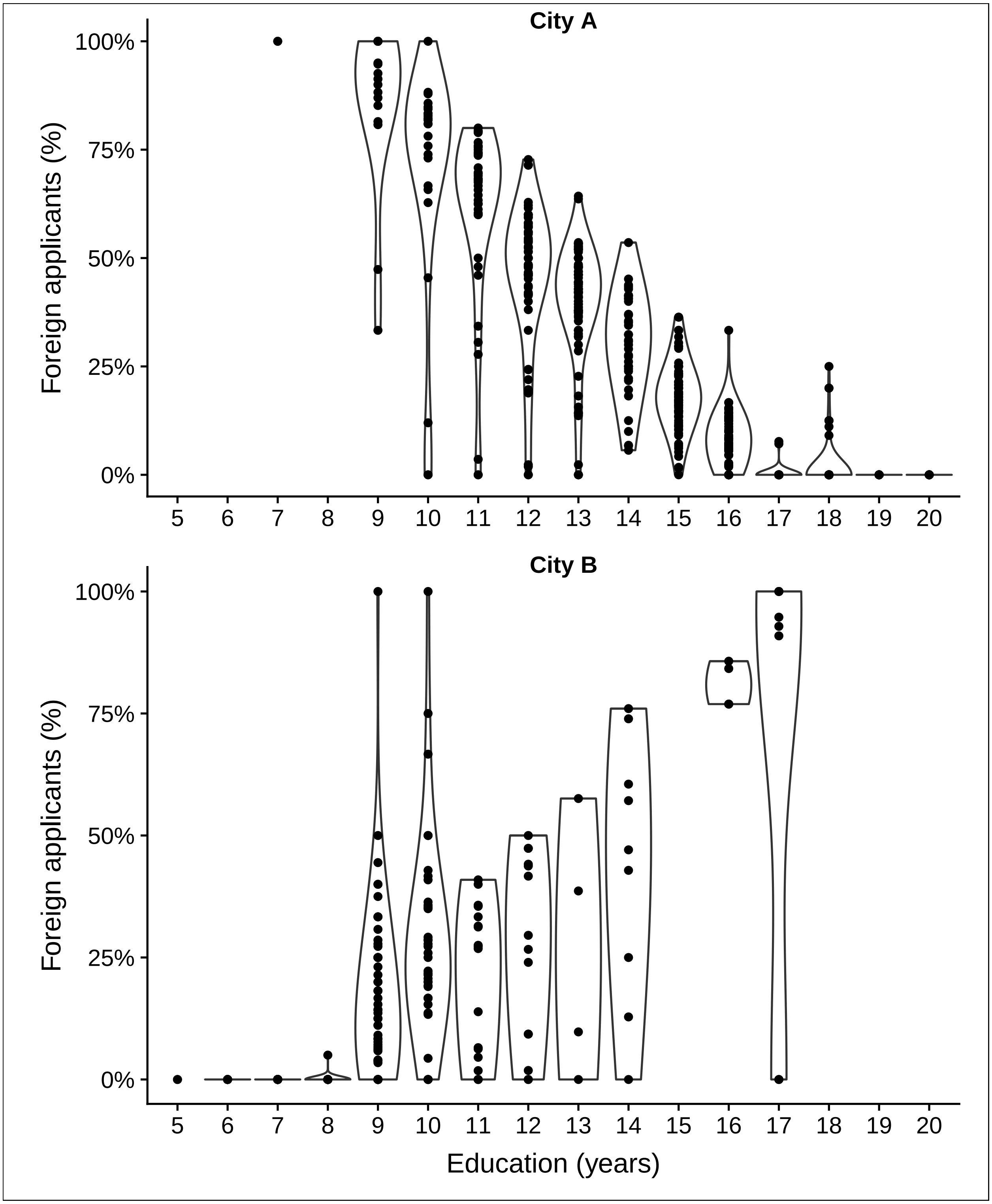


Figure 5.13: BEG: Percent foreign applicants as a function of requisite education level

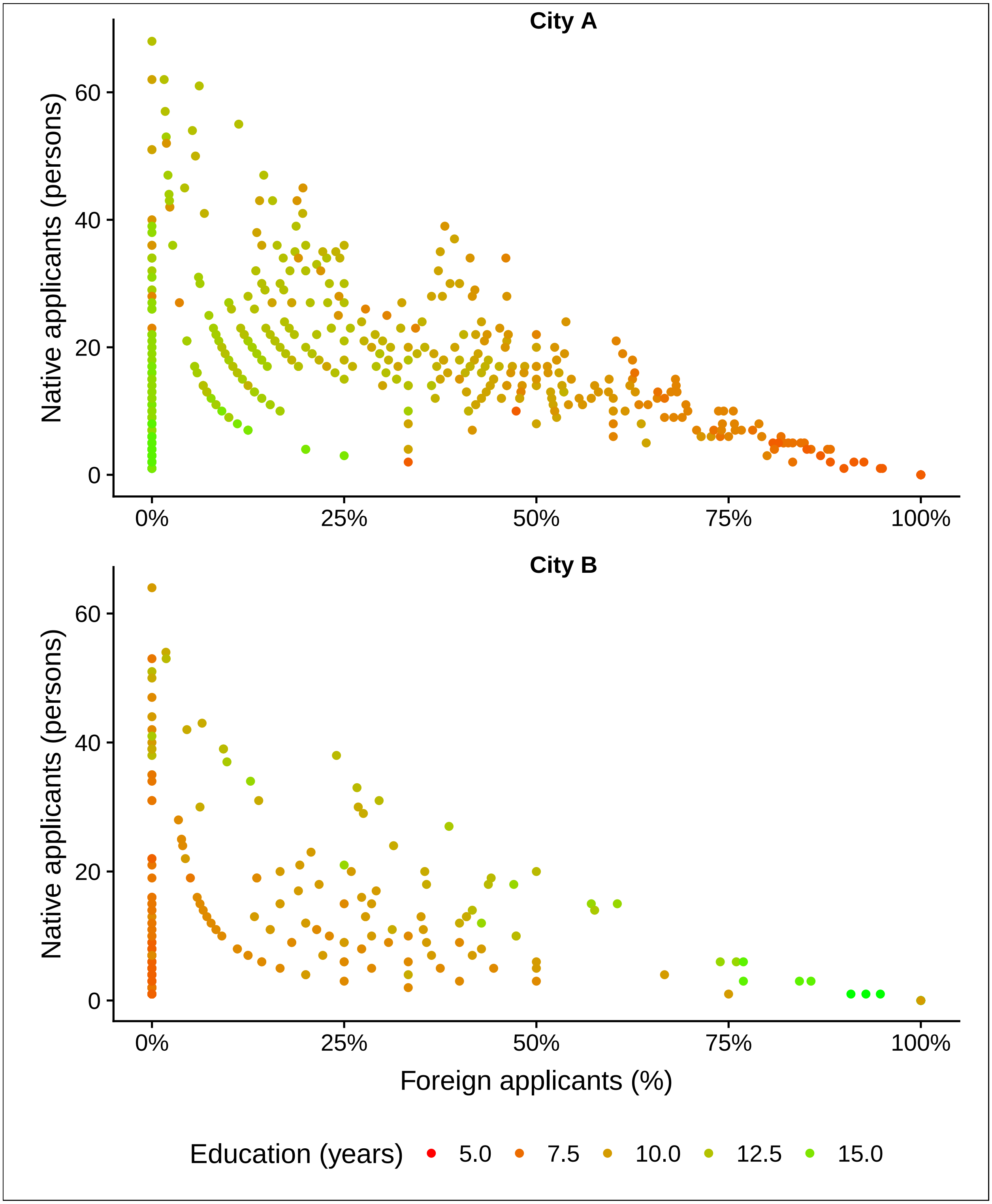


Figure 5.14: BEG: Impact of foreign applicants on native jobseekers

applicants that loses out to an immigrant competitor, the impact to the overall job market is far less than 50 natives competing against and losing to one of 50 immigrants. Figure 5.14 allows us to see how many natives are impacted by competition from migrants. I plot the number of native applicants for each opening against the percent of applicants that are foreign, and the color of points indicates the ideal education level for that job. The further upwards and rightwards one moves, the more natives are being pressured by immigrant competitors. These figures corroborate the earlier interpretation of figure 5.13.

In city A, high education jobs inherently have a high degree of competition between natives, but no additional pressure is brought to bear by migrants. This is indicated by the stack of green points with small X-axis values. Low education jobs have fierce competition among migrants but almost no natives are affected - the red dots on the bottom right of the plot. It is jobs in the 10-16 education level (orange to yellow points) in which natives face significant pressure from migrants, with between 10-30 natives facing pools that are 25%-75% foreign per job. The picture is reversed for city B, with high intra-native competition for low education jobs, low impact on natives for high education jobs, and large competition and impact from migrants for middle education jobs. Aside from the reversal of the education scale, the plots for the two cities look remarkably similar.

Figures 5.13 - 5.14 indicate that the positive and negative impacts of migration on labor in BEG align well with real world stereotypes. In the traditional receiving country (A), migrants take the jobs that natives are not willing to take, fulfilling a vital economic need. At the same time, they pressure workers in the middle of the education distribution, competing for jobs and lowering expected wages. The traditional sending country (B) is also impacted by migration, with migrants from A reciprocating pressure against middle education B natives. They also fulfill some of the need for high education persons in B.

It is important to realize, however, that these outcomes are specific to the parameters defined in BEG. For example, one phenomenon / stereotype that does *not* occur in BEG is that immigrants from B do not significantly compete for high education jobs in A, akin to the way H1-B visa holders do in the real-world US. Given BEG's parameters, it is clear why this does not occur here. In both cities, there is a shortage of high education persons relative to the number of openings available so there is little to no incentive for persons to migrate to fill those jobs in another locale. And while it is true that a job in city A pays better than a similarly high skilled job in B, one must not forget that there are also significant costs that the model imposes for moving, which keeps such movement low. The flip side to all this is that if we alter the starting conditions appropriately, by moving or widening B's person education distribution rightward, we should see such phenomena occur. However, doing so would basically change the substance of this scenario. The bottom-line is that the overlap in education distributions is what determines the sectors that are impacted, and in the case of BEG, it is the middle education natives in either country who face that pressure.

The natural next question to ask is: given these changes in the competitive landscape for jobs in both A and B, how are individuals' economic outcomes changed? The traditional, survey research based approach to answering this question would take one of two forms. One would be to regress income (or some other measure of economic well-being) on a variety of factors including the key IV of whether the person migrated or not. Another would be to do some kind of matching analysis to compare similar individuals who have and have not migrated, thereby making some inference about the effect of migrating. The major limitation of both these approaches is that while they can shed light on the effect that migrating would have had on one person given the existence of migration, they do not reveal the effect that migration has at a system level. That is to say, they cannot address the counter-factual of what a person's outcome would be if migration simply did not occur. Fortunately, with ABMs we are not subject to this limitation. We can simply run the model with the ability to migrate disabled (a simulation-level parameter), then compare the results against the exact same inputs and migration enabled.

Figure 5.15 shows the distribution of differences in lifetime earnings for each individual between the model with migration enabled, and the same model with migration disabled. That is to say, each of the 200 people in the simulation (100 from each city) has an entry in the dataset forming figure 5.15, with the datapoint equaling the difference between their (1) lifetime earnings in the migration-enabled simulation, and (2) the same in the migration-disabled one. Therefore positive values mean a person does better with migration; negatives

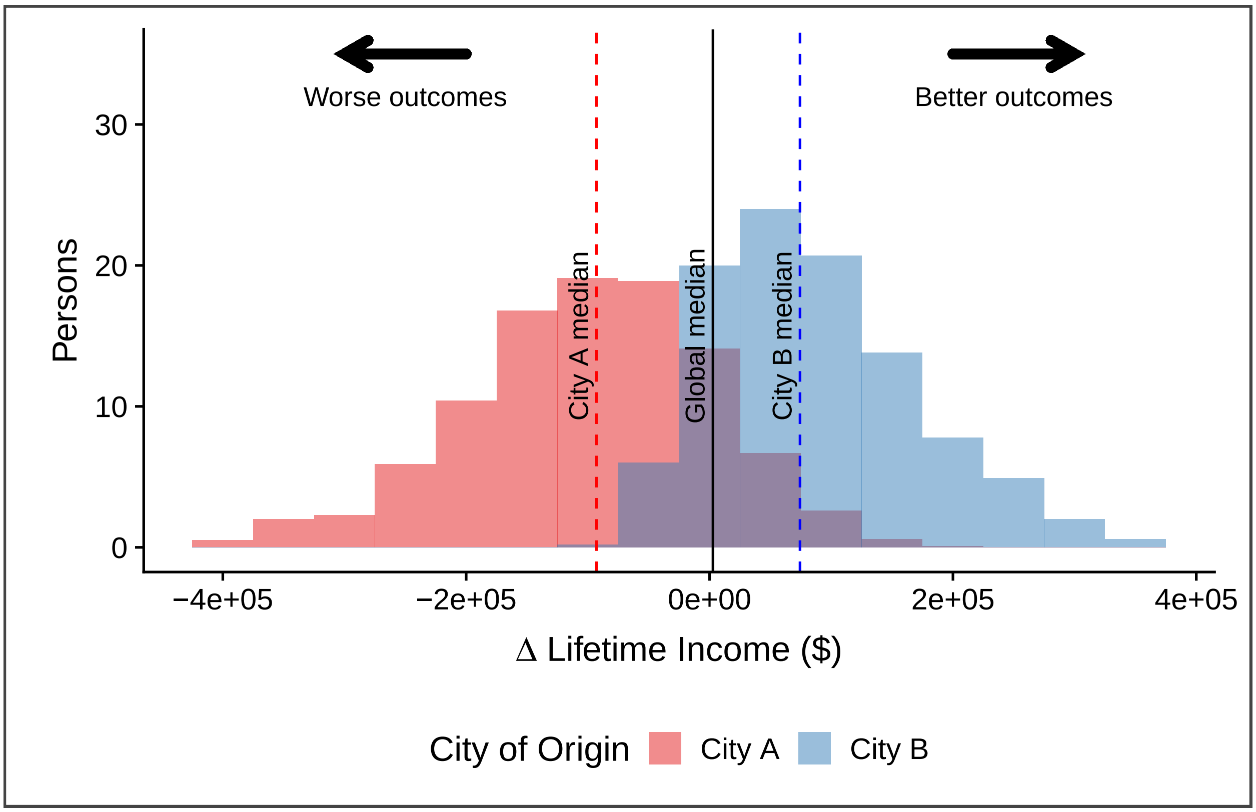


Figure 5.15: BEG: Difference in Financial Outcomes with and without Migration

	Change in lifetime earnings		
	(1)	(2)	(3)
(Intercept)	3,053.440*** (1,158.431)	13,572.940** (6,586.240)	−19,922.620 (23,700.070)
Education		−422.312 (260.290)	740.781 (811.591)
Education^2			58,931.860** (27,648.940)
From City B			−2,365.005** (1,046.481)
From City B x Education	195,330.800*** (5,695.322)	195,928.100*** (5,704.885)	−172,046.100 (187,117.700)
From City B x Education^2	−140,714.900*** (16,503.780)	−203,963.900*** (42,330.180)	33,653.280 (171,250.200)
N	2,000	2,000	2,000
Adjusted R^2	0.500	0.500	0.501

*p < .1; **p < .05; ***p < .01

Table 5.3: BEG: Change in lifetime earnings due to migration

mean they do worse.

Even upon a cursory glance, there is a clear difference in outcomes based on what city a person is from. While the median change is 2800 monetary units (MU) for all individuals (regardless of origin city), persons from city B by and large do better when migration is a possibility. For context, while these numbers are technically measured in MU, they were scaled to approximate real-world dollars when setting up the scenario, so one may think of them as such when trying to understand the magnitude of these values. Given that, 2800 across 100 ticks is insignificant on a per tick basis.

Of the persons from city B, there are some who do worse, but the worst case is relatively small (-1.10814×10^5 MU) and the median improvement is a not insignificant 7.4334×10^4 MU. The pattern is the opposite for folks from city A, with the median person clearly doing worse (-9.2885×10^4 MU) with migration than without.

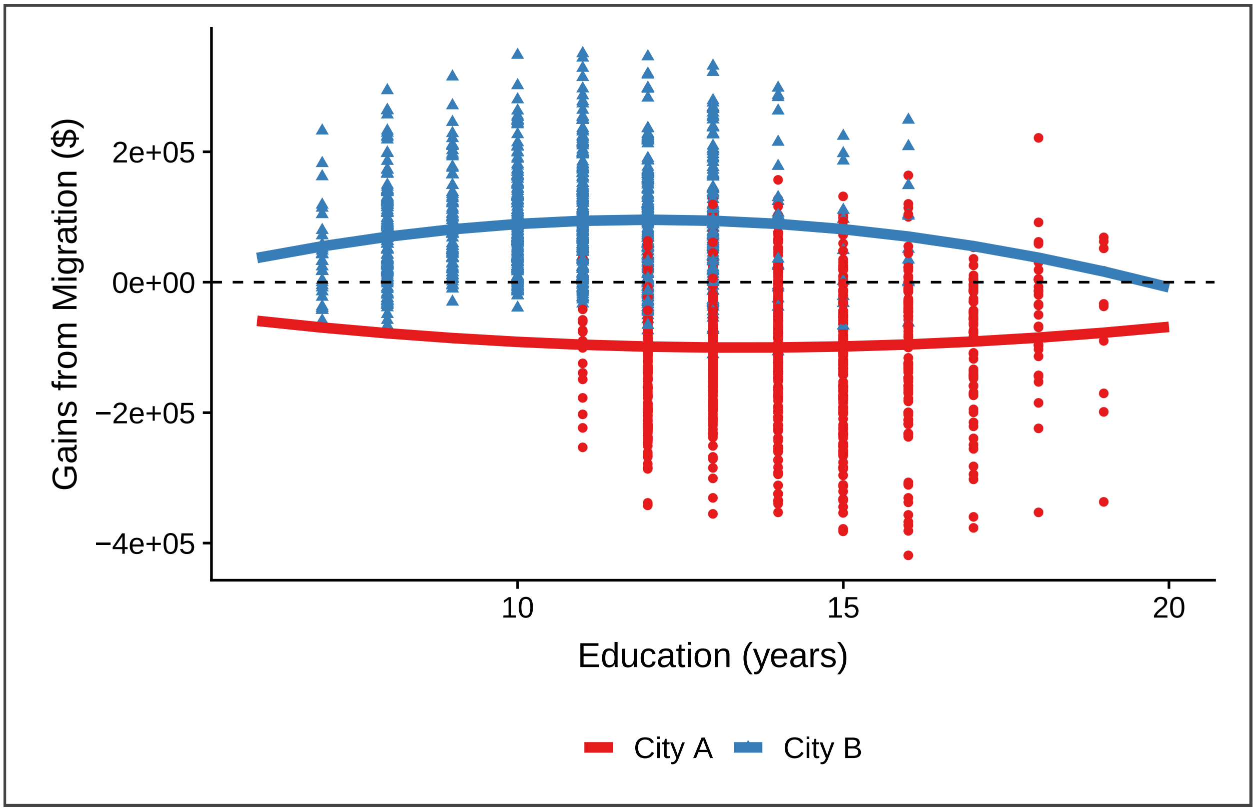


Figure 5.16: BEG: Education and Difference in Financial Outcomes with and without Migration

This result is validated through regression analyses, shown in table 5.3. These models regress the differences shown in fig 5.15 on the time-invariant characteristics for each individual (viz. starting city; education level, which MAPES assumes to not change). The predictors must be time-invariant because the outcome (difference in lifetime savings) is calculated on a per-simulation basis (not per time period) so the predictors must be as well.[6] The three models present a couple different combinations of the predictors to see whether the effect of education is modulated by city of origin. However, this does not seem to be the case; ultimately, it is mainly city of origin that explains the gaps. Over the lifetime of the simulation (100 ticks), the total difference according to the models in table 5.3 is between 1.94575×10^5 and 1.95928×10^5 MU over the lifetime of the simulation, or 1945 and 1960 per tick. The estimates of model 3 are shown in figure 5.16. This figure confirms that education has some effect but it is overshadowed in magnitude by city of origin. The bottom line is that *for the specific scenario modeled in BEG*, allowing migration between A and B hugely benefits the population of B, while it has a fairly negative impact on lifetime earnings for persons from A.

5.4 Persons

Thus far, we have examined the impact of migration on politics, economics and populations, but we have not looked at the impact on the persons themselves. Which persons are doing what, and what are the impacts of these actions on their outcomes?

Let us begin by examining who is migrating. Figure 5.17 shows the frequency of migrations by city of origin for the pooled sample of all persons across all simulations. To clarify, each unique individual appears 10 (number of simulation runs) times in the sample for figure 5.17 since each run starts with an identical set of agents. The points are the actual data and the lines show the NLS (non-linear least squares) regression fit to an exponential curve for each city's data. In both cities, the frequency of simulation-persons decreases with increasing

[6] Arguably, one can also include a variable like income, which can change per period, averaged to a single value for the entire simulation. However, that is not used here since averaged income is endogenous to the outcome.

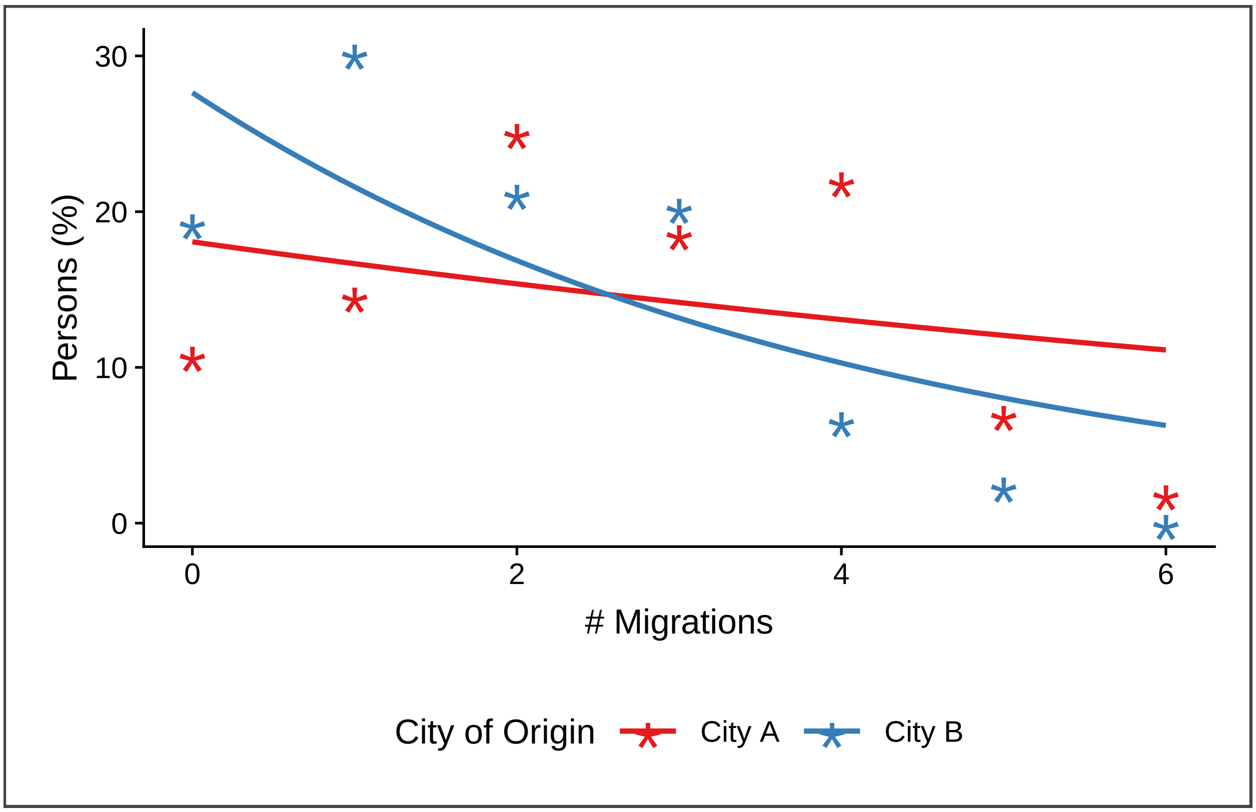


Figure 5.17: BEG: Migrations per individual

number of migrations, with the decline being more precipitous for city B. For the traditional "receiving" city, city A, 10.8% of simulation-persons migrated not at all, 39.7% migrated 1-2 times, 40.6% 3-4 times, and 8.9% 5 or more. The respective figures are 19.3%, 51.4%, 26.9%, and 2.4% for simulation-persons from city B. This figure indicates that even though we parameterized city A as the "receiving" location, in fact persons originating from city A tend to move more often than persons from B, the "sending" city. This is a surprising result given that we know there is a net flow from B to A over the duration of the simulation. To understand why, several interrelated points must be considered.

First, it is important to note the implications of figure 5.17 being based on a pooled sample. To say that 25% of simulation-persons from city A didn't migrate is not necessarily the same as saying that 25% didn't migrate in any given simulation. It could be that the same 25 individuals consistently didn't migrate across all runs, or that each of the 100 people originating from city A stayed put in 25% of simulations. These are very different propositions. The former case would suggest that there are some circumstances that are (more or less) deterministic for not migrating. The latter situation would mean that not migrating is purely stochastic. The truth almost certainly lies somewhere in between, and to see this, we may turn to figure 5.18.

Figure 5.18 shows histograms of the number of moves for each individual averaged across all simulations, by city of origin. In other words, this figure takes the data for figure 5.17 and aggregates it by individual using the mean as the aggregating function. A person who moved 0 times in 3 simulations and 4 times in another will have 4 entries in the previous figure, but a single entry for 1 (i.e. $(0 + 0 + 0 + 4)/4$) in this one. This means that if there are persons who moved not at all across most or all simulations, they will show up as close to 0 in figure 5.18. Conversely, the further away from zero someone is, the more likely it is that they moved at least once in any single simulation. Note again that this is not the same as saying that they moved more frequently within a particular simulation.

The leftmost bins of the city B histogram are less than 1 (and some even abut the 0 mark), indicating that there are persons who do not move at all in some simulations, and may even remain sedentary across *all* simulations. For the data used here, on average 1% of

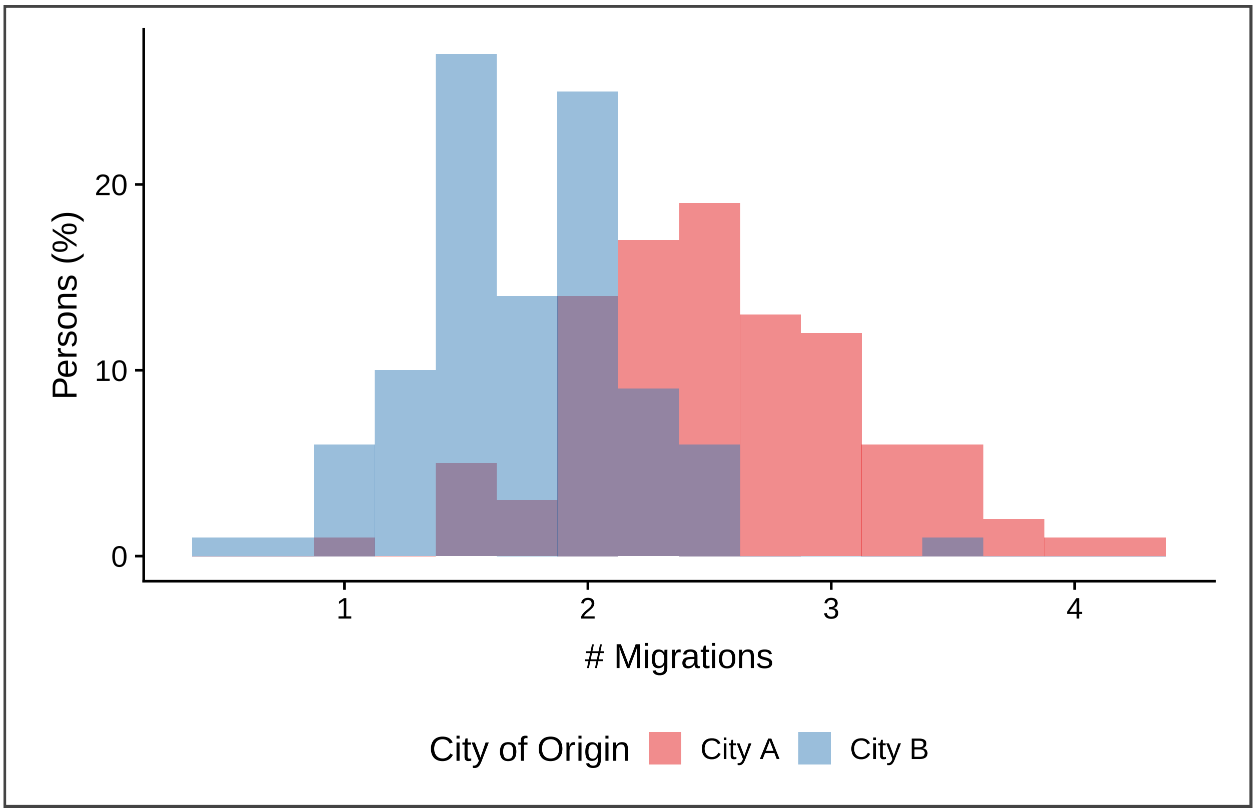


Figure 5.18: BEG: Mean Migrations per Simulation, per Individual

city B's originating population moved less than or equal to 0.5 times on average. Further, a significant portion of B (5%) moved less than or equal to once across all simulations. The flip side of this is that 95% of persons originally from city B moved more than once on average, significantly greater than what figure 5.17 would have led us to believe (80.7%). The 95% is comprised of 75% persons who migrated between once and twice on average, and 20% greater than twice, with no one having migrated more than three times on average.

The histogram for city A features several key differences from that for B. A key difference is that it shifted significantly rightward, indicating that very few (0%) moved less than or equal to once on average. 18% of people from city A moved between once and twice across all simulations, 65% between twice and three times, and 17% actually moved more than three times on average. Based on figure 5.17, we know that in some simulations, there were persons who migrated up to 6 times but figure 5.18 tells us that these were one-off occurrences that were not repeated across multiple simulations by the same individuals (else their averages would stretch further right in figure 5.18). Figure 5.18 mostly confirms what we see in figure 5.17, that persons from city A do indeed tend to move more, even when examined at the individual level (as opposed to individual aggregated across simulations). It does offer one major refinement, however, which is that almost everyone is liable to move at some point (i.e. almost no one has a mean value less than 1). In other words, the causes of migration are less deterministic for persons originating from city A – the proverbial "it could happen to anyone."

The next point we must consider is the difference between even and odd numbers of moves. Someone who moves an even number of times ends up back in their place of origin; an expatriate living abroad temporarily but eventually returning. A person who moves an odd number of times, on the other hand, is an émigré who permanently ends up in a foreign land. Figure 5.19 shows how this point applies to these data. It is similar to figure 5.18, except it takes the median rather than the mean for each individual. This small change is sufficient to show that beyond just the number of moves, the character of migration from the two cities is quite distinct. Almost a supermajority of persons from city A (46%) tend to be even numbered movers – the largest bin for city A in figure 5.19 is at 2 moves, and

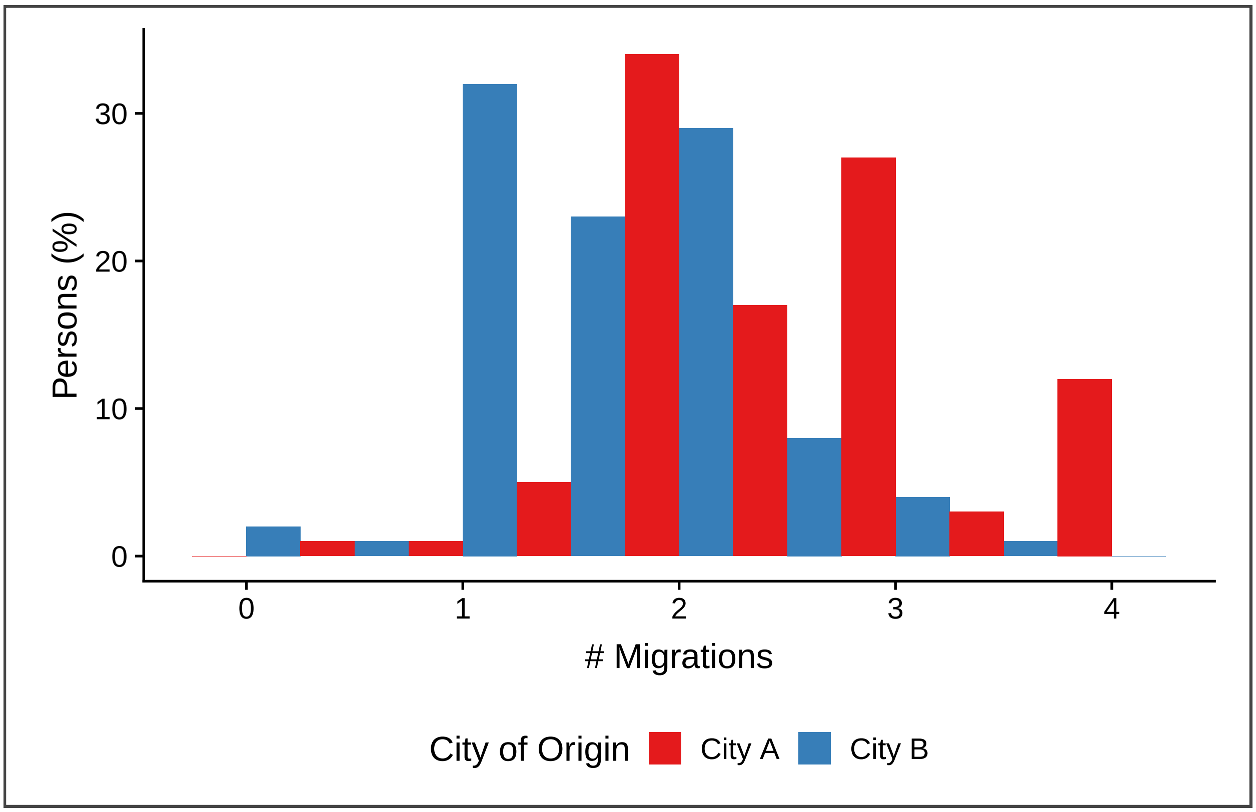


Figure 5.19: BEG: Median Migrations per Simulation, per Individual

the bin for 4 moves is not much smaller than that for 3. On the other hand, only 29% of persons from B are even movers, while 36% are B odd-movers, versus the 28% odd movers of A. This means that permanent migration (i.e. movement conducted without the intention of eventually returning) is much more common among persons from B than from A.

Figures 5.17-5.19, in toto, indicate that persons from A tend to move more overall, but their moves are less permanent than those of persons from B. Based on our knowledge of the model, we know this has to be because these migration patterns enable their respective performants to maximize utility. Some of the mechanisms involved have been alluded to in the previous sections, 5.1 – 5.3, and they will be further explored shortly.

But before we discuss what *is* causing the observed patterns, let's make clear what *is not*. That is anything not included in the model specifications covered in chapter 4. In an EBM context, even if a variable is not accounted for in analysis, the effect of the underlying phenomenon is still felt in measurements, potentially compromising that analysis. Here, because the researcher is programming the rules, "latent" variables are not just unseen – they do not exist. Therefore it is not fruitful to ask of the subsequent analysis why the impact of some factor X is not considered. More useful is to ask how the results would have changed if that factor X had been initially modeled in the ABM. While interesting, what answers may be given to such questions is more speculation than results, so are presented not here but in chapter 8.

So who exactly is migrating and why? Recall from the model description (section 4.1.1) that migration occurs as follows. Every tick, each person undertakes a life change with increasing probability in inverse relation to their average utility level (i.e. the unhappier they are). Life changes can be moving to another city, or moving jobs within the same city. Utility is composed of direct economic and social components, and an indirect political one. Social utility increases linearly with the size of the person's local (i.e. same city) social network, which grows probabilistically over time in that locale. Economic utility is calculated based on the ratio of each person's income to average local income – to a certain point, higher income leads to more utility. And since the pay offered for jobs has a positive relationship with the level of education required,[7] it is reasonable to posit a positive relationship between education and economic utility.

All of this is to say that for BEG, income, education, and social network size "should" all have negative relationships with propensity to move. However, there are a couple of caveats. First, there is potential ambiguity in the direction of causality for social network size. This is because smaller network size will definitely increase probability of migrating, but migrating will also decrease network size if the move is to a city that the person has never visited before, and few of his prior associates have moved to (if there were many prior associates in the destination, the person would move into an existing social network despite the city being completely new to them). Further, there is confounding with the fact that migrants in this implementation of MAPES only apply to jobs after arrival, so there is at least a 2-3 tick window in which economic utility will bottom out as well. With 2 of 3 utility components zeroed, there is high probability that the person will immediately choose to move again – an unrealistic case of migration begetting more migration.

Placing a timeout on migrations can help mitigate this problem, by allowing the person's utility time to recover before they consider any further life changes. The timeout allows them to (1) potentially acquire a new job, and (2) start building his social network. The level of social recovery is a function of the timeout length, and the probabilities that go into forming relationships (see section 4.1.1). MAPES by default uses a 10 tick timeout, which in conjunction with P(social interaction) = 0.5, P(negative interaction with foreigner—social interaction) = 0.5, P(form social edge—positive foreigner interaction) = 0.2, leads to $P = 0.401$ of the migrant forming at least one social tie with a native during the timeout. These values were chosen arbitrarily, but do appear to have the intended effect of preventing too much ping-ponging while not leading to too-rapid network growth.

A second caveat is that while education "should" cause migration, the relationship may not be a simple linear one. We already saw in figures 5.15 - 5.16 that education and income can have nonlinear relationships with migration-related outcomes. Further, both

[7]This is defined in the scenario-specific covariance matrix used for initially seeding jobs.

real-world evidence and consideration of model mechanics strengthen this hunch about non-monotonicity. In the real world, we know that while many migrants are from the bottom of the economic totem pole, "brain-drain" is also a real phenomenon. Within the model, we know that there are a finite number of jobs at any tier of education, so the negative relationship between education and migration will exist only insofar as a particular locale can fully absorb persons at all levels of education. If, for example, there is a shortage of jobs requiring high education, persons with high levels of education will be forced into jobs that pay less than they would be getting without the limitation. This could potentially break the negative relationship between economic utility and education, and therefore between education and propensity to migrate.

Let us start answering these questions by examining tables 5.4-5.7, which several regression models of the relationship between social network size and education, on the one hand, and individual propensity to move on the other. First, the models in tables 5.4-5.5 regress the average number of migrations undertaken by each individual over all simulations as a function of time-invariant or statically defined characteristics. As before (see the end of section 5.3 for a discussion on time invariant vs variant characteristics), this is because the outcome (average number of migrations) is calculated on a per-simulation basis, not per time period, so the predictors must be as well. As in table 5.3, income is not used here because of endogeneity: migration can be caused by a low average income, but it can cause it as well. Second, table 5.6 shows regressions modeling the per-tick incidence of migration (i.e. did a person migrate in a given period) as a function of both static characteristics as well as time-variant characteristics, such as income in a given period.

Models 1-4 in table 5.4 regress *mean* number of migrations on the variables of originating city (factor) and education (years). The models respectively use (1) city and education; (2) education squared; (3) city interacted with education; and (4) city interacted with education squared. Raw education ranges 6-20 years and education squared ranges 36-400 years^2, and using the latter permits a non-linear relationship between education and the outcome. Models 5-8 (table 5.5) use the same predictors in a Poisson regression to predict *median* number of migrations transformed to a dicrete (i.e. integer values) variable.

	Mean moves per simulation (OLS)			
	(1)	(2)	(3)	(4)
(Intercept)	0.026	0.022	0.067^{**}	1.803^{***}
	(0.020)	(0.116)	(0.030)	(0.398)
Education		0.0002		-0.060^{***}
		(0.005)		(0.014)
Education^2			-0.074^{*}	-1.697^{***}
			(0.041)	(0.464)
From City B				0.054^{***}
				(0.018)
From City B x Education	-0.743^{***}	-0.743^{***}	0.179	12.043^{***}
	(0.100)	(0.101)	(0.518)	(3.141)
From City B x Education^2	2.183^{***}	2.209^{***}	1.617^{***}	-10.834^{***}
	(0.290)	(0.747)	(0.425)	(2.875)
N	200	200	200	200
Adjusted R^2	0.394	0.391	0.401	0.450

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$

Table 5.4: BEG: Causes of Migration

	Median moves per simulation (Poisson)			
	(5)	(6)	(7)	(8)
(Intercept)	0.013	0.129	0.016	0.805^{**}
	(0.019)	(0.119)	(0.026)	(0.382)
Education		-0.004		-0.027^{**}
		(0.005)		(0.013)
Education^2			-0.006	-0.507
			(0.039)	(0.479)
From City B				0.014
				(0.019)
From City B x Education	-0.434^{***}	-0.415^{***}	-0.367	3.805
	(0.096)	(0.098)	(0.486)	(3.153)
From City B x Education^2	1.459^{***}	0.726	1.426^{***}	-4.243
	(0.277)	(0.793)	(0.363)	(2.758)
N	200	200	200	200
Log Likelihood	-370.646	-370.149	-370.636	-367.850
AIC	747.292	748.299	749.273	747.700

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$

Table 5.5: BEG: Causes of Migration

The overall picture presented by these models is that there can be a strong non-linear relationship between education and migration habits contingent on city/country of origin. First, we see from models 1 and 5 that education not interacted with city of origin has an effect on number of moves that is negligible in size and has a large p-value. This could happen for 2 major reasons: either education actually doesn't matter (unlikely) or subpopulations within the sample experience education differently, resulting in a wash. City of origin has a large and statistically significant effect (negative for being from city B), which is not surprising. The picture does not change for models 2 and 6, despite adding education squared. Things become more interesting once one interacts education with city of origin. Education in models 3 and 7 shows effect sizes that, while objectively still small, are several times the magnitude of those in models 1,2, and 4. Also in model 3, both education and the interaction term attain statistical significance, which is what we want given our prior beliefs about education mattering. Finally, models 4 and 8, which allow the non-linear effects of education to vary by city, show large and statistically significant effects for education and education squared in model 8, and almost all variables in model 4.

Models 4 and 8 clearly indicate that just as with individual economic outcomes, (1) education has an inverted U-shaped relationship with the number of moves a person is likely to make, and (2) this effect is highly contingent on whether the person is originally from the "sending" or "receiving" country. In the domain of meaningful education levels (9-20 years of school), the U shapedness is very marked for persons from city A while it is relatively shallow for those from city B (figure 5.20). Resultantly, for persons from city A, education levels in the lower 50% percentile show a strongly positive impact on the average number of moves made across all simulations. For those in the upper 50%, the effect reverses, with more education leading to fewer moves. According to the point estimates from model 4, persons originating from city B only experience a declining effect in the range of 9-20 years of education.[8] That is to say, more education results in fewer moves for them across

[8]It is important to note that the magnitudes of education values are not meaningful by themselves since they are randomly assigned to persons and the wage values they are correlated with are arbitrarily generated as well. In other words, the range of education values used here could just as easily been 109-120 rather than 9-20. The data were generated with education values ranging 9-20 because it is conveninent to think of

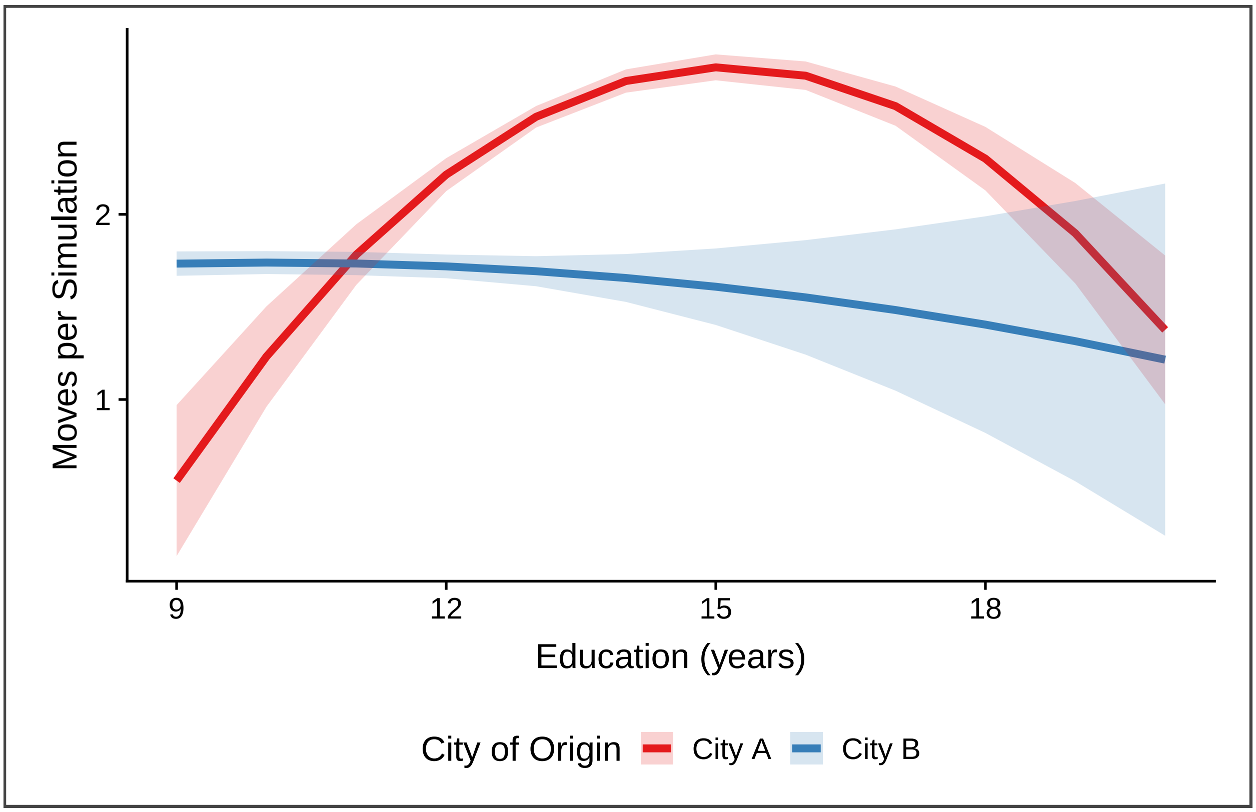


Figure 5.20: BEG: Mean Migrations per Simulation, per Individual

all meaningful levels of education. There appears to be some ambiguity for those in the upper 50% of education values because of the the huge confidence intervals in that domain. However, this turns out simply to be because the education distribution for city B specified during data generation led to few persons with education greater than 16, leading the fit in that domain to have huge standard errors (figure 5.1).

Moving on to the second set of regressions, the models in tables 5.6 - 5.7 regress whether an individual migrates in a particular tick (migrate = 1 for move, 0 for no move) on the different factors that go into the migration decision per the ABM model specification. These are utility, factors affecting utility (wage and local network size), and factors affecting mobility (wage and education). City of origin and associated interaction effects are included to account for what we learned from the static regression models. City of residence is used to see whether being *in* a particular city, independent of being *from* there, affects the outcome. Local network size is the number of social ties the person has in their current location of residence. Finally, simulation period is used as a catch-all measure for latent factors that would cause propensity to migrate to increase over time (i.e. emergent phenomena not explicitly programmed in).

Six models are given as a means of validating the existence of relationships we know to exist from the ABM specifications, and the process of model selection. Models 1-4 show different combinations of measuring income utility and social utility using the actual utility numbers (akin to asking a survey respondent "how happy are you with your financial position / social life") and using the underlying phenomena they derive from, wages and local social network size. Since they are combinations of the 4 variables, we should compare them pairwise. First, comparing model 1 to 2, and 3 to 4 shows that wages and income utility are conceptually interchangeable. This is because the negative components on both measures indicate that as we expect, the more money a person makes, the happier the person, and less likely to migrate.

On the other hand, the expected relationship between social utility and propensity to

education in terms of real-world landmarks (high school, college and graduate school graduation). In truth, however, these numbers are ordinal, not cardinal, with the actual values having no real world significance.

	Pr(migrate) in any given tick		
	(1)	(2)	(3)
(Intercept)	-11.576^{***} (3.652)	-14.556^{***} (3.687)	-14.557^{***} (3.687)
Tick	0.017^{***} (0.003)	0.017^{***} (0.003)	0.017^{***} (0.003)
In City B	-0.736^{***} (0.168)	-0.364^{**} (0.169)	-0.363^{**} (0.169)
From City B x In City B	0.461 (0.347)	1.220^{***} (0.330)	1.221^{***} (0.330)
From City B	4.259^{**} (1.968)	6.104^{***} (1.987)	6.100^{***} (1.987)
From City B x Education	-0.461^{***} (0.157)	-0.601^{***} (0.159)	-0.601^{***} (0.159)
Education	1.343^{***} (0.504)	1.631^{***} (0.510)	1.632^{***} (0.510)
Education^2	-0.046^{***} (0.017)	-0.055^{***} (0.018)	-0.055^{***} (0.018)
Income Utility		-0.057^{***} (0.003)	-0.058^{***} (0.003)
Avg. Wage	-1.176^{***} (0.075)		
Social Utility	-0.003 (0.003)	-0.003 (0.003)	
Local Network Size			-0.023 (0.023)
Has Ever Migrated?	-1.168^{***} (0.192)	-0.902^{***} (0.186)	-0.905^{***} (0.187)
N	19,800	19,800	19,800
Log Likelihood	$-1,531.966$	$-1,638.899$	$-1,638.918$
AIC	3,085.931	3,299.798	3,299.835

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$

Table 5.6: BEG: Pr(Moving) per Tick

	Pr(migrate) in any given tick		
	(4)	(5)	(6)
(Intercept)	-11.535^{***} (3.650)	-11.933^{***} (3.482)	-10.145^{***} (3.635)
Tick	0.017^{***} (0.003)	0.034^{***} (0.003)	0.017^{***} (0.003)
In City B	-0.742^{***} (0.167)	-0.366^{**} (0.171)	-0.750^{***} (0.168)
From City B x In City B	0.469 (0.347)	1.151^{***} (0.316)	0.216 (0.349)
From City B	4.232^{**} (1.968)	4.407^{**} (1.905)	3.341^{*} (1.961)
From City B x Education	-0.460^{***} (0.157)	-0.455^{***} (0.153)	-0.401^{**} (0.156)
From City B x Avg. Wage			0.664^{***} (0.159)
Education	1.338^{***} (0.503)	1.209^{**} (0.480)	1.165^{**} (0.501)
Education^2	-0.046^{***} (0.017)	-0.040^{**} (0.016)	-0.041^{**} (0.017)
Avg. Wage	-1.178^{***} (0.074)		-1.367^{***} (0.110)
Social Utility			-0.004 (0.003)
Local Network Size	-0.029 (0.023)		
Avg. Utility		-0.055^{***} (0.004)	
Has Ever Migrated?	-1.177^{***} (0.193)	-0.816^{***} (0.180)	-1.169^{***} (0.192)
N	19,800	19,800	19,800
Log Likelihood	$-1{,}531.821$	$-1{,}720.774$	$-1{,}523.866$
AIC	3,085.643	3,461.547	3,071.733

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$

Table 5.7: BEG: Pr(Moving) per Tick

	Ticks in Locale	Local Network Size	Avg. Utility	Social Utility	Income Utility	Avg. Wage
Ticks in Locale	1	0.790	0.747	0.792	0.299	0.218
Local Network Size	0.790	1	0.844	0.980	0.212	0.154
Avg. Utility	0.747	0.844	1	0.864	0.683	0.476
Social Utility	0.792	0.980	0.864	1	0.222	0.163
Income Utility	0.299	0.212	0.683	0.222	1	0.687
Avg. Wage	0.218	0.154	0.476	0.163	0.687	1

Table 5.8: BEG: Correlation of Time in Locale and Utility

migrate does not clearly materialize in these models. What is expected is that when models 2 and 3 are compared, the effect of social utility and local network size should be negative and statistically significant; this is not the case. Moreover, when models 1 and 4 are compared, the signs flip (albeit consistently between the two) and statistical significance is still not achieved. It is not completely clear why any of this is the case, given that we know social utility is used directly to calculate the migration decision. To throw more fuel on the fire, when a person's time in a locale is used in the model rather than the actual social utility or network size measures, it attains a positive sign and is statistically significant. In effect, it says that the longer a person has been in a place, the more likely they are to migrate – the exact opposite of what is expected.

All of this is quite puzzling because there are good reasons to think that these three measures (social utility, local network size, and time in locale) should all consistently decrease migration likelihood. Based on model specifications, we know that a person's utility is largely dependent on their time spent in a particular location for at least 2 reasons. First, social networks grow over time and do not degrade. Second, for persons with $Pr(\text{find employment}) > 0.5$, the probability of getting a job, and therefore getting income utility and average utility greater than 0, increases with time (not in the sense of gambler's fallacy but that $(1 - Pr(\text{not finding employment}))^t$ decreases with increasing t if $Pr(\text{find employment}) > 0.5$). Further, table 5.8 supports the existence of these relationships since there is a high positive correlation between several of the combinations of these measures.

At this juncture, there are at least three possible explanations for this puzzling behavior. One possible explanation arises from fact that, as mentioned earlier in this section, persons who move are subject to a timeout before they can move again. The purpose of this timeout is to prevent ping-ponging between locales in an unrealistic manner. However, this also has the effect of raising the probability that someone who has already moved will have greater than 0 local network size prior to moving again (since ties are formed probabilistically every tick). In short, the timeout will have the effect of creating a positive correlation between larger local network size and moving. A second similar possibility arises from the fact that in the current version of MAPES, social ties do not degrade. So with increasing time, each mover will on average have more ties than movers from earlier periods. This too will create a spurious correlation between time in locale and likelihood of moving. Finally, it may simply be the case that there is some interesting non-linear relationship that has emerged from the interplay between these various mechanisms, each of which is simple on its own but becomes complicated in combination.

The problem of social utility notwithstanding, we can next consider variables included across all the models, starting first with education. In the models where wage is not explicitly included (2, 3, 5), education behaves similarly to what was observed in tables 5.4 - 5.5, in a parabolic fashion with there being a significant difference in range for city B. In models where wage is included (1, 4, 6), the non-city-interacted effects of education lose statistical significance. That wage largely subsumes the effects of education makes sense since in the current version of MAPES, the only function served by education is to determine fit for particular jobs, and not to provide any independent input into the migration decision.

The purpose of model 5 is mainly to show that the aggregate utility measure attains the appropriate negative sign and statistical significance. It is also useful in showing that considering the individual component utilities (viz. social and income) or their underlying measures (viz. network size and income) provide better fits to the data than the comparatively more blunt aggregate measure. This is based on the fact that AIC in model 5 is significantly worse than those in any of models 1-4.

As far as differential effects arising from specific cities, it does not appear to be the case that being from or being in a city has some direct effect. Rather, the effect of these variables is to modulate those of other variables, namely education, wage and each other. This is good since a city in MAPES has no innate meaning beyond the characteristics that comprise it. While not all the terms are statistically significant, the signs on the effects are consistent across models. The gist of the story is that being from city B makes one more likely to migrate, but those most likely to migrate are those who are both from city B and are currently in city B. As discussed earlier and consistent with tables 5.4 - 5.5, being from city B makes a difference in the effect of education.

Finally, model 6 is identical to model 1 except that it additionally interacts wage and city of residence each with city of origin. The rationale for doing is that we saw in the time invariant regressions (tables 5.4 and 5.5) that persons from different cities experience the effects of education differently. Therefore, it is reasonable to expect the same of average wage. As it turns out, dollar for dollar, persons from B are less affected by increases in wage. This makes sense since there is an ever-present pressure to migrate out of B into A. So each extra dollar in wages is less incentivizing to stay put for the person from B compared to the one in A.

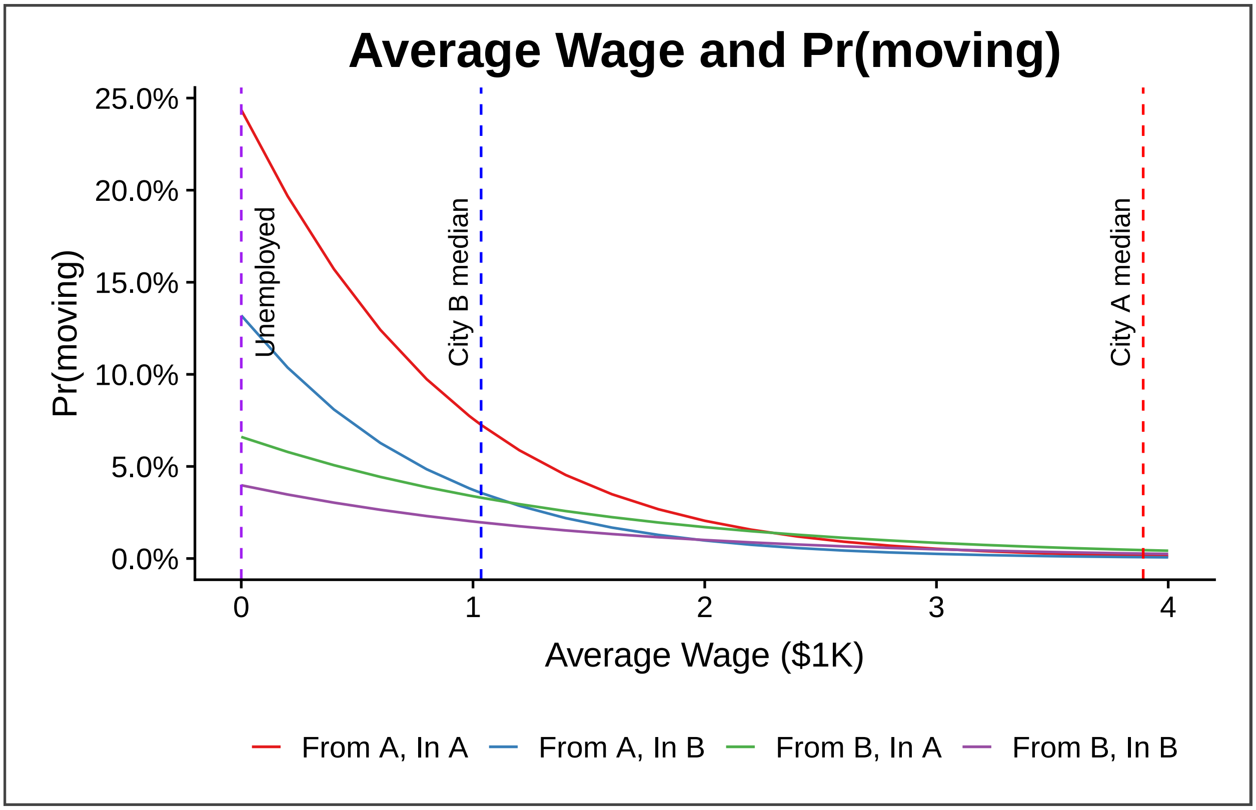


Figure 5.21: BEG: Average Wage and Pr(migrate), Model 6

Figure 5.21 shows the response curve for the probability of moving as a function of wage using model 6 of table 5.7, holding other variables at their respective median values. This figure at last enables us to see why, contrary to initial expectations, persons from city A move more often, but less permanently. The line of reasoning is a bit lengthy but it starts with the fact that people are most likely to move when they are unemployed because that zeroes out their income utility. If they are unemployed, they are on the purple vertical line in figure 5.21, where persons from city A are several times more likely to migrate than their counterparts from B, in the same city. Specifically, in city A, a person from A in A has a probability of moving of 24% while someone from B in B has a corresponding value of 4%, roughly 6 times more.

Once the person moves, if they remain unemployed, the respective probabilities of moving become 13% (from A in B) and 7% (from B in A), not a several fold difference but significant nonetheless. If they do become employed at the local median wage, then the probabilities become 4% and 0.1%, again an order of magnitude difference. What all this means is that as we observed near the start of this section, persons from A are much more likely to migrate in general, which means that they are also more likely to return to A after their sojourn in B. Persons from B are less likely to migrate, which is not to say that their probability of doing so is insignificant. However, due to the conditions demonstrated here, their moves tend to be stickier.

Beyond the substantive value of investigating the BEG scenario, the main purpose of this chapter is to demonstrate what working with a MAPES generated dataset is like. We as modelers define everything that goes on in the data generation step, so the analytical process becomes as much about validating model expectations as it is about substantive discovery. Through careful introspection of model mechanics, we are mostly successfully in predicting what should happen and validating through application of data analytical techniques. Still, even in the BEG scenario, which is the simplest of the three scenarios investigated in this dissertation, there are instances where the model gives rise to unexpected behavior. This is seen in section 5.2 in the ideological clustering issue, and in this section, with respect to social utility and local network size. It is a matter of perspective whether one views these as flaws in the model or as instances of subtantive learning. I lean towards the latter interpretation since such unexpected behavior sheds light on parameters and characteristics that were hitherto not given deep thought.

Within the overall narrative of this dissertation, this chapter on BEG can be seen as equal parts a tutorial on MAPES fundamentals as well as a substantive investigation of a standard migration situation. The subsequent chapters on the trilateral-equal and UMBC scenarios will build on the processes covered here to explore more complicated situations.

CHAPTER 6

Scenario 2: Trilateral Equal

In this chapter, I examine a scenario that I call Trilateral Equal (TE) comprised of three city-countries that start the simulation identical in nearly every regard. Each starts with the exact same citizenry, pool of jobs, and pool of legislators. By extension, also identical are the distributions of ideological ideal points, education, and incomes, etc. across the three locales. At the outset, the only difference between the city-countries is the relationships that exist between the different agents. In a graph context, the nodes are all the same but the edges are not. Those are randomly generated as part of model initialization. So while person X in city A may have job I, in city B the same job I may be held by person Y rather than that city's person X. While not terribly interesting (since the cities are identical), table 6.1 and figure 6.1 show scenario specific parameters and agent characteritics.

The fact that there are identical agents across the 3 locales means that we need to be as clear as possible about notation. When necessary, the following convention is used to refer to specific agents: $type^{city}_{id}$ where *type* is person, job, or legislator; *id* is an index into some per-city ordering for that type; and city is the initial city that agent started in. So job^A_i is the job in city A that is identical in characteristics to job^B_i and job^C_i in their respective cities. The relationship between job^A_i and job^A_j is specified in the context in which they are talked about (i.e. i and j can be ordered based on characteristics of the underlying agent such as education or wage, or just be randomly assigned IDs). Finally, as with the BEG scenario, cities and the countries they belong to are synonymous since there is only 1 city per country; therefore they are referred to interchangeably.

Table 6.1 shows the main simulation parameters for TE. The three cities are arranged in an equilateral triangle, at a distance of 100 units from each other. Each city starts with 100

City	City A	City B	City C
Country	Country A	Country B	Country C
Latitude	50	0	-50
Longitude	0	86.6	0
Living Cost (Monthly)	800	800	800
Population	100	100	100
Pop. Education Mean	12	12	12
Pop. Education Std. Dev.	3	3	3
Pop. Ideology Mean	50	50	50
Pop. Ideology Std. Dev.	15	15	15
# Jobs	70	70	70
Job Education Mean	12	12	12
Job Education Std. Dev.	3	3	3
Wage Mean	2000	2000	2000
Wage Std. Dev.	300	300	300
Corr. of Edu. & Wage	0.8	0.8	0.8
# Leg. Seats	5	5	5
# Leg. Candidates	20	20	20
Leg. Ideology Mean	50	50	50
Leg. Ideology Std. Dev.	15	15	15

Table 6.1: TE: Locale Parameters

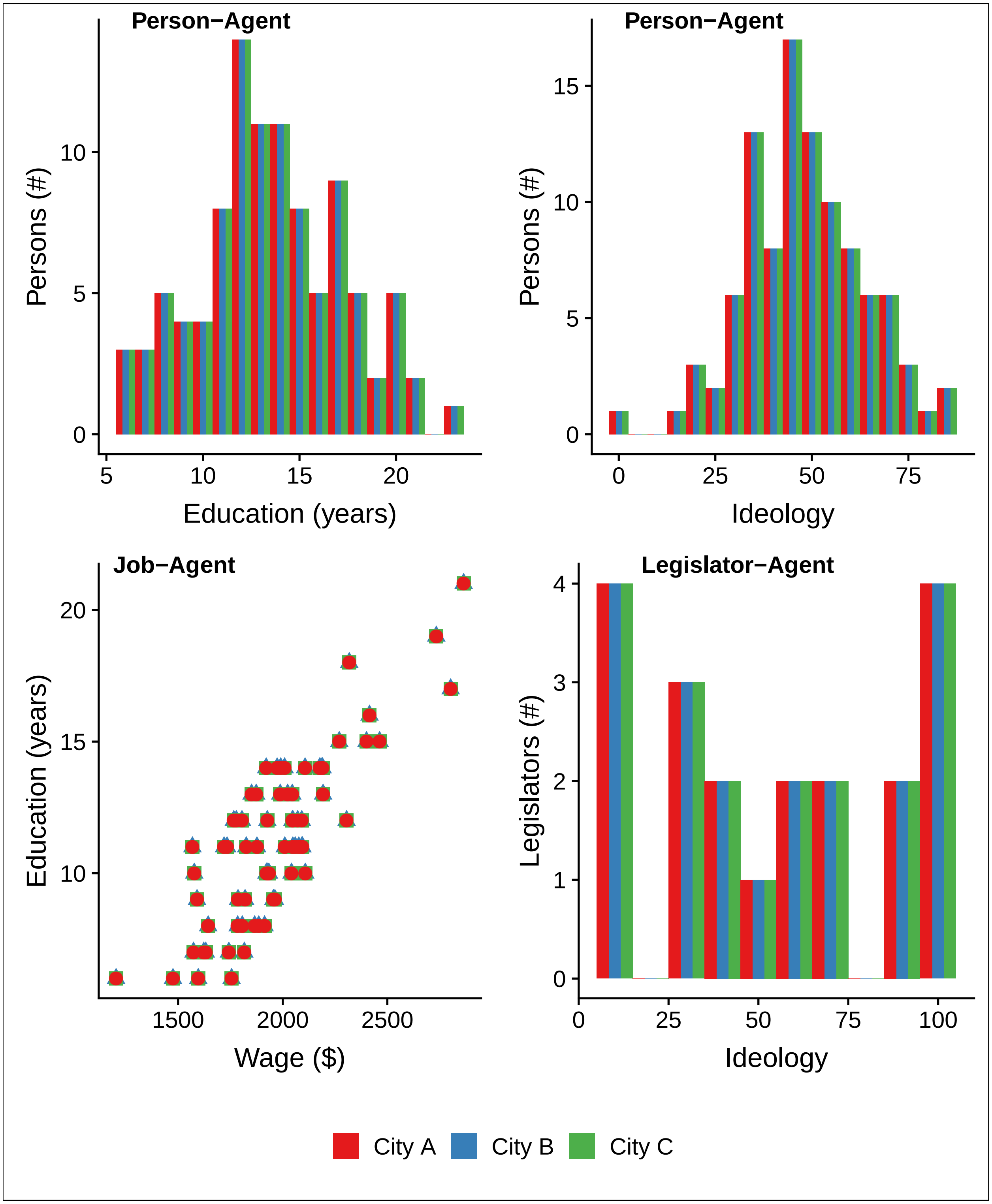


Figure 6.1: TE: Agent Characteristics

persons who have education normally distributed with mean 12 years and standard deviation 3, and ideology distributed $N(\mu = 50, \sigma = 15)$ on the 0 to 100 (open to closed) scale. Each city has 70 jobs which are distributed multivariate normal on the dimensions of education and wage, with education distributed $N(\mu = 12, \sigma = 3)$, wage distributed $N(\mu = 2000, \sigma = 300)$ and correlation between the two $\rho = 0.8$. Each city also has 20 political candidates who have ideology distributed identical to the population they represent. In each election, they vie for the 5 legislative seats available in each city.

6.1 Population

The main analytic wrinkle of this scenario is that it is uninformative to use a priori city/country names (e.g. city A, country B) as a unit of analysis. This is because unlike in BEG, the name given each city/country does not distinguish between them in terms of characteristics. By contrast, in BEG country A was always the receiving country and country B was always the sending one. This fact is seen immediately in plots of the populations of the three locales.

It should come as no surprise that the plots in figure 6.2 look very similar. Given that the three locales are identical, an event that occurs in one is equally likely to occur in another. So the consequences of such events should average out across multiple runs of the model, resulting in similar aggregate patterns across the countries such as patterns of population change. In a sense, the uniformity of figures 6.2 is an indication that the model functions properly - it would be concerning if they were not similar. However, it also doesn't take a social scientist to predict that three identical countries will average out identically. So, to get analytic value from this model, the locales must be classified in a meaningful way.

In BEG, the two countries are classified before the model is ever run, based on a theoretical understanding of sending and receiving countries, with the former being labor rich and capital poor; the latter, the opposite. By contrast, classification in TE is done ex post based on characteristics observed after running the model, namely immigration patterns. Like BEG, TE appears to develop sending and receiving countries but in addition, it develops neutral countries, which remain largely unaffected by migration and maintain constant

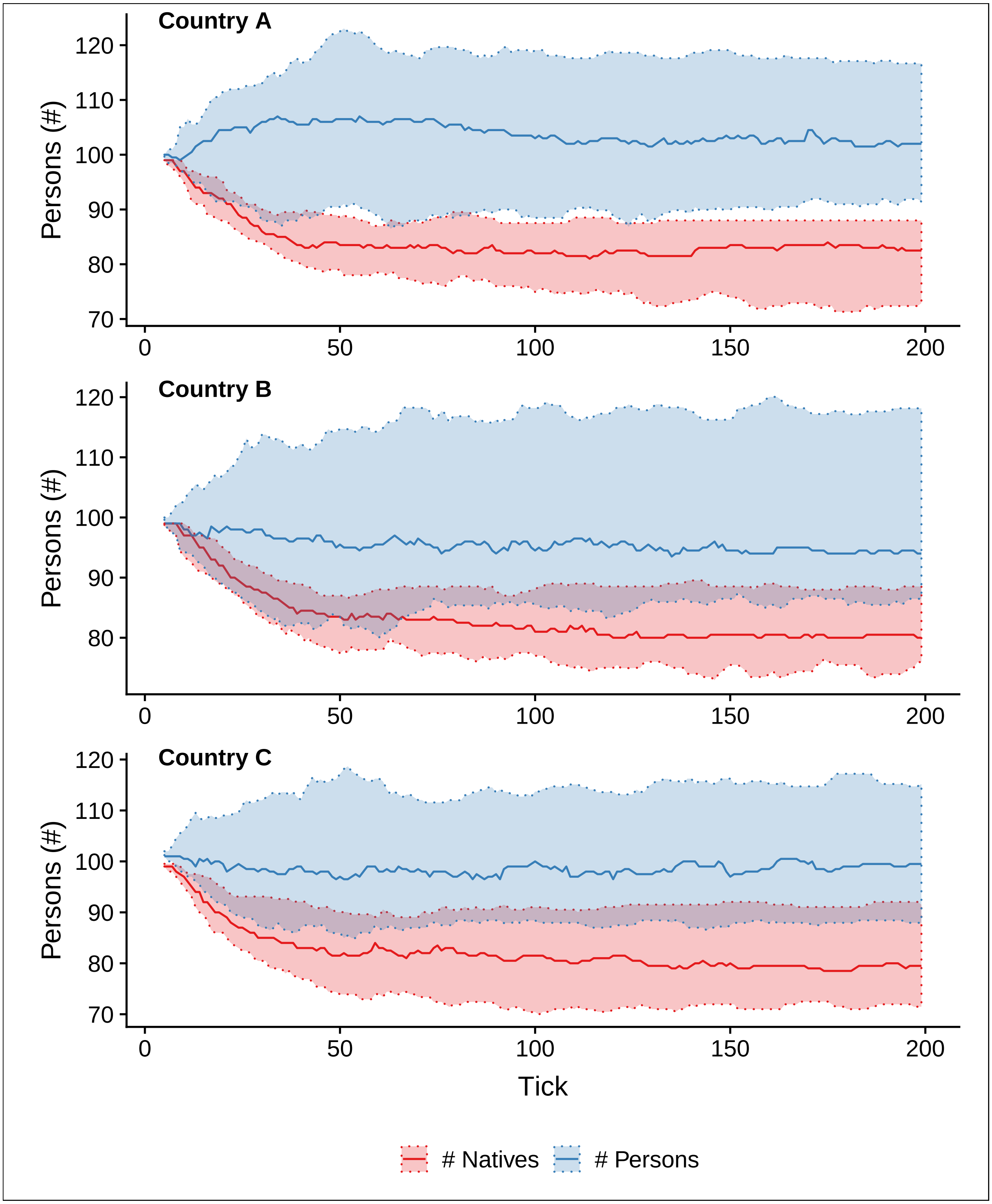


Figure 6.2: TE: Native vs. All Population by Country ID

populations in net. Therefore, the countries in TE are analyzed according to these three categories.

To apply this paradigm, there must exist well defined criteria by which a given country can be placed into one of the three categories. I use the following arbitrary but reasonable criteria to categorize each country. Receiving countries are defined as those countries which at simulation's end have a population greater than or equal to 102.5% their starting population. Sending countries are defined as those which end up with 97.5% or less. Countries falling into neither category are neutral. Figure 6.3 shows population plots by this new country type rather than country ID. Immediately, we see that these groupings result in much more distinct population plots.

All three country types appear to have a transition state up to about tick 50 and a steady state thereafter. Receiving countries stabilize around a total population of 108 and an immigrant population of 26. Sending countries reach a total population of approximately 92 and an immigrant population of 11. Neutral countries maintain their starting total population but also attain immigrant populations of around 18. Populations of natives maintain between 80 and 84 across all three categories, meaning that different type outcomes are primarily driven by receiving conditions (a particular country changes in receptivity relative to its peers), not sending ones (a country changes in likelihood of sending persons).

Figure 6.4, which compares immigrant populations across country types, underscores the presence of transition and steady states, and the relatively clean separation between categories. The dotted lines show population trajectories for individual countries (i.e. confidence intervals), while the solid lines are loess fits across all countries in a particular category. There is ambiguity distinguishing neutral countries from the others in this particular plot, but the distinction between sending and receiving countries is quite clear. This lends credence to the categorization scheme I use.

Figure 6.5 shows when people are migrating to result in the patterns observed in figures 6.3 - 6.4. This histogram shows immigration by country type, counted in 10-tick bins and averaged across all simulation-countries falling into each category (22 receiving, 26 sending,

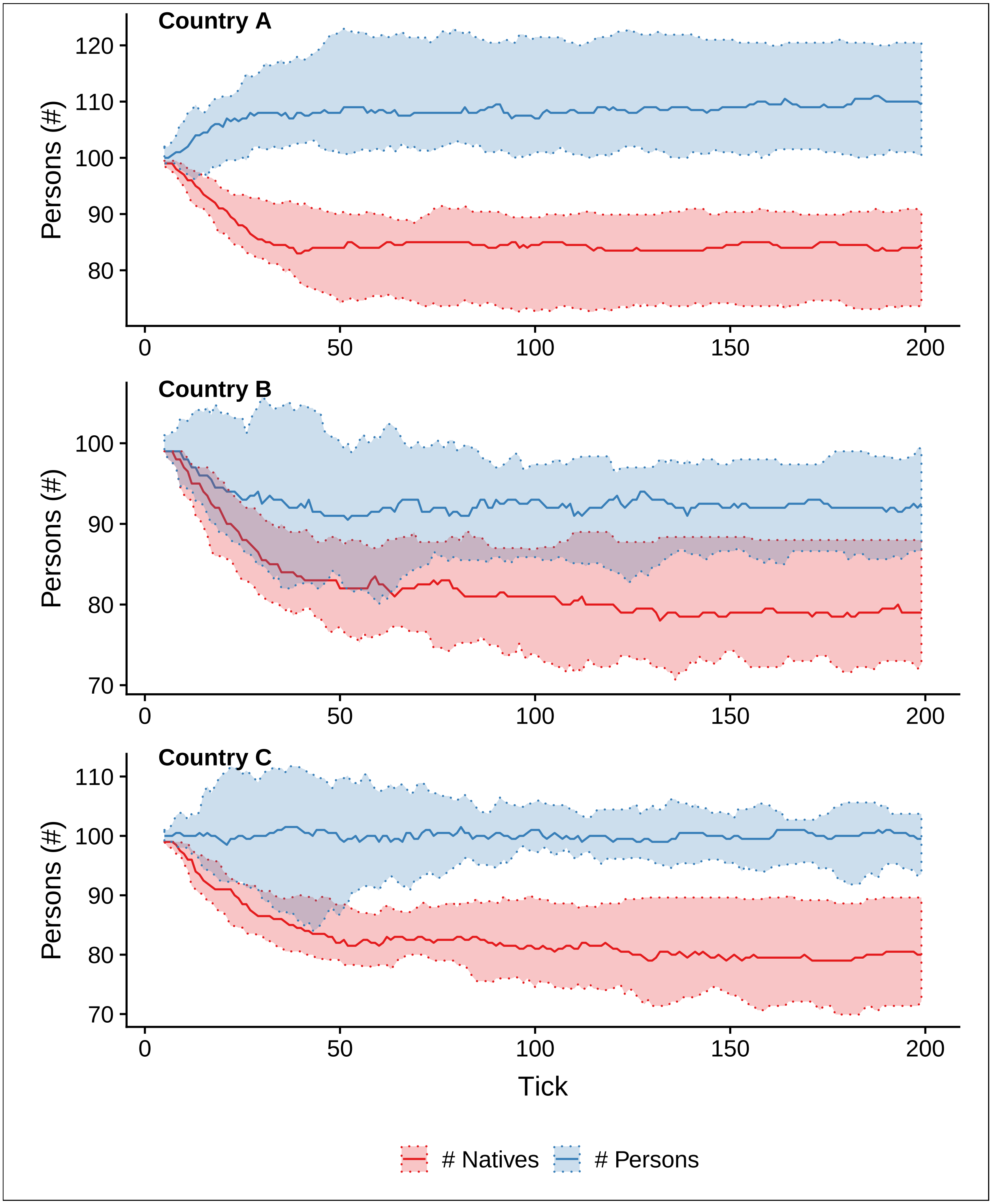


Figure 6.3: TE: Native vs. All Population by Country Type

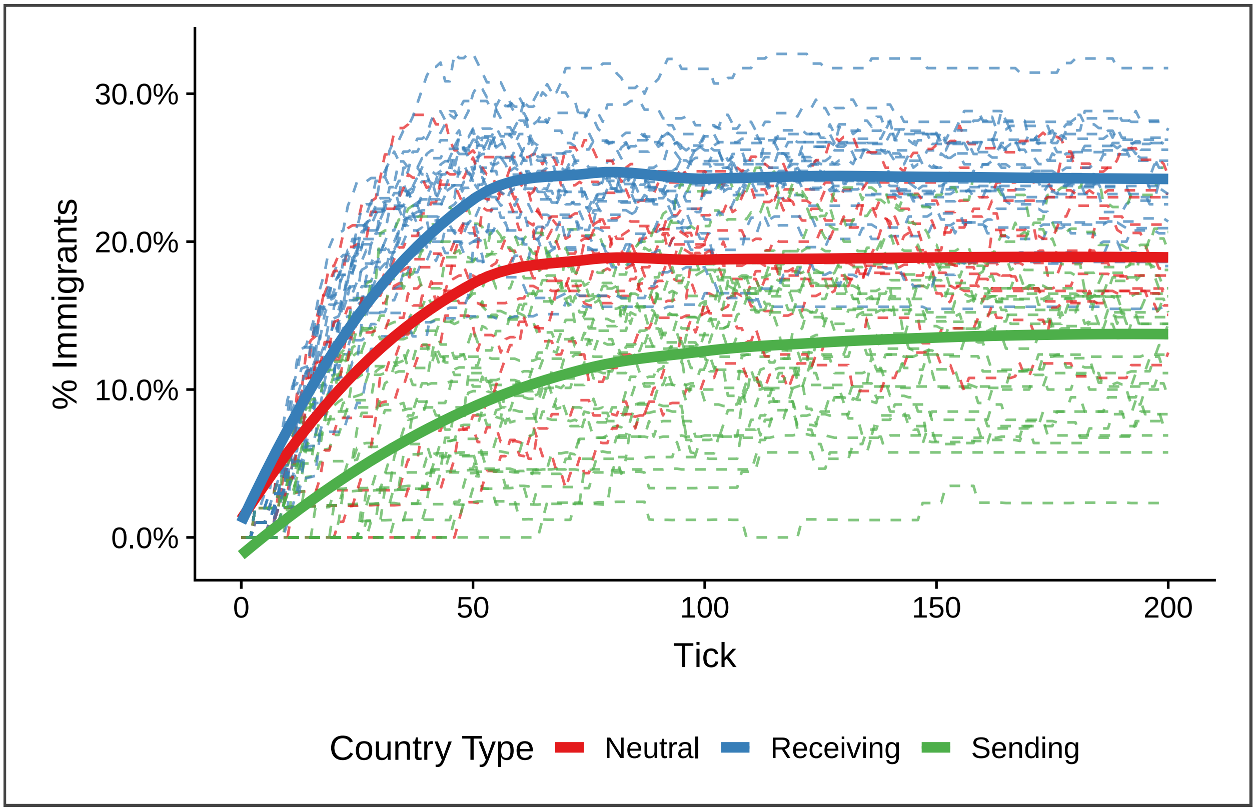


Figure 6.4: TE: Immigrant Population by Country Type

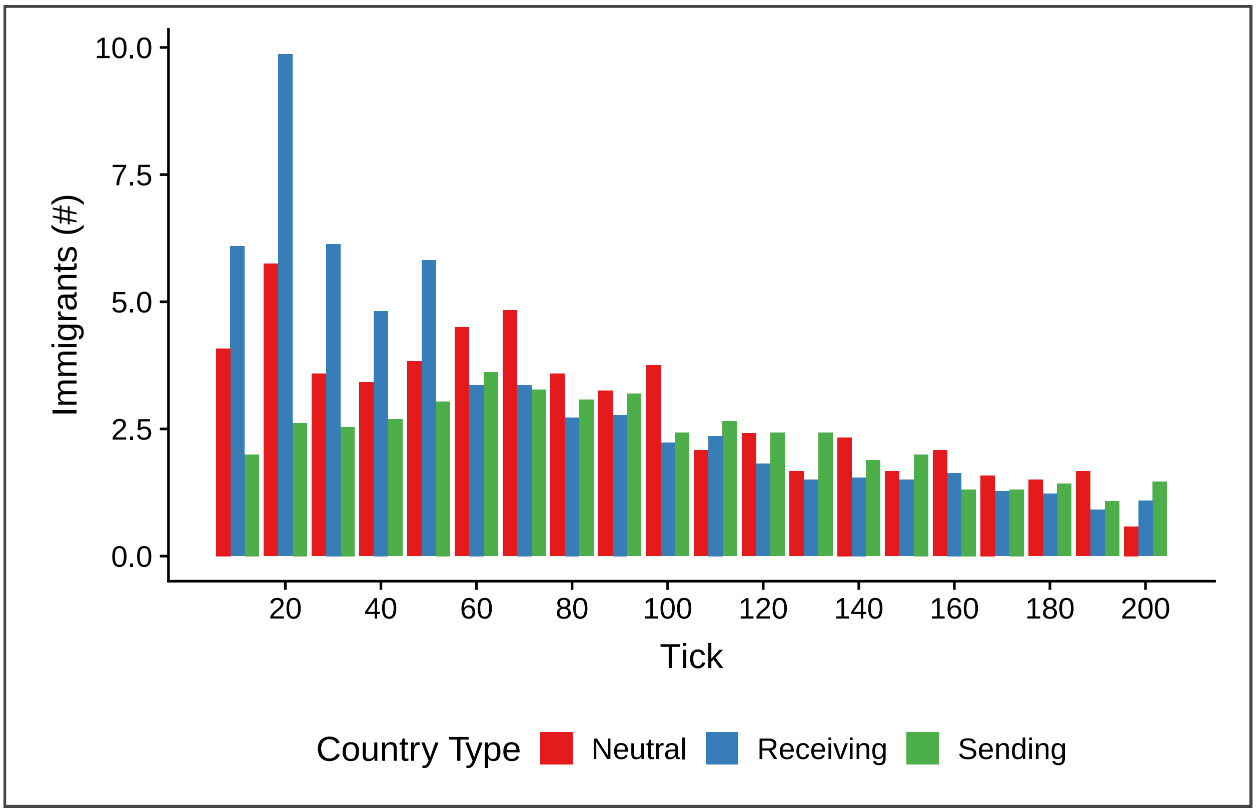


Figure 6.5: TE: Number of Immigrants per Tick, by Country Type

and 12 neutral). Receiving countries have a huge spike in immigration within the first 20 ticks, receiving over 16% of their starting populations in immigrants during that time. The influx stabilizes to the same level as neutral countries by around tick 100, halfway through the simulation. Neutral countries are second to receiving countries in gross immigration, receiving 10% of their starting population by tick 20, and 21% by tick 50. Sending countries remain relatively stable through the lifetime of the run, averaging about 2% of their starting population in inflow every 10 ticks. Though sending countries receive far less than the other types at the outset, by the halfway point the others fall off to around the same level. After the halfway point, immigration slows down to a trickle across all types, with less than 2% of starting populations moving into each category every 10 ticks.

Once we have classified different country types, it becomes apparent that we can also organize each simulation-system (a collection of simulation-countries) based on the numbers of each country type present. For example, run #1 might have 1 sending, 1 receiving and 1 neutral, but run #2 could have 3 neutrals. In a purely combinatoric sense, there are 3^3 possible type combinations but substantively we know that many combinations are simply not possible (e.g. 3 receiving countries, 2 sending & 1 neutral). In fact, only 4 combinations occur in the simulation data. Table 6.2 lists these along with the percentage of runs falling into each category.

	Number	Percent
1 Sending, 1 Receiving, 1 Neutral	6	30%
1 Sending, 2 Receiving	4	20%
2 Sending, 1 Receiving	8	40%
3 Neutral	2	10%

Table 6.2: TE: Number of Countries by Type, All Simulations

Table 6.2 is obviously interesting because it goes against the intuition that an uninformed observer might have that the even (i.e. 3 neutral) outcome would be the most common one. After all, if all three countries are identical, shouldn't they tend to end as they started, on an even footing? For those acquainted with the immigration literature, however, the answer obviously is that path dependence renders uneven outcomes quite likely, as is borne out by our model. The question now is what causes this path dependence. Theoretically we think that micro variations compound into macro ones but showing how this happens is a more difficult task than merely hypothesizing. In the next section (section 6.2), I confront this question from a political angle, investigating how policy changes may contribute to this outcome.

6.2 Politics

Politics in TE functions identically as in BEG (the previous chapter) with one exception. Persons in TE discriminate who they form social ties with based on ideology so that improbable liaisions between persons with radically differing world views no longer occur. This is accomplished by modifying the model such that for each individual, only persons with ideology similar to their own (viz. within a threshold of plus or minus 5) are considered for social ties. This leads to the formation of distinct ideological groups, alleviating the clumping problem that, in BEG, unrealistically gives rise to a single ideological mass encompassing the entire polity.

With this mechanical alteration in mind, I begin my examination of politics in TE by looking at the ideological development of individuals. As with figure 5.8, figures 6.6 - 6.8 show the progression of individual preferences across the simulation, clustered by ideology at simulation's end. Clustering is done using the R library, "Ckmeans.1d.dp", which calculates the optimal number of clusters to minimize intra-group distance. Note that the colors in these figures are randomly assigned to each cluster and hold no substantive meaning across plots.

Visually examining these plots gives the impression that ideological outcomes in the trilateral scenario are complicated at best and unpredictable at worst, unlike the straightforward patterns we saw in the BEG scenario. By "ideological outcome," I mean the general appearance of the plot of ideology against time as in figures 6.6 - 6.8; a definition which is vague but sufficient for the time being. While this initial reaction turns out to be false, there

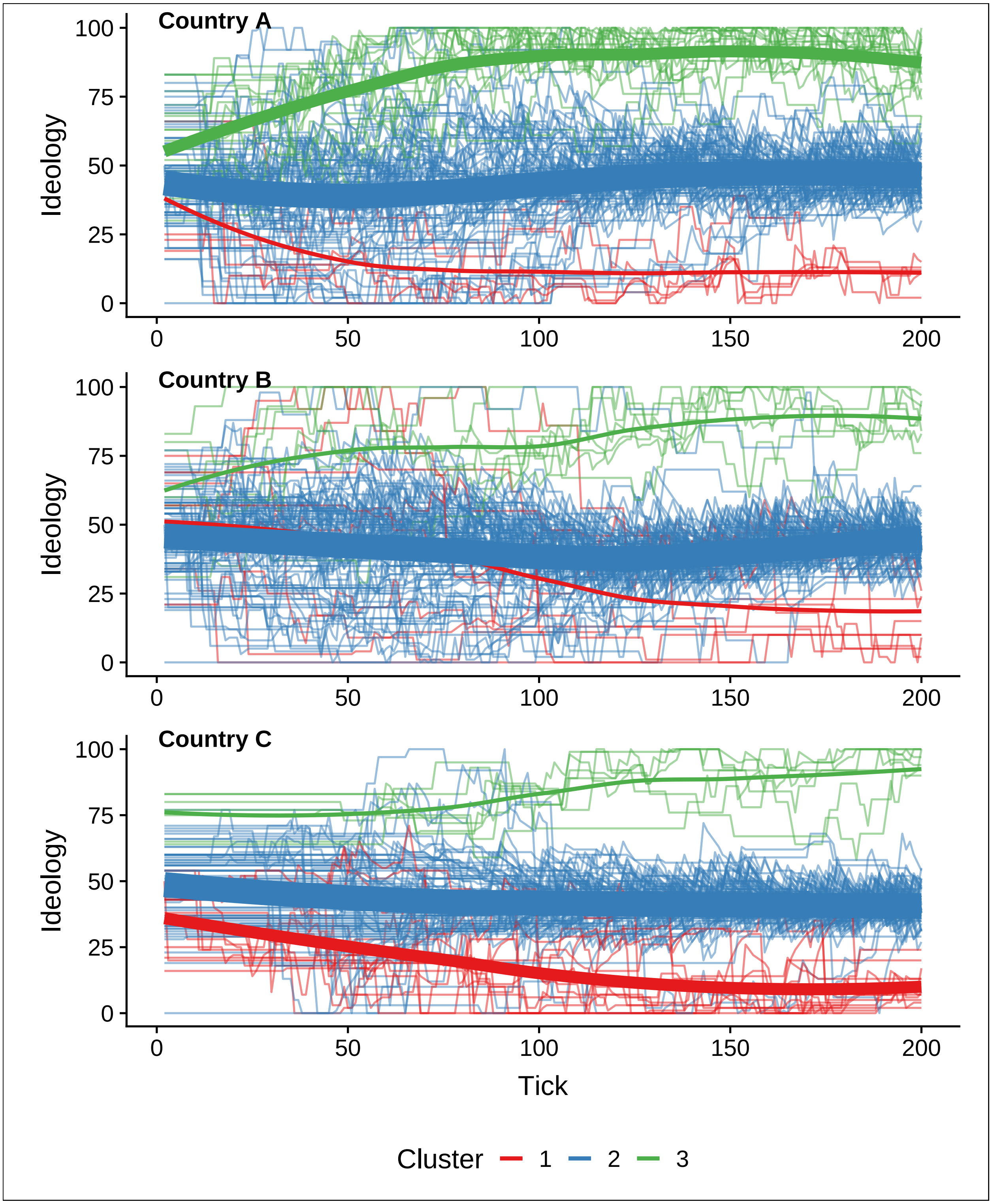


Figure 6.6: TE: Ideological outcomes, individual and clustered, simulation 1

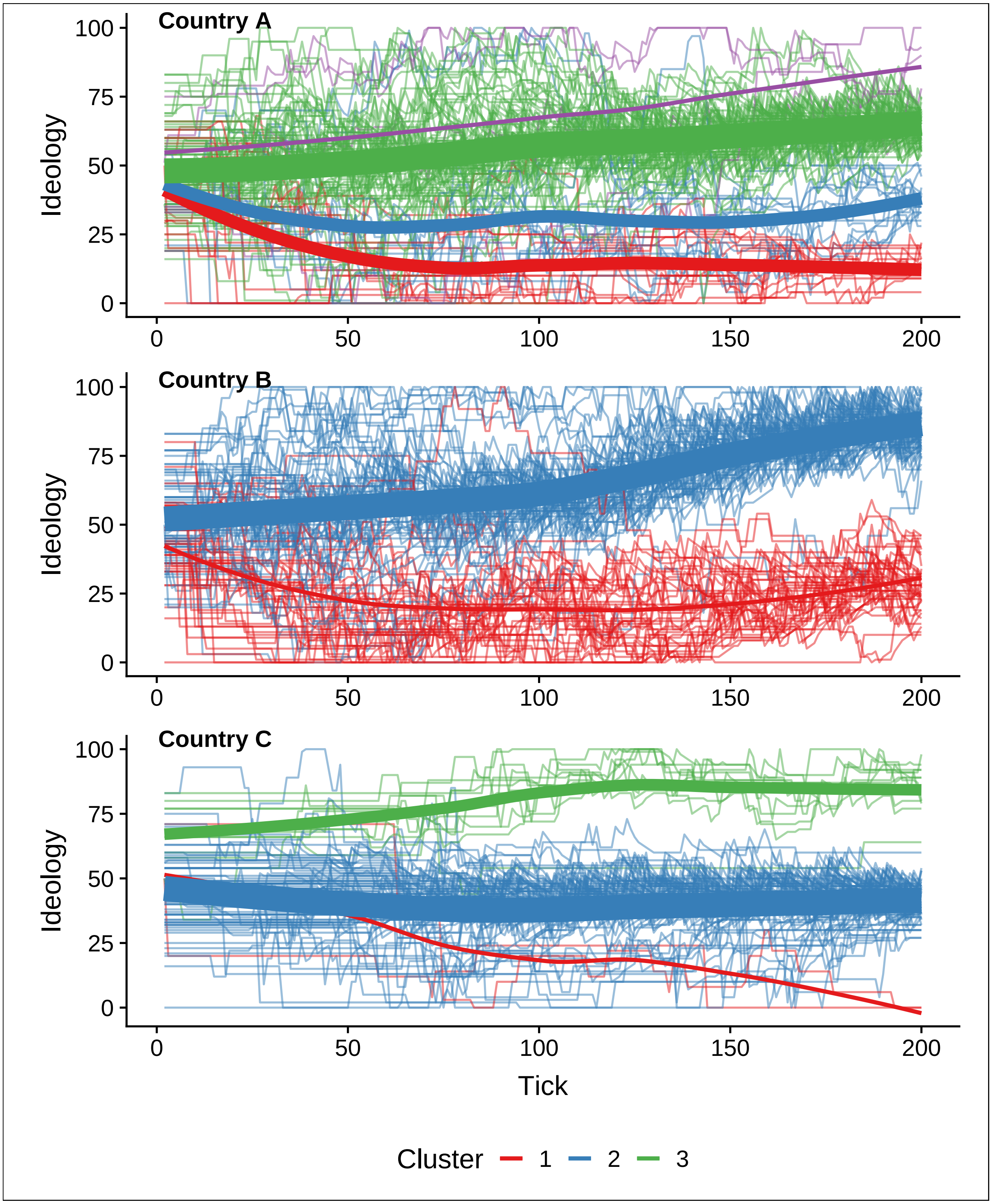


Figure 6.7: TE: Ideological outcomes, individual and clustered, simulation 2

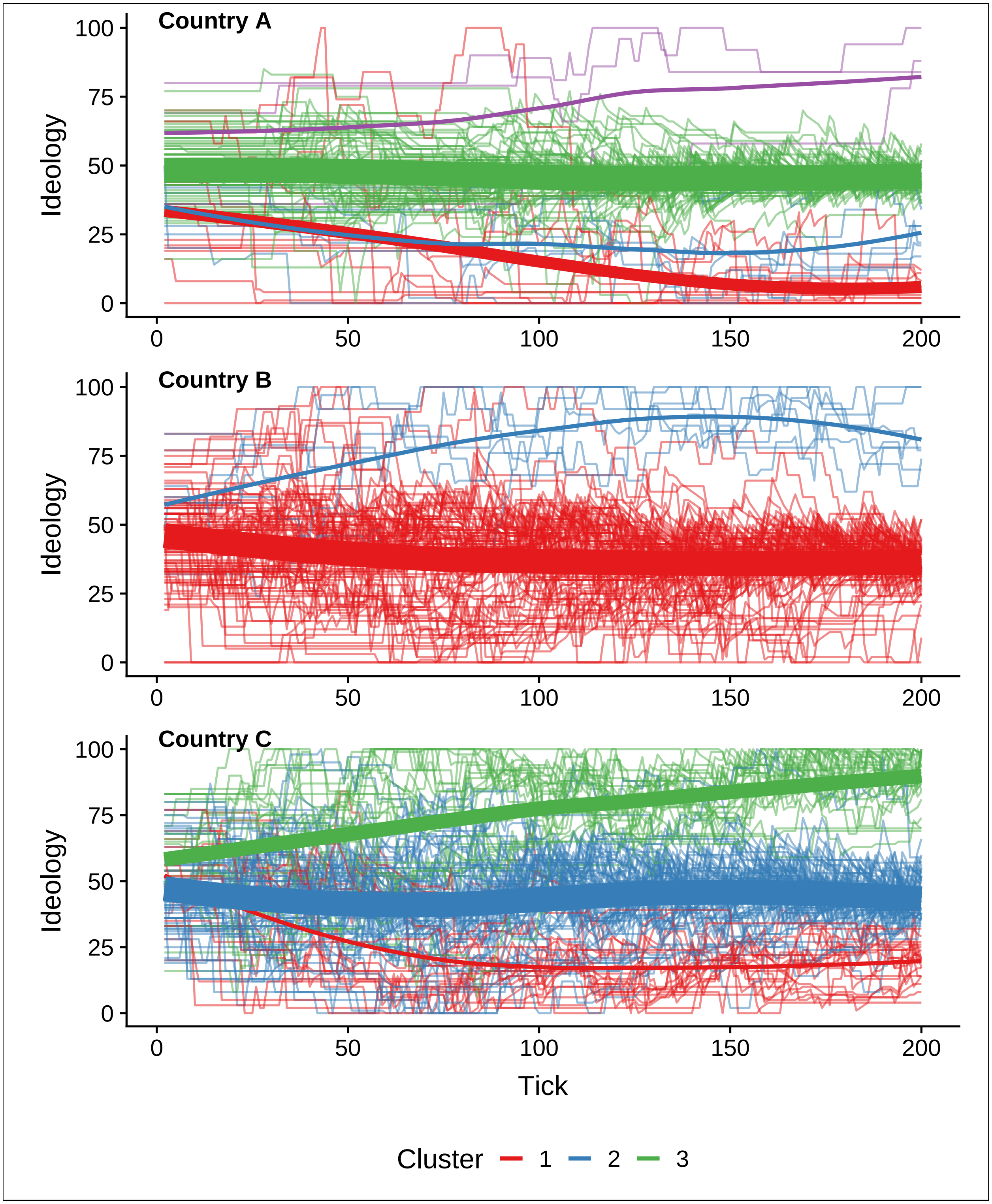


Figure 6.8: TE: Ideological outcomes, individual and clustered, simulation 3

are many valid reasons to think so at this time.

First, the ideological outcome for any particular country appears stochastic across simulations. In one simulation, country A might develop 3 ideological clusters, with one being larger than the other, while in another simulation it might only develop 2 evenly sized clusters on either end of the ideological spectrum. The same country that, in one simulation, devolves into extreme polarization around the issue of migration could just as well have become multipartite and ideologically diverse in another.

Further, the outcome displayed by a particular country in one simulation is not just assumed by another country in a different simulation. That it is not is counter-intuitive because the cities are literally identical at the outset (since that is the premise of TE). We can visually confirm this by observing that the distribution of starting locations (i.e. the y-value of the left-most point of each line) is identical across the 3 cities. Given this fact, it is not unreasonable to hypothesize that the outcomes displayed by country A in one simulation may be displayed instead by country B or C in another. In essence, countries would be exchanging roles. This, of course, assumes that there is such a thing as a set role for countries to play in this trilateral scenario, and that there is some logic to their appearance. For example, one might hypothesize that any trilateral system in which there is a polarized receiving country must also have a polarized sending one. However, as the randomness of figures 6.6 - 6.8 makes fairly clear, these kinds of expectations will not be fulfilled. In short, similar initial conditions do not lead to similar ideological outcomes when considering aggregate patterns in individual preference.

Another important point is that the same country type (viz. sending, receiving, neutral) often leads to very different patterns of ideological development, even within the same simulation. For example, in figure 6.7, both A and B are receiving countries (this is not indicated in the figure but is known from the underlying data), but while the majority in A conservatizes gradually by a small amount, the majority in B conservatizes very strongly by halfway through the simulation. On the other hand, there are also two sending countries in figure 6.8 (A and B). While B develops only 2 clusters (a large centrist cluster and smaller conservative one) the other (A) develops 4 clusters evenly distributed across the ideological spectrum. So country type, as well as initial conditions, does not appear to be a great predictor of a country's ideological outcome, at least in the vague sense in which I use that phrase.

The consequence of the previous three points is that for the TE scenario of three identical countries sharing migration flows, there seems to be little that is politically preordained. One simulation might have all three countries achieving an even distribution of 3 clusters, while another simulation might have 1 country with 2 ideological clusters, and yet another simulation may have 2 countries with 4 uneven clusters. This stands in stark contrast to the situation in BEG, where the receiving and sending countries have strict political roles which derive from their predefined position in the migration system.

The obvious caveat to all this is that our analyses may change once we start categorizing by ex post country type rather than arbitrary country ID. As I showed in the previous section, meaningful patterns may emerge once these are considered. The question then arises, what measures shall we examine for these potential patterns? Certainly we would like more precision and measurability than the vague notion of appearance-cum-ideological outcome used so far. I proceed by considering: number of clusters, spread in cluster size, spread in cluster means, and overall mean ideology. They are good candidates because they are easily quantified and substantively meaningful, and in combination, they quantify the fragmentation, polarization, and general direction of ideological preferences for each country.

	Num. Clusters			
Country Type	1	2	3	4
Neutral	0 (0%)	3 (25%)	8 (67%)	1 (8%)
Receiving	2 (9%)	6 (27%)	12 (55%)	2 (9%)
Sending	1 (4%)	2 (8%)	21 (81%)	2 (8%)

Table 6.3: TE: Cluster Counts by Country Type

Table 6.3 presents the number of clusters across all country-agents by country type along with corresponding percentages by country type (i.e. summing to 1 by row). The majority of countries tend to have 3 categories of ideology. There are countries which develop 1, 2 or

4 clusters, but they are the minority. In table 6.3, the standout fact is that sending countries overwhelmingly tend to have 3 clusters (81%), while both receiving and neutral countries show far smaller percentages of the same (67% and 55%). We can easily test whether this apparent relationship between country type and number of clusters holds up in a regression setting.

	OLS	Logistic	
	(1)	(2)	(3)
(Intercept)	2.833*** (0.190)	0.693 (0.612)	0.357 (0.348)
Receiving	−0.197 (0.236)	−0.511 (0.747)	
Sending	0.090 (0.229)	0.742 (0.789)	1.078* (0.607)
N	60	60	60
Adjusted R^2	0.005		
Log Likelihood		−35.525	−35.763
AIC		77.049	75.526

*p < .1; **p < .05; ***p < .01

Table 6.4: TE: Median Legislature Ideology

The leftmost results in table 6.4 are for a simple OLS of cluster count on country type. We know this model is not going to be a great fit since the outcome variable, cluster count, is not monotonically increasing. The middle column changes the model to a logistic regression with the outcome variable indicating whether a country attained 3 clusters or not. The final column does the same as the previous with the sole predictor changing from country type to a dummy variable indicating whether a country is a sending country. Only in the third and final model does country type attain weak statistical significance, at the 0.1 level. However, we do see across all three models that sending countries tend to have higher incidence of 3 ideological clusters, as we expect based on table 6.3. Substantively, this suggests that receiving and neutral countries are more liable to become polarized by migration than sending ones.

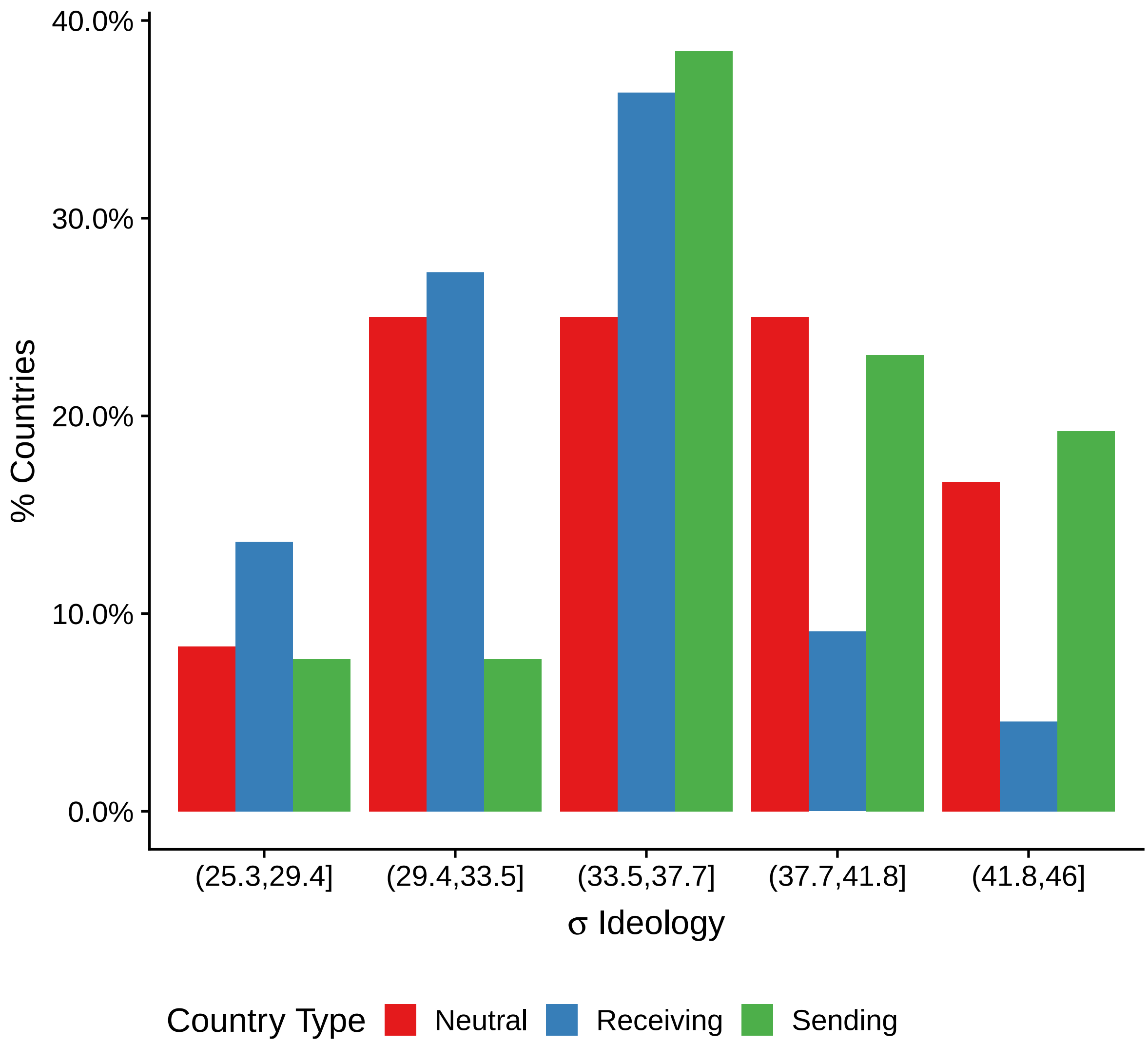
	(35.1,40.7]	(40.7,46.2]	(46.2,51.8]	(51.8,57.3]	(57.3,62.9]
Neutral	3 (25%)	3 (25%)	4 (33%)	2 (17%)	0 (0%)
Receiving	0 (0%)	3 (14%)	11 (50%)	4 (18%)	4 (18%)
Sending	12 (46%)	5 (19%)	5 (19%)	4 (15%)	0 (0%)

Table 6.5: TE: Mean Individual Ideology vs. Country Type

Further, since country type appears to have an effect on number of clusters, perhaps it also has an effect on other aspects of ideological outcome. Figure 6.9 shows the relationship between the spread in cluster centers and country type; figure 6.10 shows the spread of individual ideology; and figure 6.11 shows mean ideology. From figure 6.9, we see that clusters in receiving countries tend to be less spread out than those of sending countries, while neutral countries have an almost uniform distribution across the range of observed spread (viz. standard deviation) values. Figure 6.10 shows that ideology among individuals in sending countries tends to be marginally less spread out than among those in receiving and neutral countries. Finally, figure 6.11 shows that mean sentiments are clearly more open in sending countries than in neutral and receiving, in that order.

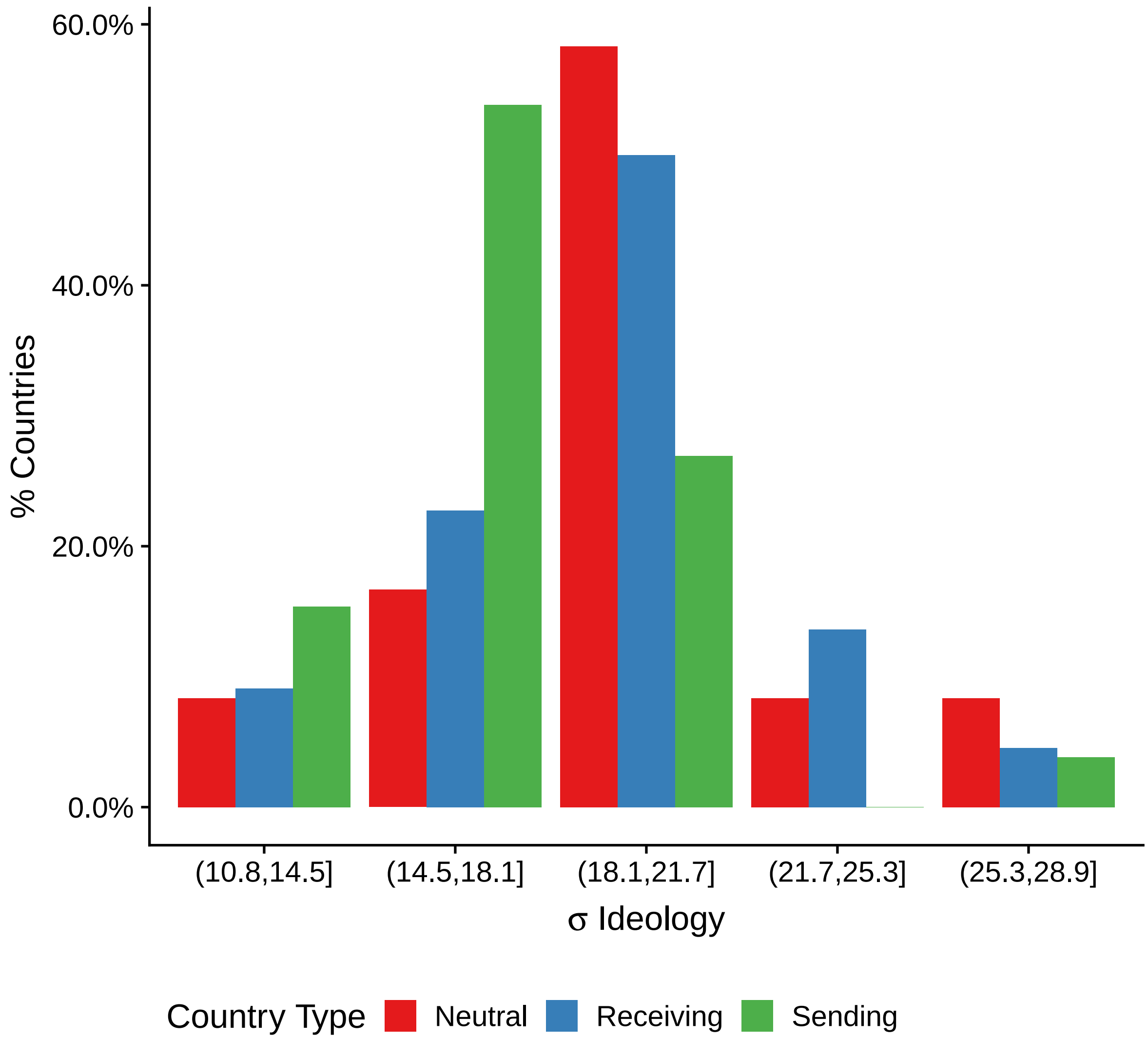
Tables 6.4-6.5 and figures 6.9-6.11 show that contrary to our expectations from the initial naïve graphical approach, there are in fact noticeable relationships between country type and ideological outcomes. Sending countries tend to be more open in their attitudes about immigration on average but there is more diversity of opinion. Receiving countries have tighter consensus on a prefererence for more closed policies. Neutral ones fall between.

Next, let us see how these patterns impact politics at the legislative level, where legislators reflect the sentiments of their constituents and compete for seats. Based on the evidence so far, country type will almost certainly be a primary determinant of outcomes. Accordingly, figure 6.12 shows the progression of median legislative ideologies separated by country type. The dotted lines are chamber medians on a per simulation basis, with color indicating country type. There are 60 such lines since there are 20 simulations used for these analyses with 3



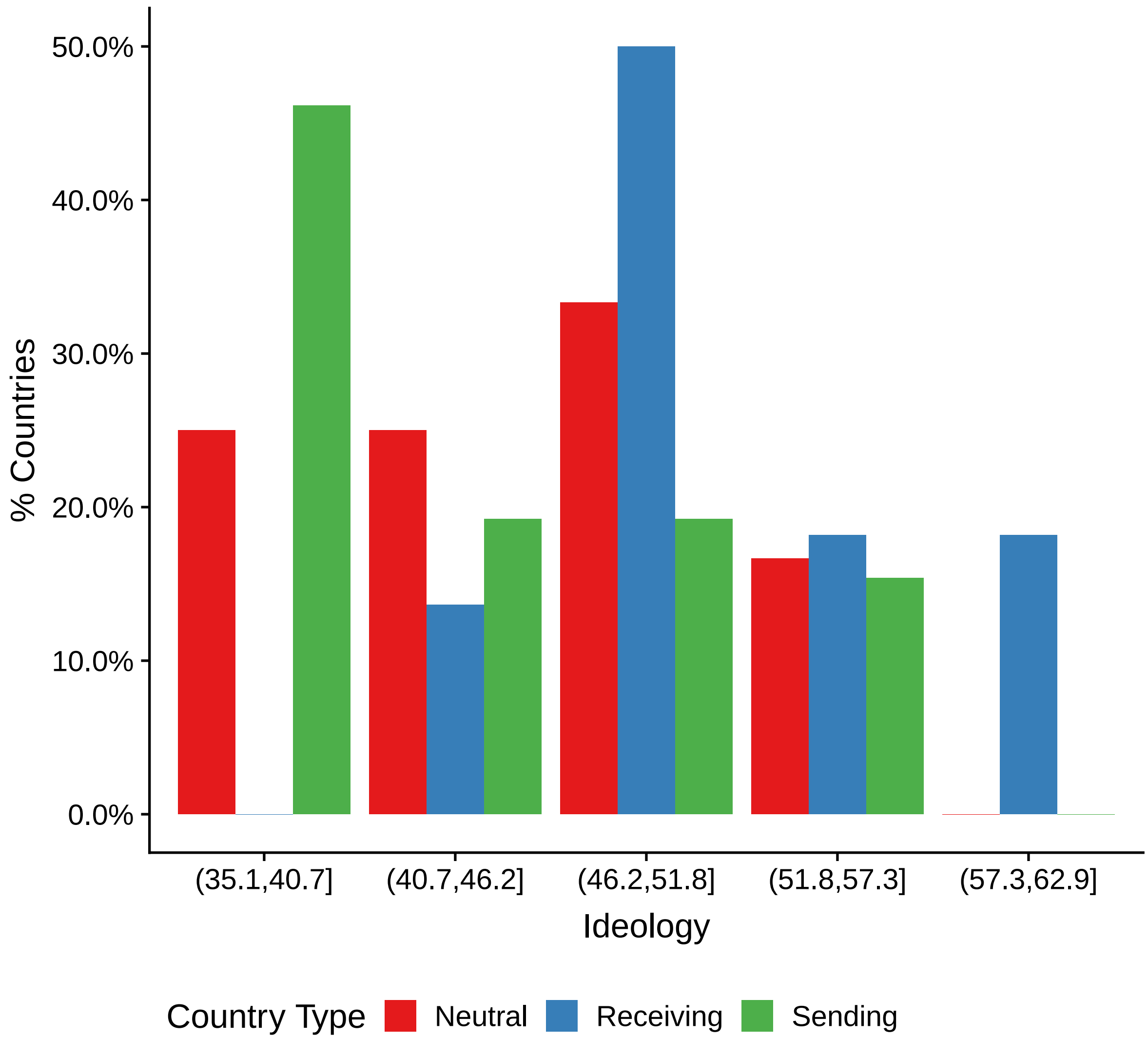
	(25.3,29.4]	(29.4,33.5]	(33.5,37.7]	(37.7,41.8]	(41.8,46]
Neutral	1 (8%)	3 (25%)	3 (25%)	3 (25%)	2 (17%)
Receiving	3 (15%)	6 (30%)	8 (40%)	2 (10%)	1 (5%)
Sending	2 (8%)	2 (8%)	10 (40%)	6 (24%)	5 (20%)

Figure 6.9: TE: Spread of Cluster Centers by Country Type



	(10.8,14.5]	(14.5,18.1]	(18.1,21.7]	(21.7,25.3]	(25.3,28.9]
Neutral	1 (8%)	2 (17%)	7 (58%)	1 (8%)	1 (8%)
Receiving	2 (9%)	5 (23%)	11 (50%)	3 (14%)	1 (5%)
Sending	4 (15%)	14 (54%)	7 (27%)	0 (0%)	1 (4%)

Figure 6.10: TE: Spread of Individual Ideologies by Country Type



	(35.1,40.7]	(40.7,46.2]	(46.2,51.8]	(51.8,57.3]	(57.3,62.9]
Neutral	3 (25%)	3 (25%)	4 (33%)	2 (17%)	0 (0%)
Receiving	0 (0%)	3 (14%)	11 (50%)	4 (18%)	4 (18%)
Sending	12 (46%)	5 (19%)	5 (19%)	4 (15%)	0 (0%)

Figure 6.11: TE: Distribution of Mean Individual Ideology by Country Type

countries apiece. The bold solid lines are the average of these lines for each country type. So, the dark green line indicates the average chamber median for all receiving countries, the red for all neutral countries, and the blue for all sending countries. This plot enables us to see not just the distribution of different ideological trajectories that legislatures take in these simulations but also the more universal patterns indicated by the bold lines.

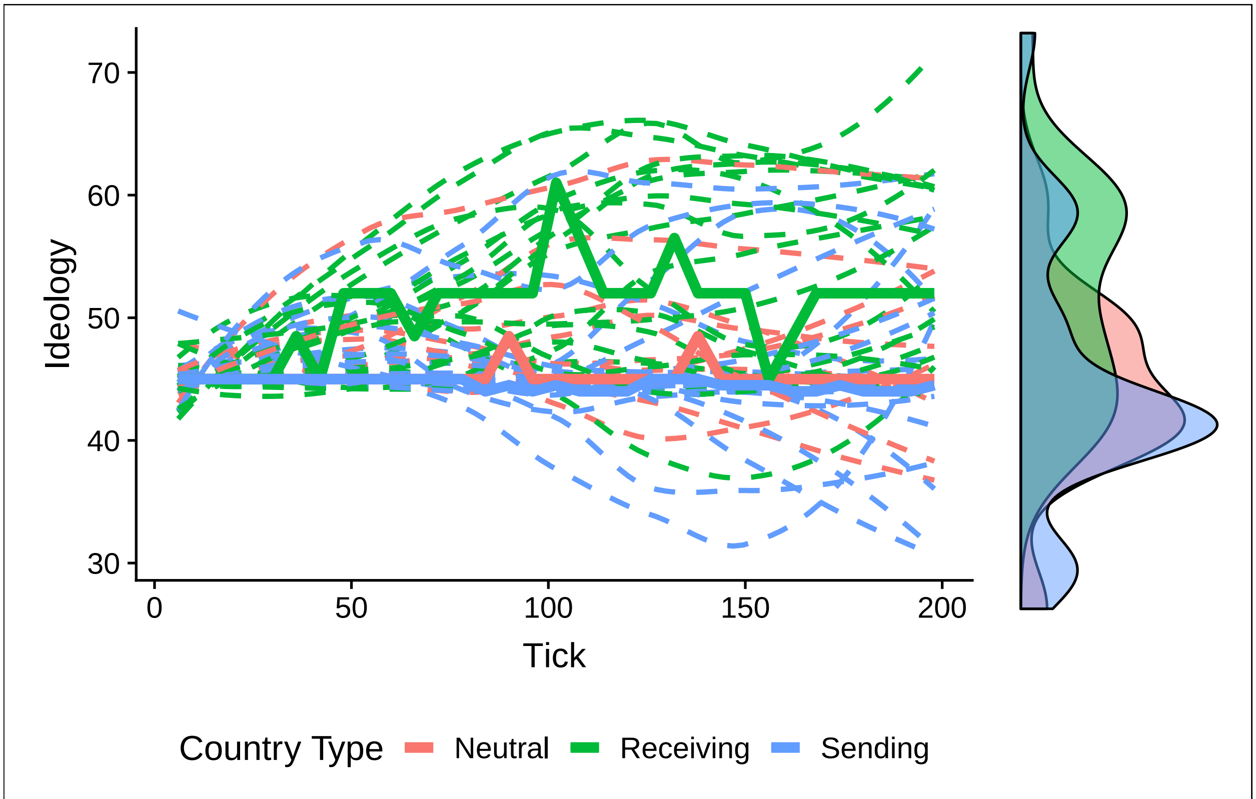


Figure 6.12: TE: Median Legislature Ideology over Time

Unsurprisingly, legislative medians correspond to the results we see at the individual level and the progression of ideology appears gradual and reasonable for all the country types. All three country types maintain for the first quarter of the simulation similar median ideologies of about 45. Receiving countries then begin to shift upwards (more restrictive) in their median ideology, attaining a steady state around 52, about 7 points more restrictive than their starting point. Sending and neutral countries share roughly similar behavior, maintaining chamber medians around their starting values. The vertical density plot on the right of figure 6.12 shows the final distribution of chamber ideological medians across all simulations. There is clearly a wide diversity of chamber medians possible for all three country types, but there are still peaks in the final density that mirror the findings from the examination of individual ideology.

All in all, our survey of individual and legislative ideological preferences in TE presents an interesting relationship between country type and ideological outcome. Speaking broadly, averaged across all our simulations, the three country types end up along an ideological continuum with sending countries tending to be least polarized and most open, receiving countries most conservative and polarized, and neutral countries in between. Based on our understanding of the model and the real phenomena involved, there are at least three possible explanations for the relationship, and the remainder of this section is spent considering a couple of these possibilities.

One possible explanation focuses on returnees – those who emigrate to another country, then return to their country of origin. The idea is that returnees are more sympathetic to the plight of migrants since they have experienced the challenges of that status. When they return to their home countries, they take those liberal (relative to their non-migrant countrymen) sentiments with them to the ballot box. A explanation focusing on returnees is intellectually appealing because of its parsimony: they immediately and directly affect legislative preferences in their country of origin through the vote. By contrast, immigrants in a foreign land are only able to affect ideological preferences of those who already possess the franchise in the short-term, though they may be able to eventually naturalize and affect politics. Crucially, what links returnees to the country type/ideology continuum is that sending countries (presumably) have more returnees than the other 2 country types since they have a bigger pool of persons who initially left and are therefore eligible to become returnees. If this lynchpin holds, then we would expect to see the greater number of returnees associated with more liberal ideological preferences.

Another possible explanation is that those countries which end up as receiving countries are those who, by chance, initially have more net positive interactions between pioneer migrants (i.e. very first) and natives, which leads the latter to prefer a more liberal policy. However, as this leads to more and more migration, an increase in negative encounters between natives and foreigners causes public opinion to swing the opposite direction. If this entire narrative were true, we would expect to see an S-shaped sentiment curve for receiving countries, with an initial liberalizing swing in the initial period of the simulation reversing and moving into conservative territory. We have already seen that no such swing occurs, that sentiments move gradually in a conservatising direction for receiving countries, so we can already dismiss the overall story. However, just because we do not observe the first part of that theory does not mean the latter part, of growing negative encounters and resultant sentiment, is also incorrect. On its surface, it does not seem impossible for there to be a direct relationship between immigrant population and negative foreigner-native interactions. This is worth considering, and I do so later in this section.

A weaker version of this explanation may be that rather than the story focusing on changes in receiving countries, the opposite set of phenomena instead occur in sending countries. This seems less probable because it implies that conservatization is the norm and that extraordinary events occur in sending and neutral countries to render them net neutral in ideology. Nonetheless, the possibility will be kept in mind. To summarize then, this set of explanations focus on net changes in foreigner-native interactions driven by changes in the number of foreign immigrants.

A third set of explanations is not explored here but presents ground for future work. These reverse the direction of causality to posit that it is not changes in population driving ideology but rather, changes in ideology that drive policy and therefore migration outcomes. In investigating this line of reasoning, the most urgent imperative is to determine proper causal direction. Only if it is established that ideology precedes changes in outcomes can we proceed to determine what is leading to the former. And, even if some causal pathways are found to operate in this direction, they do not preclude explanations operating in the opposite direction like the previous two; the relationship between ideology and migration outcomes can easily be multipartite and recursive, with ideology driving migration feeding back into ideology. Some thoughts in this direction are offered in chapter 8.

6.2.1 Returnees

Let us first investigate the impact of returnees on sentiment. The proposed causal chain is that returnees' experience as migrants colors their ideological preferences. This impacts their vote in their home countries once they return. Underpinning this mechanism are several model parameters, the first of which is that a person who migrates liberalizes for some number of ideology points, 30 in the case of TE. Also, each person's ideological position is monolithic: he does not have a separate preference for each country he's been in so that if he becomes more liberal after emigrating, he maintains those preferences even if he returns.

As explained earlier, the link between returnees and the country type / ideological outcome is that sending countries must have the most returnees, followed by neutral countries, then receiving. We can verify whether this relationship holds using figure 6.13, which shows that sending countries have the most returnees on average (as a percentage of total population), followed by neutrals, then receiving. There are very few returnees until about 30 ticks in, which makes sense since there is a timeout of 10 ticks in between moves. Returnees increase quickly until about 100 ticks in, then plateau, with a very slight upward trajectory.

Figure 6.14 adds an additional layer of complexity by showing how the data in figure 6.13 vary with respect to the percentage of the population that is of foreign origin. The three J-shaped curves are the relationship between returnee and foreigner populations at a particular period in the simulation, separated by country type. The dotted black lines traversing the curves show how the simulation periods line up across the country types; these are analogous to the X-axis in figure 6.13. In general, sending countries tend to have the most returnees as a portion of their population at any given time, followed by neutral, then receiving countries. Further, this gap grows over time, with the change in proportion of returnees in sending countries decelerating less relative to the other two types. This is seen in the increasingly negative slope of the time pseudo-axis (i.e. the dotted black lines). Additionally, this plot also reconfirms the earlier finding that foreigner populations stabilize at a certain point for each country type, seen in the fact that the country-type series of figure 6.14 each approach what appears to be a unique vertical asymptote.

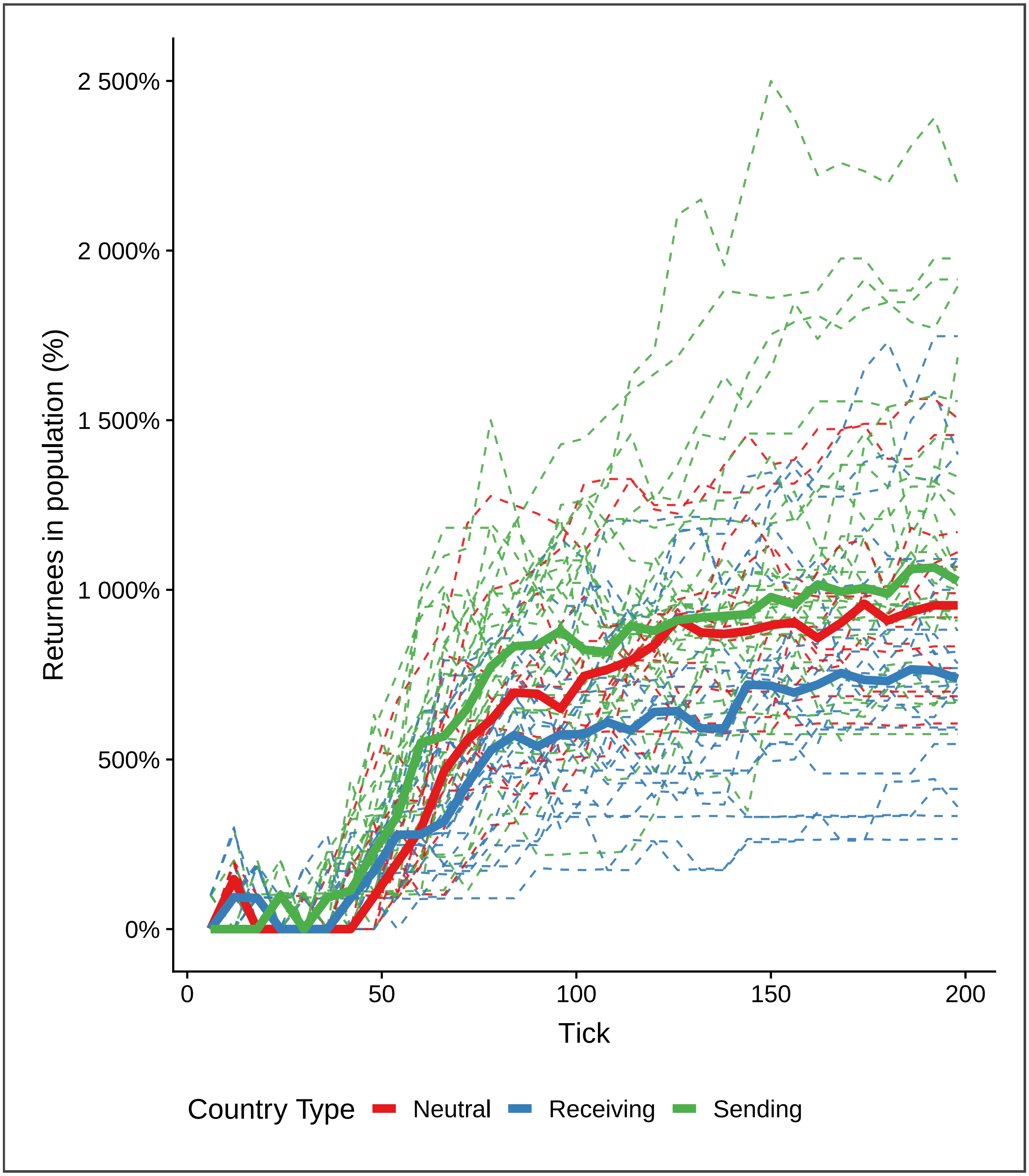


Figure 6.13: TE: Percent Returnees across Time

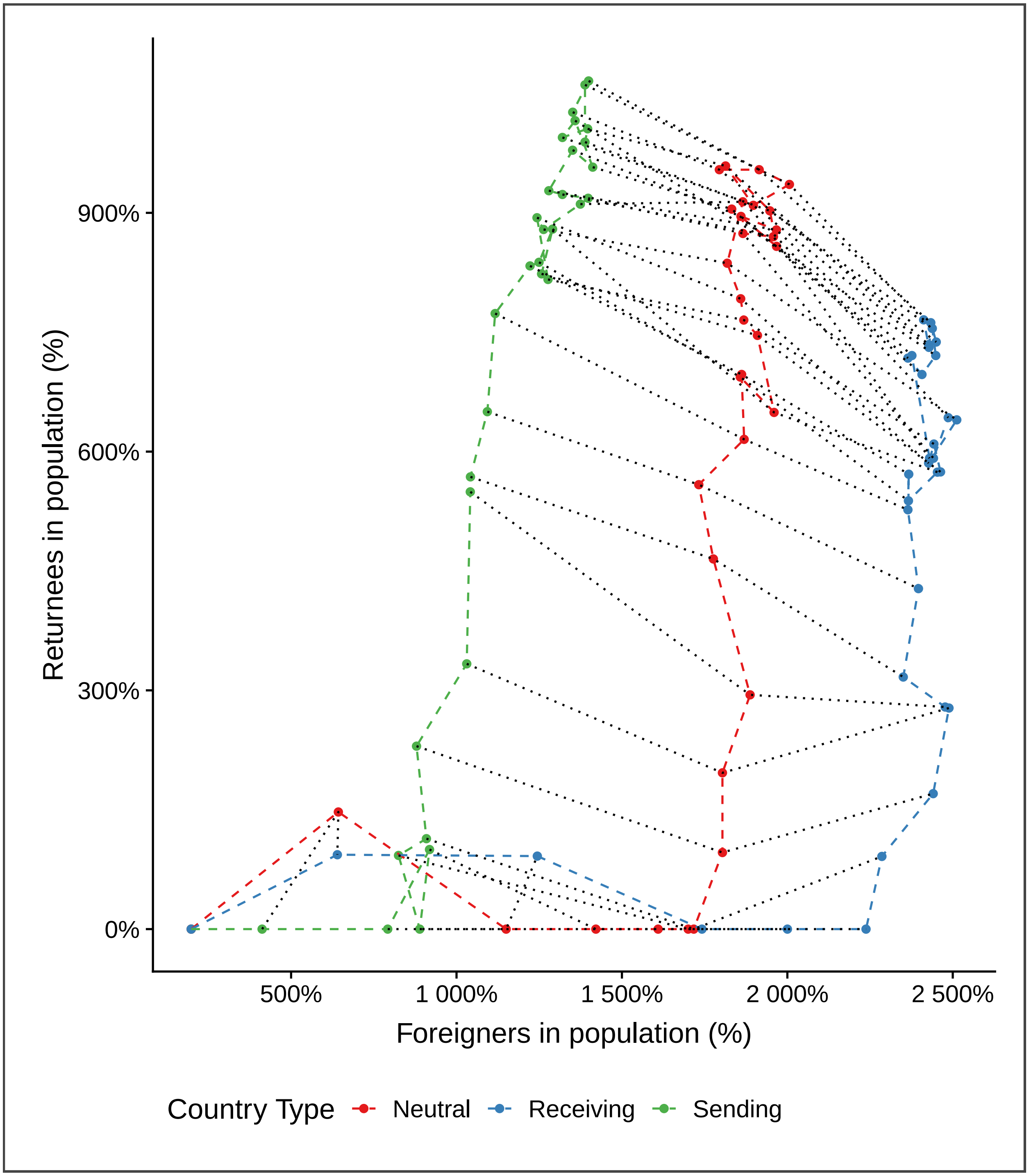


Figure 6.14: TE: Returnees vs Foreigners

	Country Type		
Migrant Type	Neutral	Receiving	Sending
Foreigner	68%	74%	59%
Returnee	31%	25%	40%

Table 6.6: TE: Percent of All Moves that are Returns

Yet another way to think about the amount of returnees is seen in table 6.6, which shows the proportion of all moves throughout the simulation that are returns to the person's originating country, by country type. Throughout the simulation, sending, receiving, and neutral countries accumulate 40.5%, 25.2% and 31.4% returns respectively. Beyond what is shown in figures 6.13 and 6.14, table 6.6 shows that it's not just some small portion of emigrants returning, but that many cycle between their home and foreign countries. The implication of figures 6.13-6.14 and table 6.6 is that returnees exist in significant enough numbers to make a difference in population-level policy preferences. Whether they do make an impact remains to be seen, which we do now using a series of linear regression models.

Tables 6.7-6.8 show several linear regression models predicting median legislature ideology as a function of country type, ideology in the previous period, and the proportion of the population that is immigrants and returnees lagged to various points in the past. The chamber's ideology in the previous period is included because beliefs are clearly autocorrelated – what I believed yesterday will have an impact on what I believe today. This predictor is lagged by a single period, which in this case is an election period (5 ticks, tunable on a per country basis if desired). Percent returnees and percent immigrants are included in instantaneous form (i.e. value for the current period without lag) as well as with various values of lag to see whether they require some time to affect the response. The latter effect may manifest if the primary effect of returnees is to spread their liberal sentiments through their social networks, as this effect will take time to subsequently affect legislative preferences. Likewise for changes in natives' sentiments arising from interactions with foreigners. If, on the other hand, the primary effect if returnees is via the vote, then we should see their presence have an immediate effect on the election taking place while they are present.

	OLS			
	(1)	(2)	(3)	(4)
(Intercept)	43.224***	42.805***	42.602***	41.502***
	(1.173)	(1.195)	(1.228)	(1.177)
Receiving	2.227***	2.233***	2.371***	1.656***
	(0.461)	(0.467)	(0.471)	(0.456)
Sending	−0.357	−0.426	−0.615	0.285
	(0.459)	(0.463)	(0.465)	(0.454)
Median Ideology (t-1)	0.044*	0.062***	0.077***	0.050**
	(0.023)	(0.023)	(0.023)	(0.023)
% Immigrants (t-1)	0.183***			
	(0.029)			
% Immigrants (t-3)		0.164***		
		(0.029)		
% Immigrants (t-5)			0.131***	
			(0.029)	
% Immigrants				0.283***
				(0.028)
% Returnees	−0.123***	−0.139***	−0.122***	−0.171***
	(0.038)	(0.039)	(0.039)	(0.037)
N	1,886	1,886	1,886	1,886
Adjusted R^2	0.123	0.119	0.114	0.150

*p < .1; **p < .05; ***p < .01

Table 6.7: TE: Median Legislature Ideology and Returnees

	OLS		
	(5)	(6)	(7)
(Intercept)	42.462***	42.704***	42.647***
	(1.182)	(1.167)	(1.166)
Receiving	2.664***	2.652***	2.741***
	(0.462)	(0.464)	(0.461)
Sending	−0.825*	−0.806*	−0.900**
	(0.461)	(0.462)	(0.459)
Median Ideology (t-1)	0.051**	0.047**	0.044*
	(0.023)	(0.023)	(0.023)
% Immigrants (t-1)	0.141***	0.144***	0.139***
	(0.029)	(0.029)	(0.028)
% Returnees (t-1)	0.056		
	(0.038)		
% Returnees (t-3)		0.046	
		(0.039)	
% Returnees (t-5)			0.089**
			(0.037)
N	1,886	1,886	1,886
Adjusted R^2	0.119	0.119	0.121

*p < .1; **p < .05; ***p < .01

Table 6.8: TE: Median Legislature Ideology and Returnees

Overall, tables 6.7-6.8 support the notion that returnees cause sending countries to liberalize more than receiving countries. First, the terms for country type show the signs that we expect across all models: receiving countries are more conservative and sending are less conservative than neutral countries. Likewise, for all models the autocorrelation term for past ideology is positive and statistically significant at the 0.1 level or better.

The percentage of the population that is immigrants is included with 3 different lag values: 1, 3, and 5 periods. For most of the models (viz. 1 and 5-7), a lag value of 1 is used, as that is the shortest lag for which the proposed mechanism of network effects can actually function. Models 2 and 3 include longer lags to see whether the effect still persists since there is no reason to think that negative sentiments that arose more than a single period ago would simply disappear. These lagged predictors still attain high statistical significance and encouragingly, have smaller magnitudes than the single period lag term. This is what one would expect to occur as sentiments become diluted over time (i.e. a negative/conservative opinion is formed, then is moderated through interactions with others with relatively more liberal attitudes). On the other hand, it is troubling that the non-lagged immigrant percentage variable in model 4 also attains such a large effect size and statistical significance because it cannot be functioning in the way theorized (since a lag is required for the network effects to proliferate). This suggests that the effects seen on the lagged variables may also be artifacts of the data or that there may be some other substantive mechanism at play. Either way, more study is required to make a determination.

Finally, for the percentage of the population that is returnees, the no lag term attains the sign that we expect (more returnees lead to more liberalization) and a high level of statistical signifiance. On the other hand, when the predictor is lagged, it loses its statistical significance in 2 of 3 cases and has an effect of the wrong direction for all three. As earlier stated, the fact that the instantaneous (non-lagged) predictor has the desired effect suggests that returnees affect ideology directly by voting their preferences. On the other hand, they appear to have an ambiguous effect via the attitudes they share with their social networks, seen in the lackluster lagged predictors.

The bottom-line is that while it is not conclusive, the models presented in tables 6.7-6.8

support the idea that returnees are a major driver of liberal sentiments in the legislatures of their home countries, and this effect is manifested directly through the vote rather than via the spread of their beliefs in their social networks.

6.2.2 Native Interactions with Foreigners

We now examine the impact of negative and positive interactions between foreigners and natives on legislative ideology. As earlier stated, the idea of there being a net positive, then swing to net negative sum of interactions is not borne out by figures 6.12 and 6.5. While the population gain of receiving countries does occur in the first quarter of the simulation (figure 6.5), this is not preceded/accompanied by a positive shift in ideology in those countries (figure 6.12). The reverse is true for sending countries. Figure 6.5 clearly shows that that this gain occurs wholly in the first quarter of the simulation. After that period, the three country types look relatively similar in terms of gross in-migration.

Further, careful consideration of the model makes it quite clear that only the first part of that causal chain can be true. Yes, it is true that more negative interactions will occur *in gross* but this does not lead to more of them *in net*. This is because, as explained in section 4.1.1, interactions are modeled to be probabilistically either positive or negative, with that probability being 0.5 by default for all scenarios (see table 3.2). In other words, more immigrants mean just as many positive interactions as negative, so the simple explanation does not make sense. And conveniently, this expectation is borne out by figure 6.15.

The top panel of figure 6.15 shows the ratio of negative to positive interactions between immigrants and natives (NPFI ratio) against time, across all simulations, categorized by country type. NPFI ratio is calculated simply as $\frac{\text{\# Negative Interactions between Foreigners \& Natives}}{\text{\# Positive Interactions between Foreigners \& Natives}}$. All values are padded with a small term to deal with periods in which there were 0 interactions of one type or the other to avoid division by 0. The short of it is that the plot is what we'd expect from model specifications. Plots for all country types are centered on 1, indicating a 1:1 ratio of negative to positive interactions. Variability around this line arises from the probabilistic way in which each interaction is determined to be positive or negative.

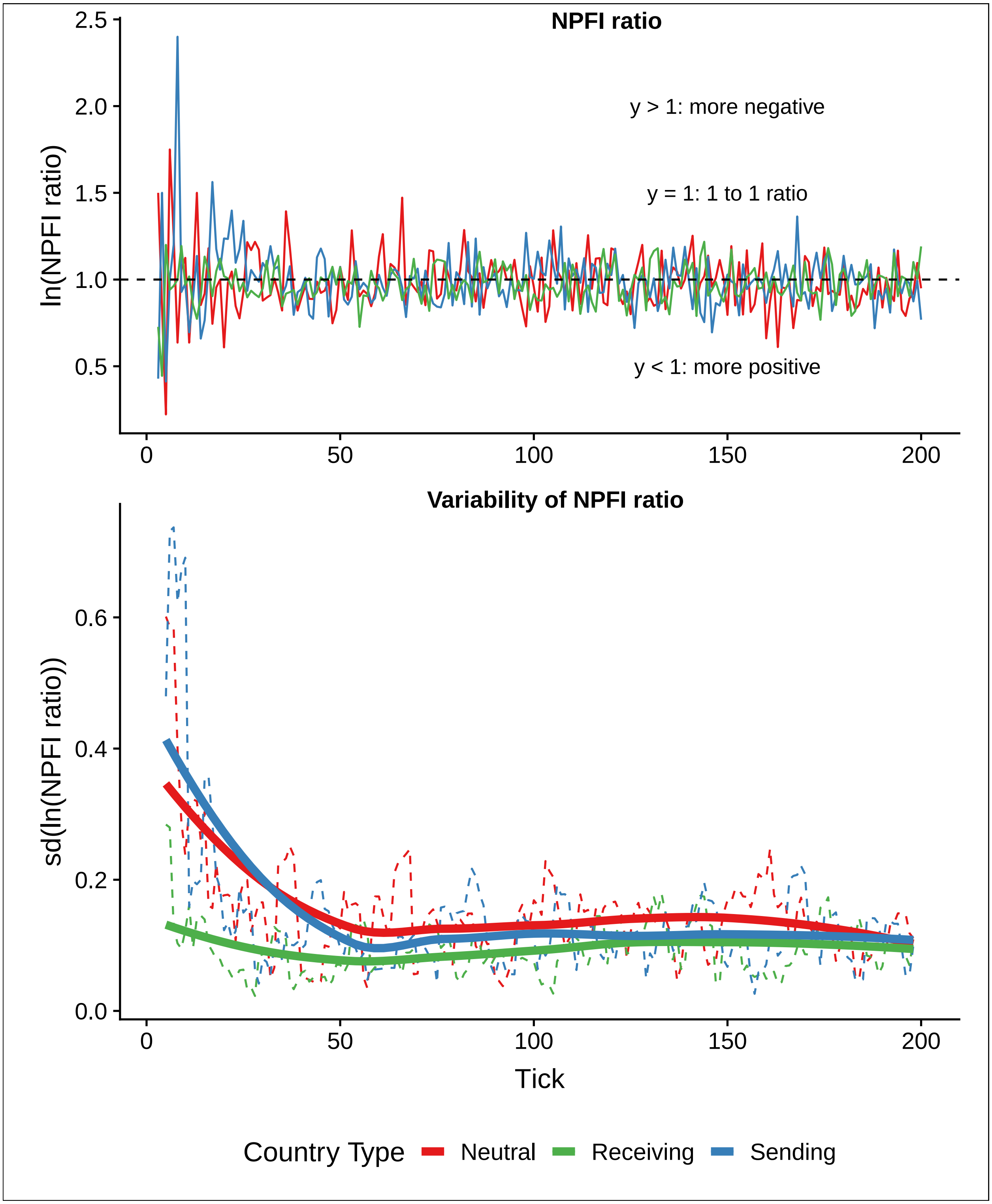


Figure 6.15: TE: Ratio of Negative to Positive Foreigner Interactions (NPFI ratio) by Country Type

The most interesting feature of the upper plot of figure 6.15 is the variability in ratios and differences in this variability by country type. Receiving countries start with a low level of variability while neutral and sending countries start much higher but eventually converge to the levels of the former. This is seen in the funneling shape of the upper plot in figure 6.15 and explicitly in the lower panel of the same figure, which displays a rolling standard deviation (binwidth of 5) of the ratios in the upper plot. It makes sense that receiving countries will have lower variability than the others from the start because it starts with a larger population of immigrants (figure 6.5) meaning each individual interaction matters less. Conversely, the impact of each immigrant on ideological preference is relatively larger in sending and neutral countries simply because there are fewer of them to average across. To put it another way, more immigrants means more stability in the effect immigrants have on native sentiments.

At the same time, while variability in the ratio means interactions will push ideological preferences in one direction or another (whenever negative and positive interactions don't cancel out), the raw number of interactions matter as well. A tick in which there are 2 negative and 1 positive interactions yields a ratio of 2 the same as one in which there are 20 to 10. The latter, however, will have a far greater effect on altering ideological preferences of the population. Therefore, we should also examine the magnitude of these interactions as in figure 6.16.

Figure 6.16 shows the sums of all interactions between a native and a foreigner by city type. The bottom panel shows the differences between the quantities of negative and positive interactions between the same. As we would expect, the raw numbers of interactions (i.e. top panel), irrespective of direction, directly parallel the number of immigrants received by each country type; receiving countries have the most, with sending and neutral following in that order. The bottom panel shows, however, that despite the differences we see in figures 6.15 - 6.16, the actual effect of negative and positive interactions is basically similar over time for all country types because high variance is balanced out by smaller raw number of interactions. Further, since the patterns for all city types regress to a zero slope, there is little chance that this metric will be able to explain the change in ideology over time we see

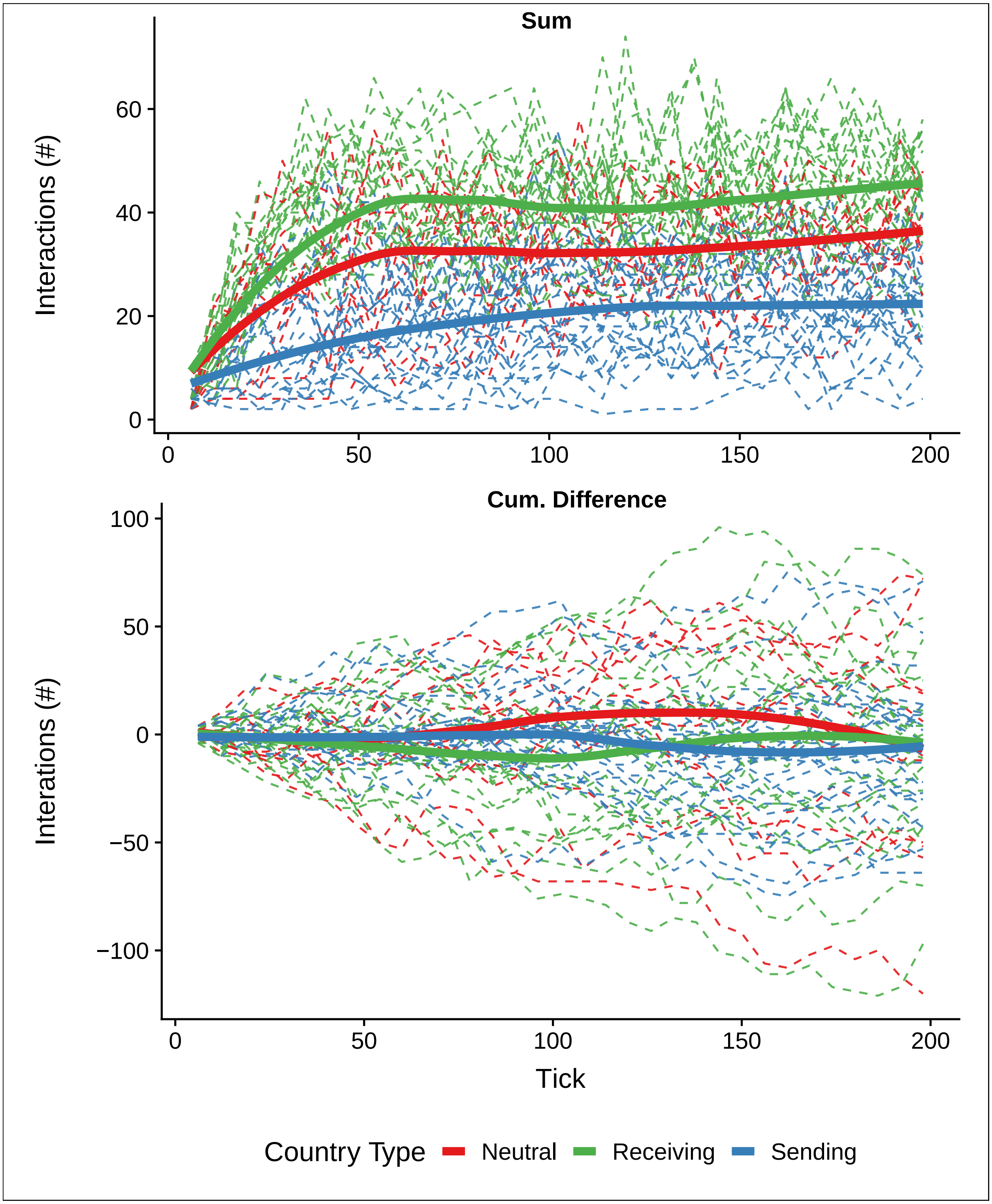


Figure 6.16: TE: Counts of Negative and Positive Foreigner Interactions by Country Type

in figure 6.12. In other words, while the patterns in interactions are interesting, it does not seem likely that the NPFI ratio can explain the correspondence we see between country type and ideology.

6.2.3 Combining Returnees and Foreigner Interactions

We can test the assertions from subsections 6.2.1 and 6.2.2 together in a regression setting. I build on the regressions in table 6.7 by supplementing with count measures of the interactions between foreigners and natives as opposed to the NPFI ratio, which we saw in the previous section is not an effective measure for the effect we are interested in. As with the regressions in tables 6.7, the main explanatory phenomena may be divided into those with instantaneous and lagged effects. As before, instantaneous effects are those which will have an immediate effect in the period they are attained, presumably by pushing persons to vote in some way. Lagged effects are those which will manifest through the intermediate step of proliferating through social networks to color the preferences of others then subsequently their vote. The intermediate step is what requires the lag time. Additionally, these variables are also examined in cumulative form, which is the sum of all values of that variable for that country until that period. The idea there would be that the buildup of particular sentiments over time is what matters, not just the amount in a particular period.

Three measures are examined in this regard: number of positive interactions, number of negative interactions, and the difference between negative and positive interactions. The first two are straightforward; if interactions matter in an instantaneous fashion, then positive interactions should be associated with a liberalizing tendency (i.e. a negative coefficient) and negative interactions with a conservatizing one. The difference measure is intended to check whether what matters is not just the incidence of positive and/or negative interactions but the relative balance between the two. If the causal mechanism works as theorized, this variable should have a positive coefficient since the measure is operationalized as the number of negative interactions minus the number of positive interactions, so positive values would indicate more negative interactions than positive.

	(1)	(2)	(3)	(4)	(5)
(Intercept)	42.033*** (1.168)	43.162*** (1.173)	42.653*** (1.182)	42.771*** (1.192)	44.728*** (1.119)
Receiving	1.581*** (0.462)	2.257*** (0.461)	1.927*** (0.468)	2.011*** (0.472)	0.965** (0.446)
Sending	0.420 (0.464)	−0.405 (0.459)	−0.007 (0.469)	−0.115 (0.472)	1.090** (0.447)
Median Ideology (t-1)	0.034 (0.023)	0.047** (0.023)	0.042* (0.023)	0.044* (0.023)	0.003 (0.022)
% Immigrants (t-1)	0.147*** (0.029)	0.227*** (0.038)	0.164*** (0.029)	0.176*** (0.029)	0.118*** (0.028)
% Returnees	−0.167*** (0.038)	−0.122*** (0.038)	−0.149*** (0.038)	−0.142*** (0.039)	−0.457*** (0.043)
# Neg. Interactions	0.172*** (0.023)				
# Neg. Interactions (t-1)		−0.054* (0.031)			
# Neg. Interactions (t-3)			0.079*** (0.024)		
# Neg. Interactions (t-5)				0.049** (0.023)	
# Neg. Interactions (cum.)					0.016*** (0.001)
N	1,886	1,886	1,886	1,886	1,886
Adjusted R^2	0.148	0.124	0.128	0.125	0.209

*p < .1; **p < .05; ***p < .01

Table 6.9: TE: Negative Interactions & Median Legislature Ideology

	(1)	(2)	(3)	(4)	(5)
(Intercept)	42.423*** (1.179)	43.253*** (1.175)	42.891*** (1.182)	42.386*** (1.203)	44.692*** (1.135)
Receiving	1.827*** (0.466)	2.234*** (0.461)	2.020*** (0.470)	1.915*** (0.471)	1.144** (0.452)
Sending	0.137 (0.468)	−0.367 (0.459)	−0.114 (0.471)	−0.010 (0.472)	0.805* (0.451)
Median Ideology (t-1)	0.038 (0.023)	0.044* (0.023)	0.041* (0.023)	0.047** (0.023)	0.007 (0.023)
% Immigrants (t-1)	0.158*** (0.029)	0.195*** (0.039)	0.171*** (0.029)	0.170*** (0.029)	0.119*** (0.028)
% Returnees	−0.152*** (0.038)	−0.124*** (0.038)	−0.145*** (0.039)	−0.149*** (0.039)	−0.403*** (0.043)
# Pos. Interactions	0.111*** (0.024)				
# Pos. Interactions (t-1)		−0.014 (0.032)			
# Pos. Interactions (t-3)			0.053** (0.024)		
# Pos. Interactions (t-5)				0.071*** (0.024)	
# Pos. Interactions (cum.)					0.014*** (0.001)
N	1,886	1,886	1,886	1,886	1,886
Adjusted R^2	0.133	0.123	0.125	0.127	0.188

*p < .1; **p < .05; ***p < .01

Table 6.10: TE: Positive Interactions & Median Legislature Ideology

	(1)	(2)	(3)
(Intercept)	43.240***	43.168***	43.203***
	(1.172)	(1.175)	(1.173)
Receiving	2.220***	2.228***	2.230***
	(0.460)	(0.461)	(0.461)
Sending	−0.354	−0.360	−0.360
	(0.458)	(0.459)	(0.459)
Median Ideology (t-1)	0.044*	0.045*	0.045*
	(0.023)	(0.023)	(0.023)
% Immigrants (t-1)	0.184***	0.183***	0.183***
	(0.029)	(0.029)	(0.029)
% Returnees	−0.123***	−0.122***	−0.121***
	(0.038)	(0.038)	(0.038)
Δ Interactions	0.050**		
	(0.021)		
Δ Interactions (t-1)		−0.018	
		(0.021)	
Δ Interactions (t-3)			0.022
			(0.021)
N	1,886	1,886	1,886
Adjusted R^2	0.125	0.123	0.123

*p < .1; **p < .05; ***p < .01

Table 6.11: TE: Net Interactions & Median Legislature Ideology

	(4)	(5)	(6)
(Intercept)	43.183***	42.771***	42.788***
	(1.174)	(1.159)	(1.160)
Receiving	2.228***	2.217***	2.215***
	(0.461)	(0.455)	(0.455)
Sending	−0.358	−0.136	−0.141
	(0.459)	(0.454)	(0.454)
Median Ideology (t-1)	0.045*	0.050**	0.049**
	(0.023)	(0.023)	(0.023)
% Immigrants (t-1)	0.183***	0.205***	0.204***
	(0.029)	(0.028)	(0.028)
% Returnees	−0.123***	−0.143***	−0.142***
	(0.038)	(0.037)	(0.037)
Δ Interactions (t-5)	−0.017		
	(0.021)		
Δ Interactions			0.016
			(0.021)
Δ Interactions (cum.)		0.036***	0.035***
		(0.005)	(0.005)
N	1,886	1,886	1,886
Adjusted R^2	0.123	0.146	0.146

*p < .1; **p < .05; ***p < .01

Table 6.12: TE: Net Interactions & Median Legislature Ideology

Table 6.9 shows the results for models examining the different measures of negative interactions. The control variables for these models (viz. country type, lagged ideology, lagged percent immigrants, and percent returnees) all remain consistent with what was seen table 6.7. As for the main explanatory variables, all the different measures of negative interactions attain statistical significance at the 0.1 level or greater, which is good. However, there is a wrinkle: the term for a single period lag has the wrong sign (the rest are consistent with the theory). This unfortunately casts doubt on all the models. After all, if a measure attains statistical significance for a coefficient with the wrong sign, then are these models actually representing what we think they are? Any number of things may be wrong - the theory, the operationalization of the concept, the data, or the model. In all cases, the doubt cast by the one model also undermines the credibility of the others.

Fortunately, the issue is clarified by table 6.10, which displays the models exploring the effect of positive interactions. At first, this table seems to only muddy the water further. As with table 6.9, most of the explanatory variables attain statistical significance but the problem here is that all show effects in the wrong direction. If these models supported our theory, we would expect the estimates on the positive interaction variables to be negative - more interactions should lead to lower (i.e. more liberal) ideology. Unfortunately, they do the opposite, and in a statistically significant way. Where this actually clarifies matters is that it forces one to consider the underlying structure of the data. Upon deeper inspection, it turns out that both the negative and positive interaction counts are highly collinear with the number of immigrants ($\rho_{neg} = 0.794$, $\rho_{pos} = 0.801$), which makes sense since they would grow proportional to the number of immigrants. So in retrospect, of course, they will track positively with ideology since that is the relationship we are ultimately trying to explain. These measures are therefore not usable in their present form.

The final set of regressions (tables 6.11 - 6.12) focus on the difference between negative and positive interactions, measured in both instantaneous and a cumulative form. These variables don't suffer from the collinearity problem, seen in the fact that their correlations with the number of immigrants are -0.001 and -0.031 respectively - essentially nil. The first of the models using these variables (column 1) suggests that having more negative than positive interactions in a period may have a conservatizing and statistically significant impact on ideology. None of the lagged versions of these variables is statistically significant and the estimated effects vary in direction, presenting an ambiguous picture of their impact (columns 2-4). Like in the first model, the cumulative version of the variable (column 5) also shows a positive and statistically significant impact on ideology. Finally, when the instantaneous and cumulative variables are both included in the model (column 6), the former loses statistical significance while the latter retains its original effect size as well as statistical significance. Bottom line, models 1 and 6 suggest that the accumulation of sentiment resulting from interactions between foreigners and natives has a clear impact on legislature composition by determining who gets voted in.

Across all the models of tables 6.9 - 6.12, percent immigrants and percent returnees retain large, statistically significant effects consistent with the models considered in table 6.7. Recall that the analysis of those models ended with the conclusion that returnees appear to have a significant effect, at least when examined as the sole driver of ideological change. That their estimated effect remains consistent even when examined simultaneously with a potential competing theory increases our confidence that they have a meaningful impact. The same can be said of the competing theory (viz. native and foreigner interactions) as well. By the same token, the positive relationship between legislature ideology and the percentage of the population that is immigrants seems to also matter independent of these two aforementioned mechanisms. This is seen in the fact that the estimate for the variable measuring percent immigrants remains large and statistically significant across all the examined models. Taken as a whole, tables 6.9 - 6.12 support the ideas that returnees and the character of interactions between natives and foreigners form at least part of the reason why receiving countries tend to be more closed than sending ones.

CHAPTER 7

Scenario 3: U.S.-Mexico Border in California

This third and final chapter applies MAPES to a scenario modeling the U.S.-Mexico Border in California (UMBC). The UMBC scenario is comprised of Southern California to the north and Baja California to the south, with 3 cities on the U.S. side, and 2 on the Mexican. I model a real-world area for two major reasons. First, it is relatively easy to perform a first pass, common-sense validation of model output for a real area since it is actually observable, as opposed to a hypothetical scenario like TE, which requires more theoretical introspection. Second, using MAPES to model a real area demonstrates its applicability to research with immediate relevance, and provides a starting template for future work. The California border region was specifically chosen because (1) the area is well studied in the immigration studies literature (e.g. Kandel and Massey (2002), D. S. Massey (1987), Cornelius and Tsuda (2004)), helping contextualize results; and (2) the cities comprising the region are diverse in characteristics, which will lead to interesting results.

The five modeled cities are Los Angeles, San Diego, Riverside, Tijuana and Mexicali. Figure 7.1 is a map of the region, and Mercator coordinates for each city are recorded in table 7.1. These coordinates are integral to the person-agents' migration decision making process as they determine the cost of moving from one city to another. As done in previous chapters, figure 7.2 shows distributions of several of the agent characteristics shown in table 7.1.

Figure 7.1: Person: Simulation Area – Southern & Baja California.

Structurally, this chapter is broken into two sections. The first elucidates the basic mechanics of the scenario, focusing on flows and outcomes for populations on both sides of the border. Emphasis is placed on what is changing when, and who is going where. The second portion of the chapter looks at the effect of a policy intervention, namely increasing the cost of movement across the border from Mexico to the U.S. The purpose of that section is to identify what impact such a change has on the outputs examined in the first section. In order to make comparative statements (i.e. how does output X change in the intervention scenario compared to the non-intervention one?) all simulations are reconducted in duplicate: once with the "normal" crossing cost, and again with an increased crossing cost, ceteris paribus. Many of the results are presented as "deltas" (i.e. changes) relative to the non-intervention scenario. Throughout this section, the former case is variously referred to as the "normal," "low-cost" or "open" version of the scenario. The latter is referred to as the "high-cost" or “restrictive” version.

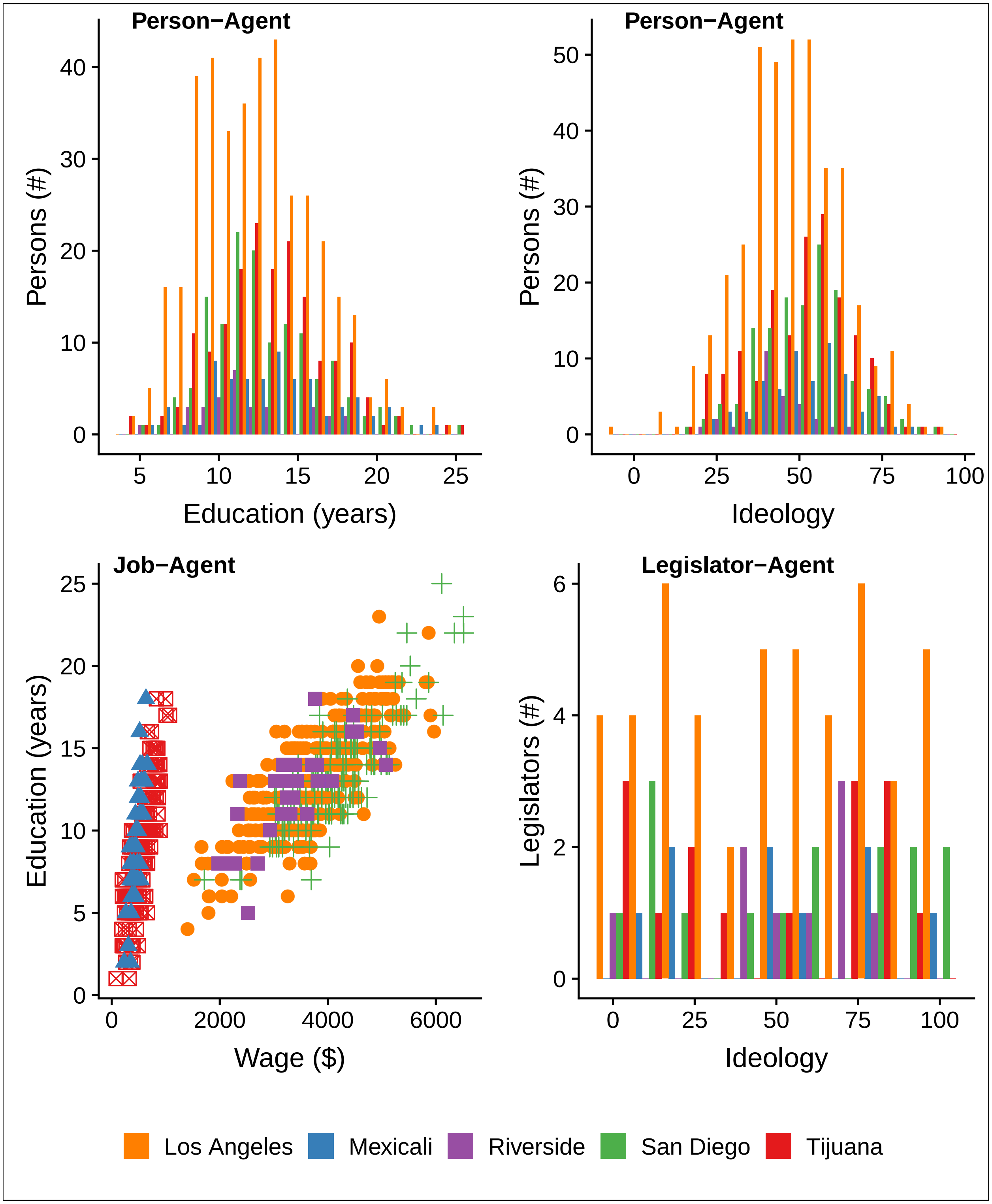


Figure 7.2: UMBC: Agent Characteristics

City	Los Angeles	Riverside	San Diego	Tijuana	Mexicali
Country	US	US	US	MX	MX
Latitude	34.05	33.9533	32.715	32.525	32.6633
Longitude	-118.25	-117.3962	-117.1625	-117.0333	-115.4678
Living Cost (Monthly)	2435	2224	1926	603	454
Population	390	31	140	170	69
Pop. Education Mean	14	14	14	9.3	9.3
Pop. Education Std. Dev.	3	3	3	3.6	3.6
Pop. Ideology Mean	50	50	50	50	50
Pop. Ideology Std. Dev.	15	15	15	15	15
# Jobs	389	30	139	163	67
Job Education Mean	14	14	14	9.3	9.3
Job Education Std. Dev.	3	3	3	3.6	3.6
Wage Mean	3801	3605	4224	537	448
Wage Std. Dev.	800	800	800	200	100
Corr. of Edu. & Wage	0.8	0.8	0.8	0.8	0.8
# Leg. Seats	16	3	5	5	3
# Leg. Candidates	48	9	15	15	9
Pop. Ideology Mean	50	50	50	50	50
Pop. Ideology Std. Dev.	15	15	15	15	15

Table 7.1: UMBC: Locale Parameters

For both the open and restrictive versions of UMBC, the cost to cross the border from North to South is just the cost of movement from one city to another. This cost is a simple multiplicative factor proportional to the distance between the cities as stated in the general model specification (chapter 4). In the normal scenario, the cost to cross the border from South to North is the distance-proportional cost plus an additional USD 500. In the high cost scenario, the cost is increased by USD 5000, which is an order of magnitude larger. These numbers were chosen as reasonable approximations of the real costs involved in crossing the border. In the low cost case, USD 500 is representative of the cost of lodging and transportation for a 1-2 day trip across the border. The figure of USD 5000, on the other hand, approximates the cost of dealing with U.S. bureaucratic processes for legal migration, or the cost of hiring a coyote for undocumented migration. Regardless the exact interpretation, the higher cost to migrate in the latter case is meant to represent the direct consequence of a policy intervention to restrict immigration.

7.1 Populations and Flows

Figure 7.3 shows the population of the region across the lifetime of the simulation for the default (open) version of UMBC. Los Angeles is basically flat the entire simulation, maintaining a median population of 393. The same is true for San Diego which centers around a population of 151. The remaining three cities have an initial period of population change that lasts until about tick 50, followed by a relative steady state. In the case of Tijuana, the transition period sees an initial dip followed by a gradually decreasing increase to a plateau around 143. Both Mexicali and Riverside show early spikes in population that stabilize to 46 and 54, respectively.

An interesting connection derived from figure 7.3 is that similar to what was shown in BEG (section 5.1), there is a relationship between the number of available jobs and steady state population values. In BEG, the population of country A, the receiving country, stabilized a shade above the number of jobs available in that country. There was a steady rotation of people from country B to country A in order to fill the labor needs of the latter,

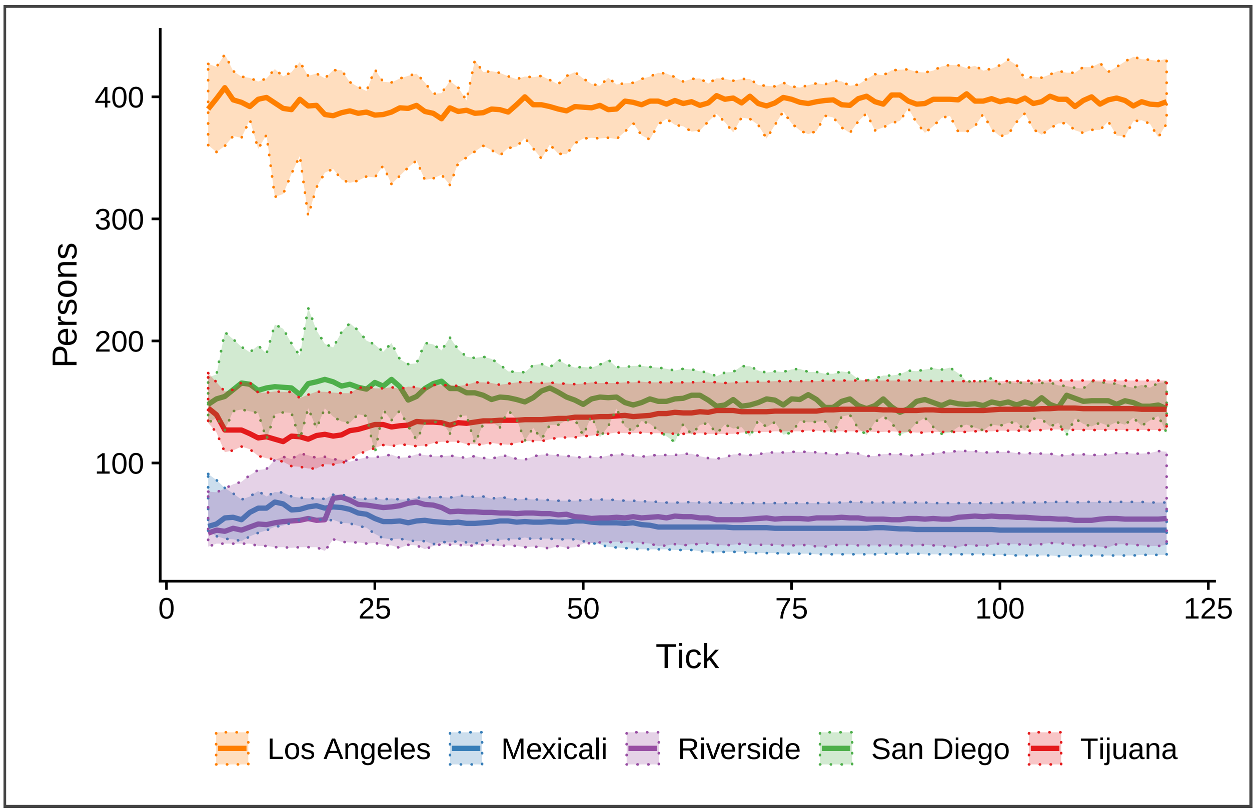


Figure 7.3: UMBC: Population by City, Open Immigration Policy

keeping it near full employment. A similar phenomenon is seen in UMBC as well. From table 7.1, we know that Los Angeles, San Diego, and Riverside have steady state populations (393, 151, 54 respectively) that are slightly higher than their numbers of available jobs (389, 139, 30). Conversely, Tijuana fall into a state of worker scarcity, with populations of only 143 and 46 available to fill 163 and 67 jobs. So, to apply the paradigm from BEG, the U.S. cities clearly fall into the role of receiving cities just as city A, while the Mexican cities are akin to the classic sending city, B.

Another noteworthy point about figure 7.3 relates to the ribbon around the median response which, as noted in section 5.1, is the range of all observations for the point estimate across multiple simulations. They are of particular interest in UMBC because the population sizes of the cities are all different and vary quite widely. Resultantly, even if the bandwidths of responses across cities are roughly similar in magnitude, as a percentage of initial population size, they can present differences on orders of magnitude. This point clearly manifests in figure 7.3. The range intervals for all cities except Riverside are roughly similar, with Los Angeles on the wider side (mean bandwidth 52.8), and Mexicali on the narrower (mean bandwidth 37.4). Riverside stands out with a relatively large bandwidth of 71.3; almost twice that of the smallest value. However, these differences are compounded when one considers them as a proportion of population. In that case, the bandwidths for all cities but Riverside range between 14% for Los Angeles and 54% for Mexicali, as a percentage of initial population. Riverside's is a massive 230% of initial population. What this means is that while Los Angeles's population fluctuates across simulations within a little more than a tenth of its initial population, Riverside's ranges between doubling and almost zeroing out. So, clearly, while there is some degree of change across all the cities in the region, Riverside exhibits population instability that is notably higher than all the other cities in the scenario.

How do these cities come to have the population patterns we see in figure 7.3? One way to answer this question is by examining the flow of persons from a network perspective. Figure 7.4 shows the median flow of persons over the lifetime of the simulation as a network graph. Nodes are cities, and edges are movements between them. Each city pair has two edges indicating movement in either direction, and arrows indicate the direction of these flows. Each edge is marked with the total number moves in that direction and does not account for individual persons, so an edge value of 4 could be 4 persons moving once, or a single person moving 4 times.

One stand-out point is that the vast majority of moves (94.8%) occur within countries (between Tijuana and Mexicali: 8.8% ; between Los Angeles, San Diego, and Riverside: 86%). Of those, most occur between Los Angeles and San Diego (76.6%), with Tijuana and Mexicali coming up a distant second (8.8%). There are several edges that have little or no movement - Riverside and Tijuana (0.3%), Riverside and Mexicali (0.4%), and Mexicali and San Diego (0.7%). It is pretty clear that the amount of traffic in a city pair is related to the size of the cities involved; all three of the low traffic edges involve either Riverside or Mexicali. However, it is unclear at this point what is the exact causal mechanism. Beyond city size, other potential causes include city distance (Riverside and Mexicali are relatively far from the geographic center of this network), as well as having the lowest median wages. MAPES could certainly be adapted to investigate this issue in future work (more on this in chapter 8)

Another interesting point deriving from figure 7.4 is that the individual components of bilateral flows tend to be similar in magnitude. That is to say, while there are net movements from one city to another, most aren't hugely imbalanced e.g. 100 travel from city A to B, but 10 travel back. We can quantify this statement using the table and plot in figure 7.5, which shows imbalances in pairwise flows. The "ratio" column in the table is the ratio of the smaller value of a pair of flows to the larger value. For example, if there are X gross movements from city A to B, and Y from B to A where Y is greater than X, then the "ratio" is simply X/Y.[1] The point of the "ratio" is to measure imbalances as a percentage of the populations involved. A problem with this measure is that it is highly sensitive when the number of movements involved is small. A one movement imbalance when $A = 1, B = 2$ leads to a ratio of 0.5, which seems very different from the one movement imbalance resulting from $A = 101, B = 102$, even though they are not substantively different. On the other hand,

[1] X/Y is used rather than Y/X because while $Y \geq X$, this metric is normalized 0 to 1.

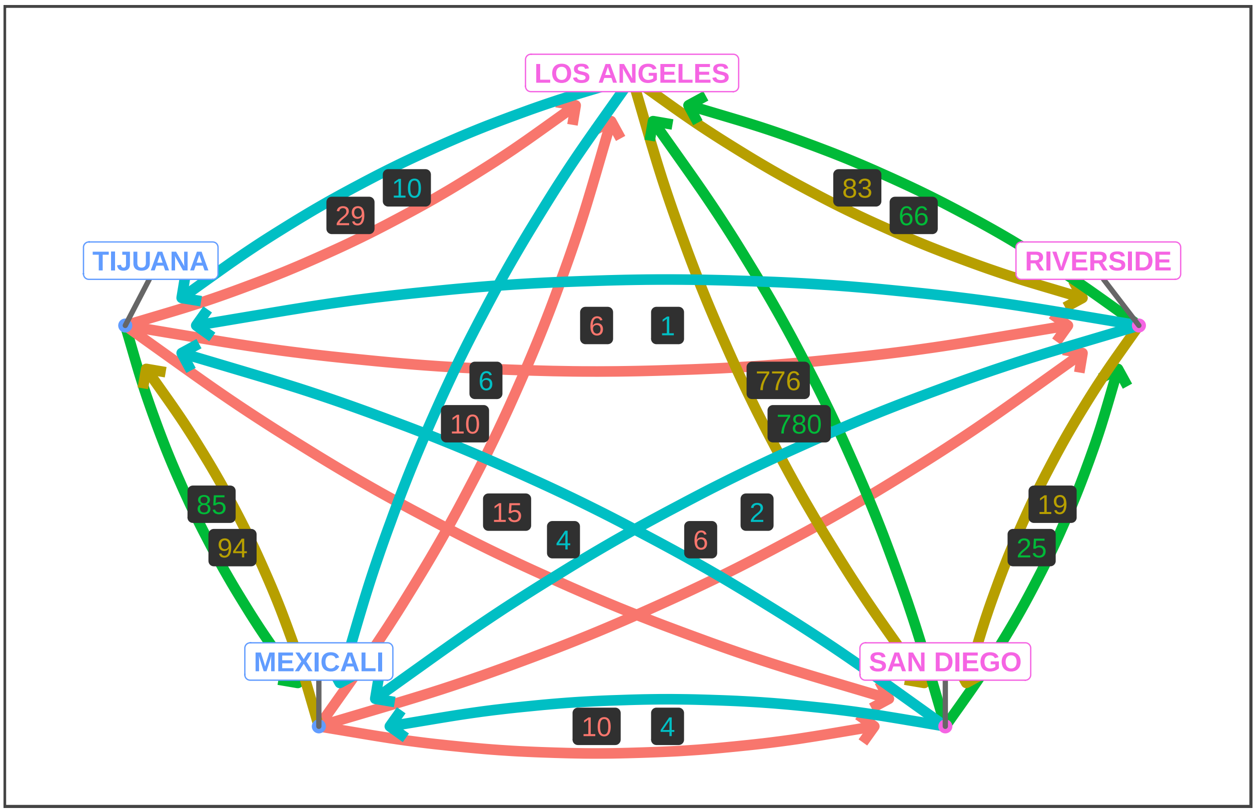


Figure 7.4: UMBC: Gross Moves Across the Region

using the raw difference as in the “diff” column, which is the difference between the larger and the smaller flow in the pair, leads to an overweighting of larger flows. What we really care about are imbalances that are large in magnitude as well as a percentage of the number of persons involved. So, the ratio and difference metrics are combined as $-1 * \text{diff} * (1 - \text{ratio})$ in the “metric” column. Note that the table in figure 7.5 shows exactly half the possible flows; each city pair has two flows associated with it (A to B, B to A) and only the lesser of the two is shown since including both would show duplicate information.

Figure 7.5 shows that there are two flows that are exceptional in their imbalance, which is to say that there is significant net movement from one city to another. These are the flows from Los Angeles to Tijuana, and San Diego to Tijuana which respectively send back only 34% and 27% of the persons who come to them. Arguably, the flows from Riverside to Tijuana, San Diego to Mexicali, and Riverside to Mexicali are just as imbalanced. But as explained earlier, it is hard to say whether this is due to the small magnitude of the populations involved, or for substantively meaningful reasons. Whether we consider only the top two or all five, it is striking that all the imbalances lead to net flows from Mexican cities to U.S. ones. Stereotypes notwithstanding, this makes sense because jobs in the U.S., even in Riverside which has the lowest median salary of the three, pay dramatically more compared to jobs in Mexico. As we saw in BEG, this is more than enough to motivate net movement so long as the cost calculation makes sense (viz. utility gained exceeds financial cost and utility lost from moving).

The information in the table of figure 7.5 can also be presented in graph form as in figure 7.6, which shows net flows rather than gross flows as in figure 7.4. Here, unlike in the previous graph, there is only one edge per city pair because the two flows for each pair are consolidated into a single net flow. For example, if there are 10 moves from A to B, and 12 from B to A, figure 7.6 shows a single flow of 2 from B to A.

Surprisingly, Riverside is the biggest net gainer from with 32 moves in and no net movement out. San Diego and Los Angeles, on the other hand, attract large numbers of moves in but also lose almost as many out (27 versus 17, and 17 versus 10, respectively). It also looks like there is a clear directionality to the flow of persons: from Mexico, to San Diego, thence

	From	To	Ratio	Diff	Metric
1	Los Angeles	Tijuana	0.3448276	−19	12.44827586
2	San Diego	Tijuana	0.2666667	−11	8.06666667
3	Riverside	Tijuana	0.1666667	−5	4.16666667
4	San Diego	Mexicali	0.4000000	−6	3.60000000
5	Riverside	Los Angeles	0.7951807	−17	3.48192771
6	Riverside	Mexicali	0.3333333	−4	2.66666667
7	Los Angeles	Mexicali	0.6000000	−4	1.60000000
8	Riverside	San Diego	0.7600000	−6	1.44000000
9	Tijuana	Mexicali	0.9042553	−9	0.86170213
10	Los Angeles	San Diego	0.9948718	−4	0.02051282

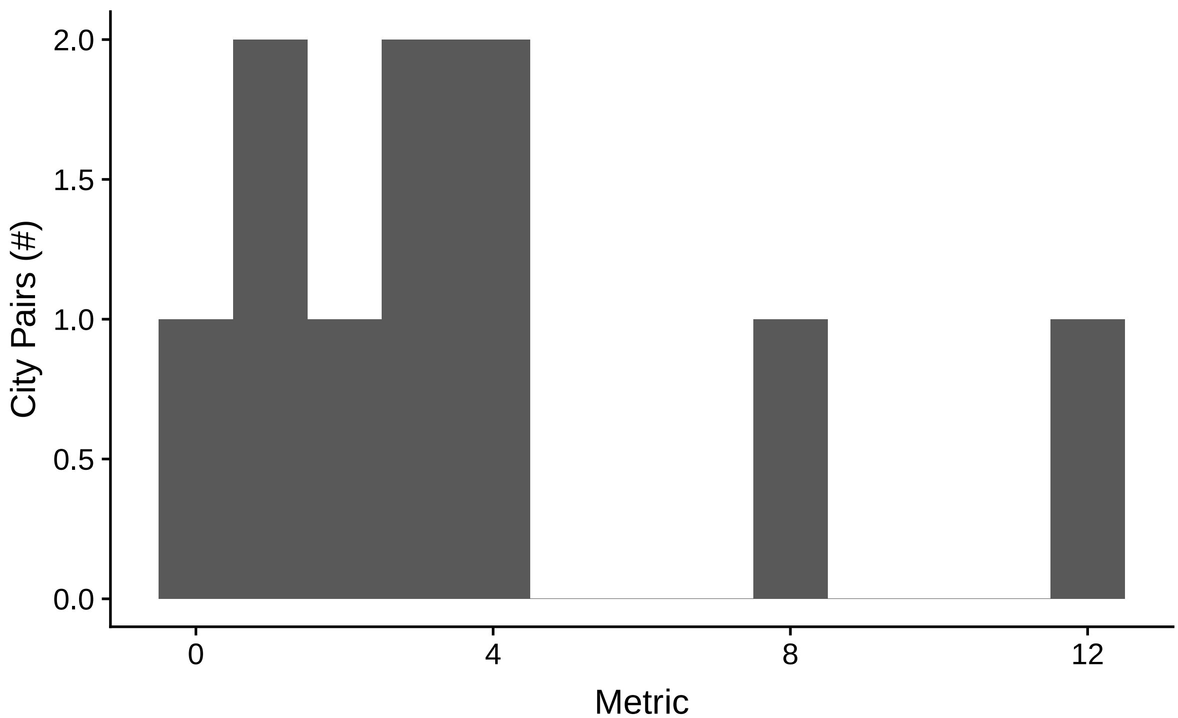


Figure 7.5: UMBC: Imbalanced Bilateral Flows

to Los Angeles, then Riverside. It is important to note that this does not necessarily mean there is a metaphoric tunnel where individuals enter in Mexico and emerge in Riverside. The more apt analogy is that each city is like a glass of water being filled past capacity. Some of the overflow results from the original contents and some from the inflow.

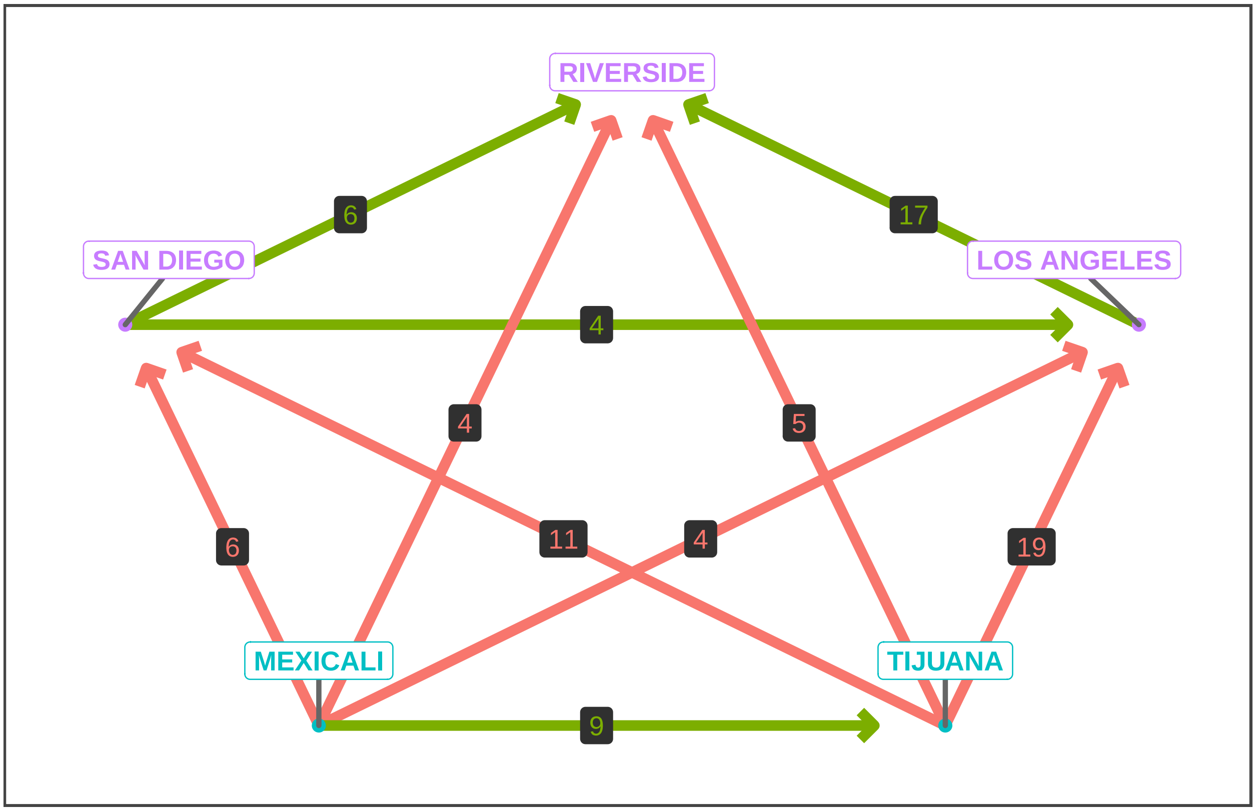


Figure 7.6: UMBC: Net moves across the region

To verify that this simulation is functioning per the water glass paradigm rather than that of the tunnel, we can examine table 7.2. Table 7.2 is a raw count of people by city of origin (rows) and city of residence at simulation's end (columns). It's immediately apparent that flows do not cleanly tunnel from one city to the next. For example, based only on Fig 7.6, one might reasonably hypothesize that 5 move from Tijuana to Riverside since the numbers of net movement plausibly support such a claim. Further, the idea of a migration pipe from Tijuana to Riverside via San Diego and Los Angeles is intuitively attractive. Table 7.2, however, immediately refutes this picture; only 5.35 persons in Riverside are originally from Tijuana. So even if all 11 net moves from Tijuana to San Diego are originally from

Tijuana, much of that initial stream sloughs off either in San Diego or Los Angeles so that by the time the “pipe” approaches Riverside, the flow is an amalgam of persons from all cities along the way.

	City From				
Final City	Los Angeles	Mexicali	Riverside	San Diego	Tijuana
Los Angeles	262.30	12.35	13.70	84.05	27.05
Mexicali	4.65	13.80	0.90	2.40	28.10
Riverside	24.55	3.75	9.75	9.55	5.35
San Diego	89.45	4.80	5.60	41.10	10.65
Tijuana	9.05	34.30	1.05	2.90	98.85

Table 7.2: UMBC: Final city populations by individuals’ originating city (mean across simulations)

Further insights are gained by transforming Table 7.2 to form tables 7.3, 7.4 and 7.5. Table 7.3 shows the percentage of each city’s population at the end of the simulation by the city in which each person originated. The rows are the cities of residence and the columns, the cities of origin. So for example, row 2, column 3 shows the percentage of Mexicali’s final population that originated in Riverside (1.8%). Table 7.4, on the other hand, shows the percentage of each city’s starting population that ends up in each destination. In this case, row 2, column 3 shows the percentage of Mexicali’s original population that ended up in Riverside (5.4%). Finally, table 7.5 is the percent growth of each city across the simulation.

Examining tables 7.2 – 7.5 in concert, one sees several interesting patterns and ways to characterize cities and flows. First, there are cities that we might think of as transit cities. Tijuana is one example. Fig 7.6 shows a net movement of 11 from Tijuana to San Diego but table 7.2 shows that only 10.65 in San Diego are originally from Tijuana. So the remainder must be coming from other cities that use Tijuana as a transit point on the way to San Diego. The obvious candidate is Mexicali, but we see that only 4.8 are from that city (table 7.2: row 4, column 2) so at the very least some of the traffic from Los Angeles and Riverside to San Diego must occur via Tijuana. In other words, Tijuana is a transit city in the sense

	City From				
Final City	Los Angeles	Mexicali	Riverside	San Diego	Tijuana
Los Angeles	65.7%	3.1%	3.4%	21%	6.8%
Mexicali	9.3%	27.7%	1.8%	4.8%	56.4%
Riverside	46.4%	7.1%	18.4%	18%	10.1%
San Diego	59%	3.2%	3.7%	27.1%	7%
Tijuana	6.2%	23.5%	0.7%	2%	67.6%

Table 7.3: UMBC: % of city X originating in city Y

	Final City				
City From	Los Angeles	Mexicali	Riverside	San Diego	Tijuana
Los Angeles	67.3%	1.2%	6.3%	22.9%	2.3%
Mexicali	17.9%	20%	5.4%	7%	49.7%
Riverside	44.2%	2.9%	31.5%	18.1%	3.4%
San Diego	60%	1.7%	6.8%	29.4%	2.1%
Tijuana	15.9%	16.5%	3.1%	6.3%	58.1%

Table 7.4: UMBC: % of city X ending in city Y

that people travel there not as a final destination but as a means of transiting to another one.

While Tijuana is a pass-through point for people from other cities, Tijuana's native population is also affected by migration. This is easily seen from Table 7.5; Tijuana shrinks to 86% of its original size. Thus, Tijuana is simultaneously a transit city as well as a sending one. While the observation is obvious, it is worth explicitly pointing out because Tijuana could theoretically have stayed net neutral or grown instead of shrinking. It indicates that whatever dimension of migration is captured by this idea of "transitory", it is to some extent orthogonal to the basic distinction between "sending" and "receiving" locales. While I do not investigate it here, one could imagine that it is some confluence of economic and population conditions combined with geographic circumstance that determines a locale's position amidst

Los Angeles	Mexicali	Riverside	San Diego	Tijuana
102%	72%	171%	108%	86%

Table 7.5: UMBC: Population percent growth by city

the two dimensions (sedentary versus transitory, sending versus receiving) described.

Cities that are of special attraction are also worth identifying. Observationally, one might describe attraction in a couple different ways. First, attraction can be thought of in terms of retention: if a city is "attractive," it stands to reason that more of the city's original population would want to stay there. In that sense, Los Angeles (67.3% city native) and Tijuana (58.1% city native) are the most attractive cities in the present simulations. The remaining cities retain less than 31.5% of their starting populations, with their remaining populations seeking greener pastures in the other cities. However, this kind of thinking is potentially problematic in that cities may retain their populations not because there is something positively attractive about them but because there are barriers to exit. On the one hand, a city may entice persons to stay because it presents a relatively high expected wage or large social network. On the other, they may impede outflows by imposing a high cost of leaving (whether economic, social or political).

Another way to think about attraction is in terms of attracting inflows, which is related to but still distinct from retaining the already present population. In the simplest sense, people can be driven to move by the same factors driving others to stay put. Without being too Western-centric in outlook, more political empowerment is desirable whether one is a pauper or a prince. Likewise with expected wealth or social utility. However, such a blunt analysis only works up to a point. If we think about locales as having finite capacity (whatever the cause may be) and consider this capacity to have structure in terms of population characteristics (e.g. gender, education, profession, politics), then the picture becomes more complex. Take the case of Riverside. Riverside does a poor job retaining its original population, with only 18% of its starting population choosing to remain. However, the population at simulation's end is 171% of its original size; clearly, there is something attractive about Riverside to somebody, even if not its native sons.

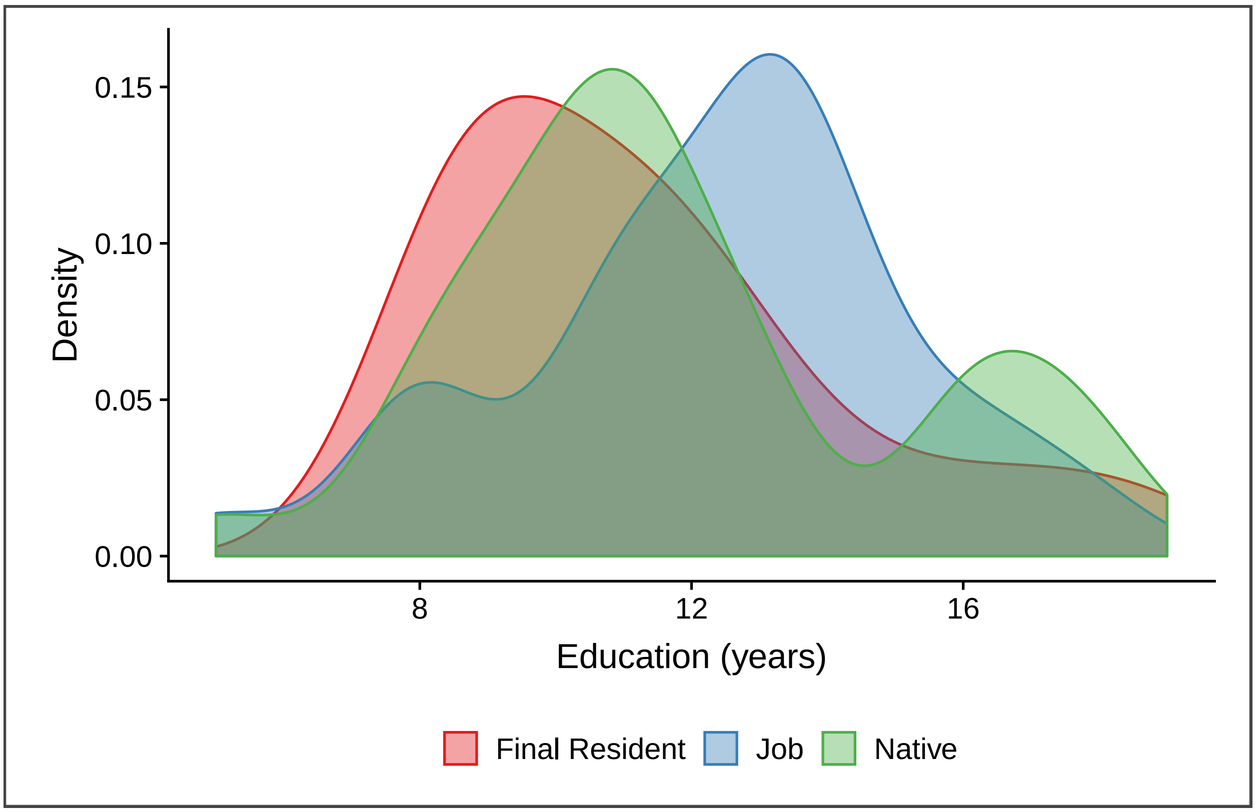


Figure 7.7: UMBC: Riverside Education Distributions

One part of the answer is that the distribution of available jobs in Riverside is mismatched against its initial endowment of labor (Figure 7.7). So to the extent that persons are dissatisfied with the jobs they have, or get fired for not performing to par, they will migrate to another city where they can find a better match. The remaining vacuum will be filled by people from other cities where the education match or utility payoff is even worse; the same distribution of jobs driving some people out entices others in. However, we know this is not the entire story because the education distribution of the population in Riverside at the end of the simulation actually diverges even further from the job distribution than the initial population! To put it another way, the people living in Riverside become a worse fit to the job market of that city as time progresses. This indicates that Riverside is not serving as a primary destination for many people but an alternative, a holding bin similar to what we saw with City B in BEG (section 5.1).

Further, the natural implication of the holding bin effect is that optimizing fit between labor and job pools is a dynamic, multistage process. Even for those people who ultimately leave Riverside, while they are there, they consume part of the available economic pie out of necessity, unpalatable though it may be. Until they leave, jobs that may be more optimally filled by outsiders are unavailable to those outsiders. And unless there is a more attractive alternative for the suboptimal natives, they will not move. Resultantly, the attractiveness of a locale for a particular individual is not just a function of that individual and the locale but also the individuals already at that locale and the options available to them. Therefore impediments to optimizing the fit of labor to jobs are felt throughout the network, like a clot impeding circulation. To be clear, critical thought would have elucidated this even in BEG, the simplest two location scenario (chap 5), but it is far more apparent in the context of the present multi-locale simulation.

Finally, tables 7.3 and 7.4 show another interesting characteristic of the flows in this model. That is that every city has one or two major destinations, and minimal traffic to the remaining cities. For example, in the case of Mexicali, 67.3% remain in Mexicali (or return by simulation's end), and 49.7% end up in Tijuana, which is the main destination of the majority of persons from Mexicali. The remaining cities (viz. Los Angeles, Riverside, and

San Diego) each receive less than 17.9% of Mexicali's starting population. The inverse is true for Tijuana, with 16.5% ending up in Mexicali and less than 15.9% in any of the US cities.

A similar phenomenon is seen in the American cities, with one additional nuance. In the case of persons originating from Riverside, both Los Angeles and San Diego serve as major destinations, though Los Angeles receives 3 times as many in raw numbers. So it's not necessarily the case that each city only has 1 major destination. Rather, what these links have in common is that they are all within country, which is another way of saying that domestic migration exceeds international migration. This is what we observe in the real world, so is a good thing to see reflected in our simulation results.

7.2 Policy Intervention

Now that we understand how the UMBC scenario generally functions, we can see how the results we've seen thus far change when the cost of crossing the border from Mexico into the U.S. is increased. The upper panel of figure 7.8 is similar to figure 7.3 but it shows population patterns for the restrictive scenario instead of the open. Of the five cities modeled, four appear virtually identical in population patterns whether the border is open or closed (though there are actually differences at a small scale). For example, Los Angeles has an initial population of 390, a mean population of 393 in the open scenario, and mean population 395 in the closed. San Diego, Mexicali, and Tijuana all show similar correspondence between the open and closed scenarios. Riverside is the only city which presents a marked difference in population patterns when the border is restricted. In the open version, it experiences a sharp population increase in the first quarter of the simulation, which it does not in the closed. Resultantly, by the time the population stabilizes in both versions of the scenario, there is a marked difference in population: 54 for the open and 42 for the closed.

The lower panel of figure 7.8 shows just how much of a difference the policy change makes for Riverside, as well as the other cities. This plot shows the difference in population per tick in each city between having a high entrance cost versus a low one. Positive Y-

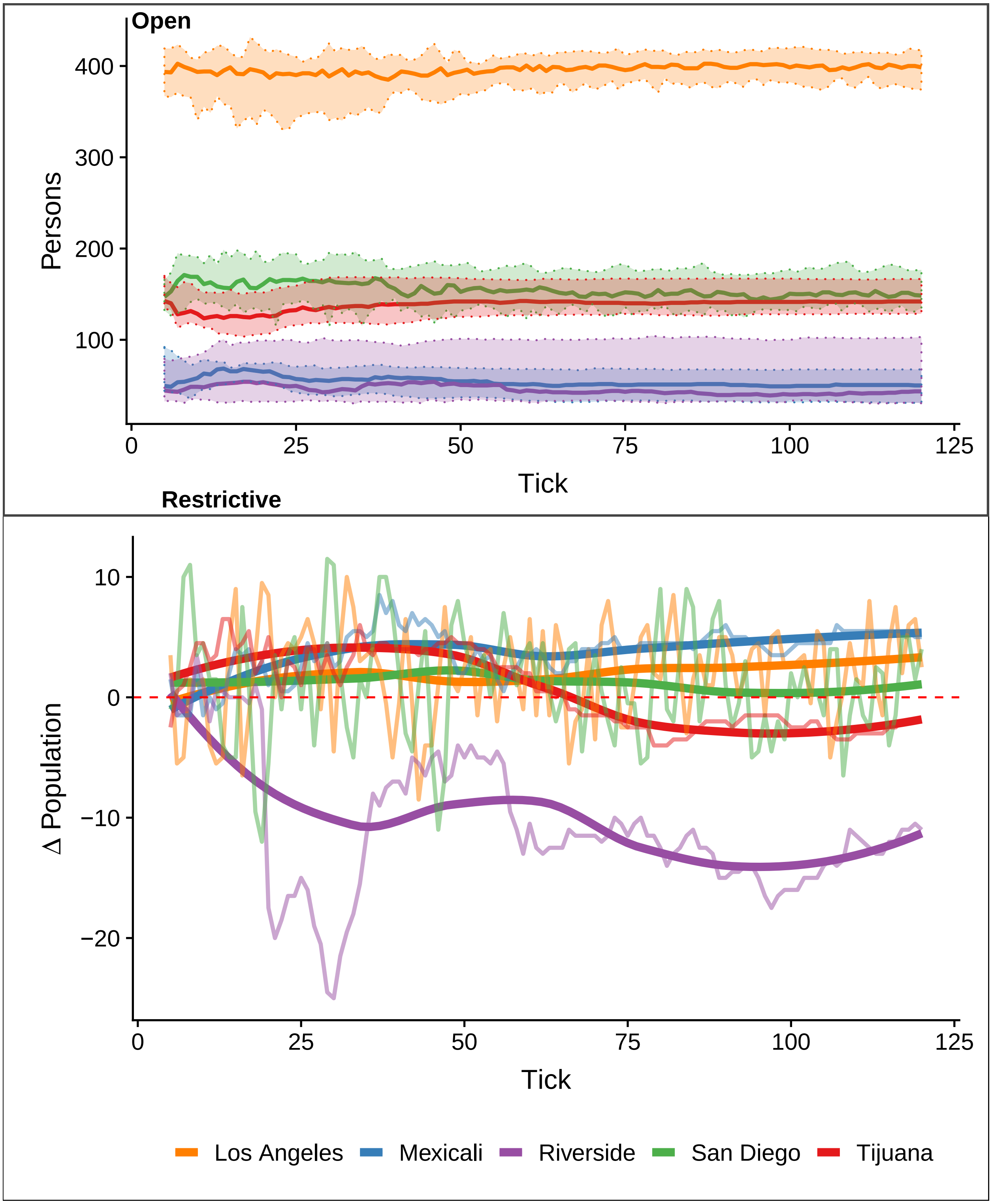


Figure 7.8: UMBC: Changes to Population Due to Restrictive Immigration Policy

axis values indicate more people under the higher cost. With an increased cost of crossing from Mexico to the US, Los Angeles gains 2 on average, an almost imperceptible 1% of its starting population. Mexicali also gains about 4, which is a slightly larger 5% of its starting population. San Diego has a small influx but by simulation's end, essentially returns to where it started. Tijuana ends the slightly significantly down, losing -1 or a paltry -1% of its initial population. As we concluded from the upper panel, the changes for these four cities are more or less insignificant. Not so for Riverside, however. Riverside loses a massive -39% (-12) by simulation's end when the border is closed.

Why is Riverside losing so much of its population under the restrictive regime? We can begin to answer this question by examining figure 7.9, which shows how different segments of the population are impacted by emplacing a higher entrance cost into the U.S. The topmost panel shows the median difference in foreign migrant population by city between the closed and open scenarios, with higher Y-values indicating more under the closed scenario.[2] Predictably, Riverside decreases in number of foreign immigrants, as does Tijuana. Los Angeles starts with fewer but the difference zeroes out by simulation's end, while San Diego remains basically neutral. Mexicali increases in foreign immigrants.

The middle panel of figure 7.9 shows the corresponding information for non-local natives (i.e. domestic migrants). The only country that shows a marked difference between the closed and open scenarios throughout the simulation is Riverside, which hosts markedly fewer domestic migrants. Finally, the bottom-most panel of figure 7.9 shows the population of locals. Los Angeles has a slight surplus and San Diego a slight deficit. There are no other notable differences resulting from the restrictive policy. So from these three panels, we have a better understanding of the patterns we see in figure 7.8. The restrictive policy has the (presumably) intended effect of decreasing the stock of foreign migrants in Riverside compared to the open policy. The policy's effects are mixed for other cities. However, the intervention surprisingly also decreases the number of native migrants who stay in Riverside

[2] *Population* is distinct from arrivals and departures. A positive value of V at tick T in this figure means that between simulation start and tick T, V more foreign migrants entered. It does not mean that V more are entering every tick. Likewise, negative values do not mean that V are departing per tick, but that V more have left between tick 0 and T.

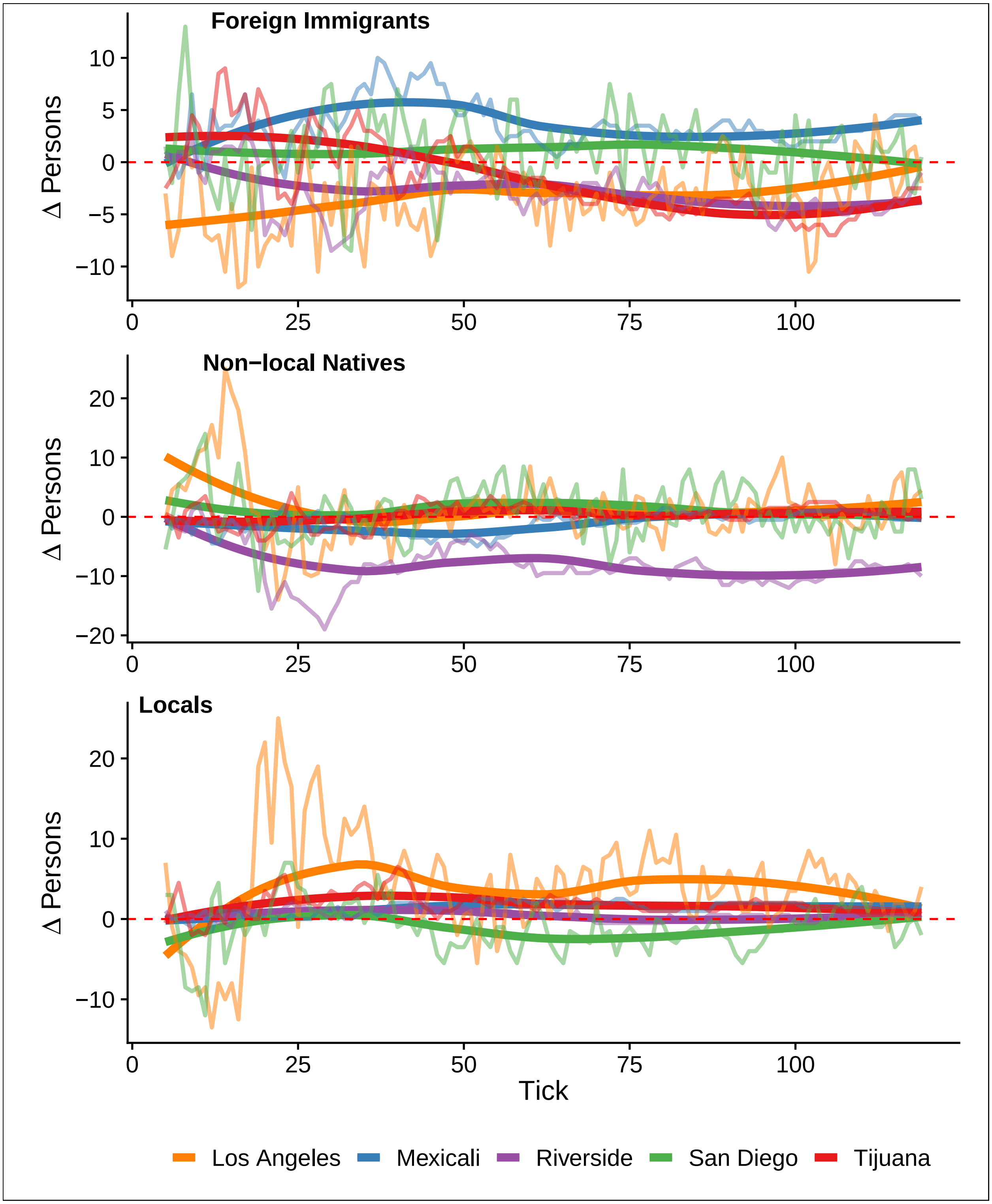


Figure 7.9: UMBC: Impact of High Entrance Cost on Different Sub-Populations

as well. As a result, that city suffers a dramatic overall population deficit under the restrictive regime compared to the open one.

Now, we know that Riverside's lower population is caused by decreases in both foreign and native migrant populations, but we can drill down even further. As we saw in section 7.1, a net change in population can result from a change in in-migration, out-migration, or both. Which is it in this case? In the previous section, we examined net changes in population in the context of network graphs. While that visualization is informative, one downside is that it does not enable us to see how the change appears in the time domain. Additionally, now that we are concerned about making a comparison between two scenarios, it is difficult to use a network graph to present the requisite information. Therefore, here we will use more conventional plots of change over time to address these shortcomings. In figures 7.10 and 7.11, the left column shows the number of arrivals and departures (respectively) over time under the open policy, broken out by city, differentiating between domestic and international moves. The plots in the right columns show the difference over time in each of the left-column figures when the closed policy is enacted. To put it a different way, adding the values in the right column to the corresponding figures on the left results in arrivals and departures under the closed policy.

At first glance, the right columns of figures 7.10 and 7.11 seem to contradict figure 7.9 since figures 7.10 and 7.11 indicate that San Diego and Tijuana decline in arrivals and departures under the closed policy, while figure 7.9 shows them to be unaffected or increase in specific subpopulations. It is important to remember, however, that arrivals and departures are gross measures of *rate* of change in population not amount, so that even with declines in both, net neutral or increase in populations can certainly result.

With that caveat in mind, several points become apparent from these plots. First, regardless whether the entrance policy is open or closed, international departures and arrivals decline dramatically during the initial transition period of the scenario and zero out completely by the steady state period for all five cities. This is seen in the steep downward trajectory to flat line of the "international" line across all left side panels in figures 7.10 and 7.11. In essence, this means that any utility to be gained from international movements are

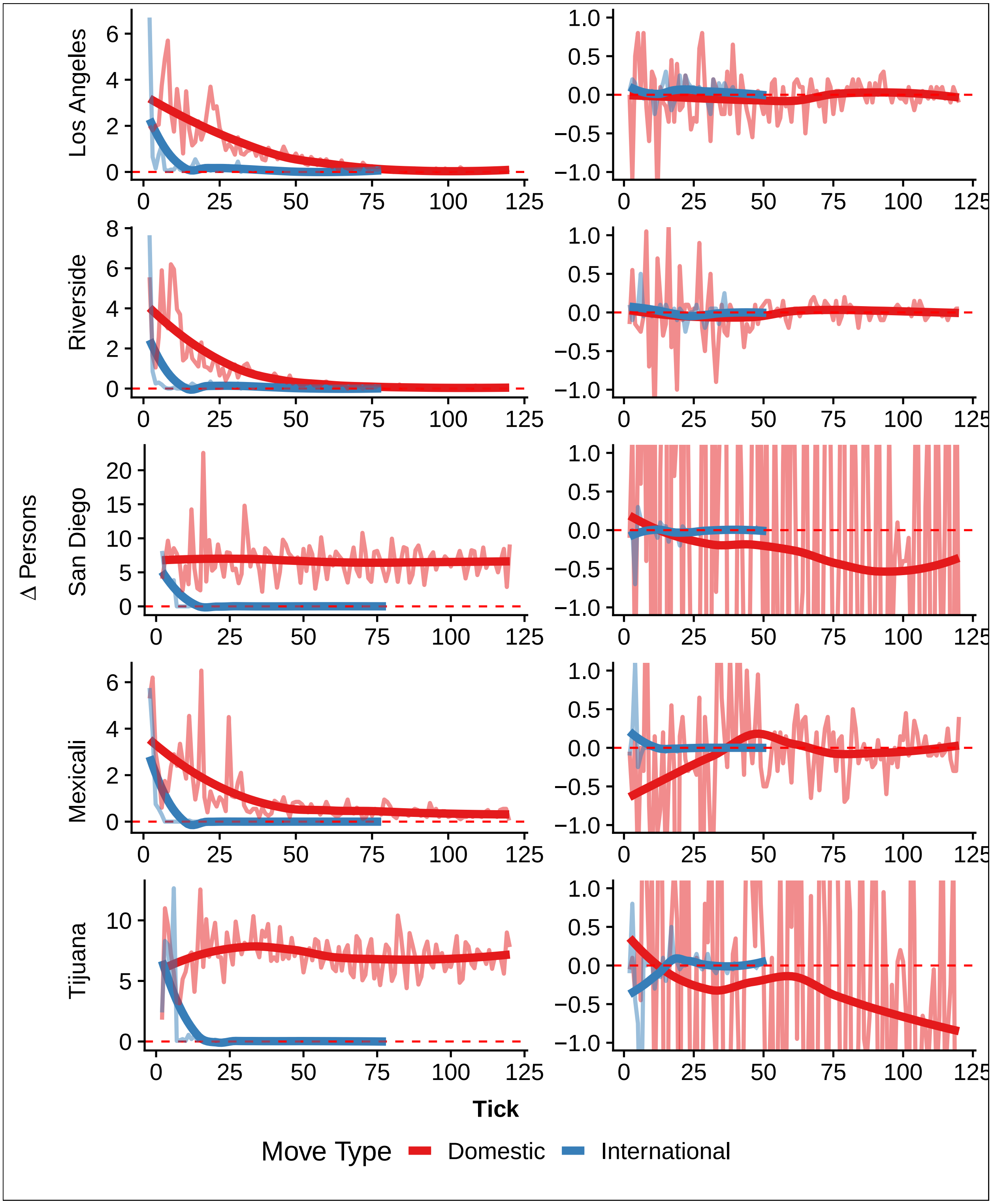


Figure 7.10: UMBC: Change in Arrival Rates Due to Intervention

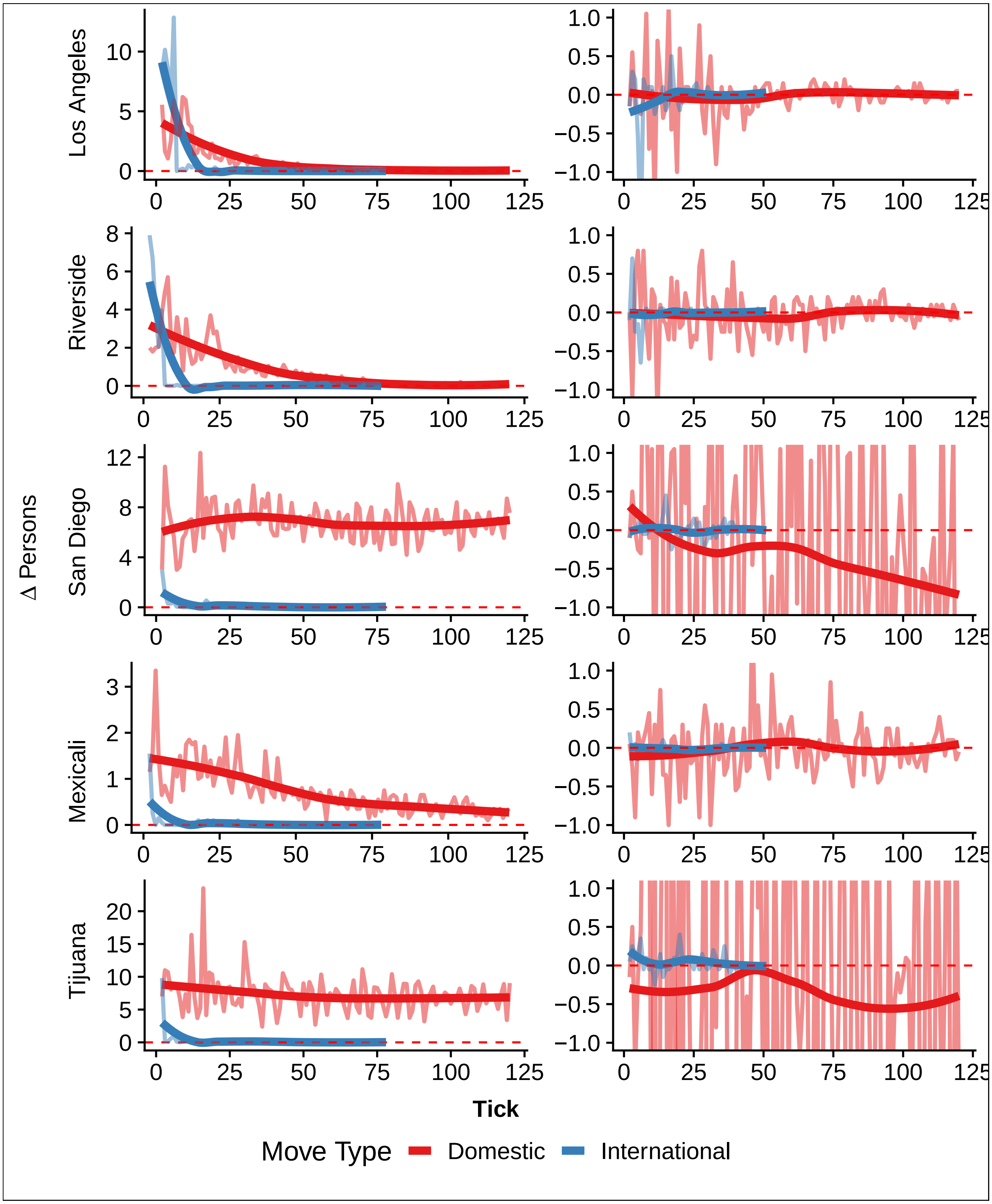


Figure 7.11: UMBC: Change in Departure Rates Due to Intervention

all obtained shortly after simulation start. Logically, this also means that despite the fact that there are many dynamic components to MAPES, none of them are capable of maintaining gradients across the border (beyond what exists from UMBC's initial conditions) sufficient to motivate international movement. We know this is true for both open and closed versions of the scenario because the "international" lines for all right-side panels are basically flat, indicating no change between policies.

On the other hand, domestic departures and arrivals over time do not zero out for all cities in the open scenario (left side panels). In the case of departures (figure 7.11), Los Angeles, Riverside and Mexicali do decline over time, though more gradually than for international movements. San Diego and Tijuana maintain steady departure numbers for the lifetime of the simulation. Note that even though Los Angeles, Riverside and Mexicali do decline over time, they do not zero out. We know this is true because the "domestic" lines for those cities continue to simulation's end, meaning there are arrival and departure events throughout. By contrast, "international" lines stop partway because there are literally no more international moves after a certain point in the simulation. Further, non-zero domestic movements must be true for Mexicali because we know there is domestic movement in and out of Tijuana, the only other Mexican city. By a similar logic, the same must be true for San Diego and at least one of the other two U.S. cities.

As far as the impact of the closed policy on the rate of domestic movement, it is felt mainly for San Diego and Tijuana. In both cases, there are fewer domestic departures and arrivals, while there is basically no impact for the other three cities. Why this happens is a difficult question to answer, and defies common sense to a certain extent. On the one hand, it is intuitive that Los Angeles is basically unaffected by increased entrance costs. It is attractive to both American and Mexican persons since it has the largest number of jobs and mean wage is second only to San Diego. As long as migration of any kind remains rational, it makes sense that that city should continue to motivate movement. However, it is somewhat puzzling that San Diego suffers a significant decline in domestic movement while Riverside does not. First, why does changing the cost of international movements affect the number of domestic moves? Second, Riverside has fewer jobs and a higher cost of living than San

Diego, as well as the lowest mean wage of all the US cities. So one would think that dollar for dollar, San Diego is a better destination than Riverside, yet it is the former that suffers a decline.

There are many possible explanations to the peculiar behavior of San Diego and Tijuana, and the explanation for one does not necessarily need to be shared by the other. One easy answer for San Diego is that Riverside is far closer to Los Angeles, and something about increasing the cost of Mexico to U.S. crossings makes Riverside's distance advantage sufficient to overcome San Diego's larger number of total jobs and higher mean wage. That something could stem the observation from the previous section that San Diego and Tijuana serve as transit hubs for international movement. In this line of reasoning, the increased cost of border crossing decreases the number of moves from Tijuana to San Diego, which decreases the number of job openings in Tijuana for persons in San Diego and therefore decreases reciprocal moves. This in turn makes the labor pool in San Diego more stagnant so fewer Americans from other cities come to San Diego and vice versa. In other words, stagnation begets more stagnation as logjams in one city are felt throughout the migration system. And as transport hubs, San Diego and Tijuana feel the impacts of such logjams more acutely than other cities. If true, this explanation probably is limited in effect given that the impact of the policy intervention on international departures and arrivals in the two affected is small (per figures 7.10 and 7.11).

Another possible explanation is that the decline in movement in and out of the border cities is less about what is going on in them and more what is going on in the rest of their respective countries. This story would be driven by the same mechanisms as above except here, the contagion effect spreads from the periphery to the border rather than vice versa. This too, however, seems limited in scope since domestic movement in those other cities does not appear strongly affected by the closed policy (seen in the relatively flat "domestic" lines of the right-side panels for those cities). So perhaps, it's a combination of the two, or something else completely. The bottom line is that the answer is not obvious and almost certainly complicated given the number of moving pieces in this sytem - a ripe topic for future research.

Since we are now on the topic of unanswered questions, this is an opportune time to mention the big limitation of this section. It is that while I have enumerated in this section *how* populations and flows change due to a policy intervention, I have not really answered the question of *why* the changes occur. There is a difference between delving into subpopulations to show which are causing changes at higher levels of aggregation, and actually explaining the mechanisms that lead to changes at those lower LoAs in the first place. Answering this question requires a different research design which is outside the scope of this chapter but is discussed in chapter 8. Regardless, this examination is valuable in demonstrating that a restrictive policy intervention has differential, surprising effects on cities on both sides of the border.

CHAPTER 8

Discussion

In its conception, the principal aim of this study was to present and validate an ABM that can facilitate the study of international migration by providing a framework within which theories from different fields can be easily reconciled. A more ambitious goal was to make novel substantive contributions to the study of international migration through the analysis of model output. It is my belief that the contents of this study have satisfied both goals.

I presented the details of the MAPES model in chapter 4 and justified its mechanics on the basis of typical modeling considerations and the substantive background provided in chapter 2. I presented the results of running the model in chapters 5-7, demonstrating that its output is internally valid and appears to be externally valid in most regards as well. As to new discoveries, while it cannot be said that there is any single landmark finding, each of the results chapters focused on one to two central themes for which I provided numerous smaller novelties and refinements to existing theory. I will now outline and discuss the implications of these findings.

8.1 Findings

In the most general terms, the main finding of this study is that political, economic, and sociological theories of migration can be modeled simultaneously to give rise to reasonable outcomes that do not diverge seriously from observed reality. When the model is implemented correctly, migration occurs in expected directions given economic conditions, legislators and their constituents respond reasonably to their incentive structures, and social factors shape flows as theory predicts (though nowhere near to the degree of economic and political factors). This conclusion was not a given, even for the limited number of theories modeled here, because the interaction of simple mechanics can lead to complex and unexpected behaviors. Indeed, such emergence is one of the draws of the ABM approach, but it also means that theories that work fine on their own could potentially break each other when modeled together. That they did not is a relief because it provides support for both established theories of immigration mechanics as well as the novel modeling approach I advocate in this study.

8.1.1 Findings: Movement Patterns

Many of the findings deal with the conditions under which movement occurs and who it is that moves. Across all three scenarios (BEG, TE, UMBC), gross movement occurred throughout the simulation even after net movement stopped. This was seen in the plots of population over time (figures 5.2, 6.3, 7.3), where minor fluctuations around some asymptote occur after a larger steady state is reached (the fluctuations indicating movement). This makes sense because migration is ultimately an individual decision and movement should occur so long as it is rational for an individual, independent of whether it is so in the aggregate. While this is not surprising in itself, an interesting consequence of this phenomenon was that net movement occurred at all in the trilateral equal scenario. In that scenario, the three locations start each simulation in what appears to be a steady state in the aggregate but actually transition out as a result of compounding differences from random individual moves (figures 6.1 and 6.3).

While it is easy to overlook, it is also noteworthy that the migration systems in all of these scenarios eventually reached steady state populations, notwithstanding the ongoing small deviations around those values. This provides equivocal support of the overall approach of the macro-economic framework, which frames migration as a reaction to imbalances. These imbalances concern only wages in the classical theory but are generalized to a multi-component utility in the MAPES model.[1] Regardless, when the imbalances disappear, so should net movement. What is really interesting is that while population does stabilize,

[1]Specifically, utility is calculated as a function of wages and social network size.

other metrics do not, including some that should impact population flows. For example, the median ideologies of the legislatures in BEG tend to stabilize midway through the simulation but then show large-scale changes that do not settle by simulation's end (figure 5.6). This should affect population by altering the costs of (1) movement and (2) being a foreign resident, yet it does not, at least in the simulations I observed. It could be the case that the simulations were not run for enough cycles or that the effects are just not big enough to matter. But in either case, this lack of an impact presents a puzzle.

As to who is moving to and from which countries, each scenario presented unique results. In the case of BEG, a major finding was that the types of persons moving differed from country to country as did the character of their moves. For persons from city A (the receiving country), those in the middle of the education distribution tended to move more often per simulation (figure 5.20). However, whether a particular person moved was far from deterministic: nearly all persons from city A had the potential to move, given the right circumstances, regardless of education level (figure 5.18). The same was true for those from city B (the sending country). Where the two cities differed was in the character of moves. More of the movers from A tended to return to their origin country, while more of those from B settled in the foreign land of A (figure 5.19). This makes sense because we know from the population-level plot in figure 5.2 that A is a net gainer in persons. But we also saw through the in-depth examination of individuals (figure 5.21) that this pattern is the result of changes in the probability of moving as a function of country of origin and earned wage. In short, I showed through my analysis of BEG not only how much aggregate movement there is, but also who is moving, when and why.

In contrast to the orderly results of BEG, the most interesting finding from TE regarding movement patterns is that under at least some conditions, initial conditions are basically uninformative. To put it another way, the outcome for any specific country is unpredictable except on average (figure 6.2). On the other hand, once one has some information on the role of that country in a migration system (i.e. after some number of ticks into a simulation), certain predictions can be made about that country as well as the others in the system. For example, as shown in table 6.2, once we know what type of a role a country assumes in its migration system, we can probabilistically predict what other roles will be filled by the other countries in that system. As an interesting side effect of doing so, I also showed that the most common state of countries in TE is not one of equivalence, but one in which there are clear net senders and net receivers (table 6.2).

Finally, I showed in chapter 7 the predicted flow patterns when MAPES is used to model the real-world scenario at the U.S.-Mexico border in Southern California (UMBC). As we expect from the real world, the vast majority of moves are domestic (94.8%) but what net movement there is moves from Mexico to the U.S. (figure 7.6). I also demonstrated that the net flow of persons from one city to another in such a multi-location system is not a linear one. Many persons have significantly long stops in intermediate cities before they reach their ultimate destination (e.g. stopping in Riverside on the way to Los Angeles), and conversely, some cities serve more as rest areas than final destinations (e.g. Tijuana, Riverside). Perhaps the most interesting set of findings from UMBC concern Riverside, which is the smallest American city modeled in that scenario. Observationally, Riverside is the most unstable in outcomes across simulations (figure 7.3). In some runs of the normal scenario (as opposed to the high-cost one), the city more than doubles in size, while in others, it becomes a ghost town. While this large variance is not fully explained, in the normal version of UMBC, the variance is likely due to a combination of different factors including distance from other cities, job endowments and relative size. When the policy intervention is applied to UMBC, Riverside's population maintains its high inter-simulation variance but is also the only city that suffers a constant and significant loss in overall population relative to the normal version of UMBC. In other words, the impact of raising the cost of South to North border crossings is to shrink the population of an already unstable (in terms of population size) American city! This is not exactly the effect that one predicts from such a policy, and demonstrative of the surprising, emergent behaviors that can be studied using MAPES.

8.1.2 Findings: Politics

In the political realm, this study made novel contributions both in model building and in substance. In the first case, much of the discussion in section 5.2 showed how an oversight in the model-building process yielded problematic results, which led to an update that improved the results of later chapters. Specifically, person-agents initially formed social ties with other person-agents with uniform probability. This turned out to be problematic because it allowed the formation of improbable social networks between individuals with ideologies that were extremely different from one another. This in turn resulted in distributions of ideology that regressed to singular means because these extremists moderated each other; this is unrealistic in my opinion. In detecting and diagnosing this problem, I showed how the process of working with MAPES is an iterative process in which the modeler must make reasonable guesses to combine disparate theories into a cohesive whole but those guesses do not always behave as expected on the first pass. When the appropriate adjustments were made (by weighing probability of forming social ties by ideological distance), the resulting distributions of ideology across populations better reflected reality by giving rise to multiple mezzo-clusters, which can be thought of as proto-parties, rather than a single “hive mind” (figures 6.6-6.8).

With respect to substantive findings, there were at least two clear takeaways from BEG. First, at least during certain stages of migration, there is a direct relationship between foreigner inflow and the preference of natives for more restrictive immigration policy. This is demonstrated both directly in the ideology of individuals (figure 5.8) and indirectly in the preferences of their elected officials (figure 5.6). In the real world, we typically associate such preferences only with traditional receiving countries (A) such as the U.S. and France. But in BEG, we saw that such countries do not have a monopoly on the mechanism and that even the traditional sending country (B) experiences the phenomenon once it begins to receive migrants from A. The other political point of interest in BEG is that there is significant room for slippage between the electors and their representatives. This is seen in the clear contrast between figures 5.6 and 5.8 in the later ticks of the simulation. While the ideologies of the populations of both A and B increase monotonically (partly due to the clustering problem discussed in the previous paragraph), those of their legislatures fluctuate widely right up until simulation's end. This means that the legislature is not a perfect representation of the popular will and it also means that the legislature can have a moderating effect on the more extreme tendencies of the electorate. Neither of these is necessarily unrealistic nor undesired, but it is also unclear whether this phenomenon is substantively meaningful or is simply an artifact of model mechanics. If the latter, it is likely due to the static nature of politicians in this version of MAPES (discussed later in section 8.2), but if not, the issue should be further examined in future work.

In the case of the TE scenario, there were also two sets of findings of a political nature. The first: just as population trajectory cannot be predicted for a specific country in TE before actually running the simulation, neither can patterns in ideology. With ideological trajectories, however, there is even more complexity than in dealing with population patterns. In the case of population, once one sees the output from the initial ticks of a simulation, it is not difficult to determine whether a country will become a sending, receiving, or neutral one (figure 6.4). In the case of political ideology, the patterns are not obvious to the naked eye even after the simulation is fully run (figures 6.6 - 6.8). Not only are there more possible distributions of ideology as a whole, there can also be variation within the same population category, even within the same simulation (e.g. two sending countries in the same simulation may have completely different ideological patterns). That is not to say that they are completely random; as shown in figures 6.9 - 6.11, there are definitely relationships between ideological clustering and means on the one hand, and country type (viz. sending, receiving or neutral) on the other. Specifically, (1) sending countries tend to be more open in their attitudes about immigration, but there is more diversity of opinion; (2) receiving countries have tighter consensus on a preference for more closed policies, and (3) neutral countries fall between.

The other set of political findings from TE arose from exploring the relationship between country type and ideological patterns. It is that the ideological trajectory of a population can be significantly affected by both (1) returnees – migrants who return to their country of origin – and (2) random imbalances in the nature of interactions between foreigners and natives. In the first case, I tested the effect that returnees have on political sentiment in their countries of origin in a series of regression models. I found that having more returnees results in an immediate impact on the median ideology of the legislature, causing it to liberalize on the order of 0.1-0.2 ideological points per percent of the population that is returnees. This is balanced out by the fact that each percent of the population that is foreigners increases conservative preferences by 0.1-0.3 points. A corollary of these models was that lagged measures of the number of returnees did not appear to matter, which I interpreted to mean that the returnees did not significantly cause the rest of the population to liberalize via their social ties (more on this in subsection 8.1.4).

In addition to returnees, I found that the difference in the number of negative and positive interactions between natives and foreigners has a clear impact on legislative ideology: more negative interactions in net leads to more conservative preferences. This effect was estimated to be about 0.036 ideology points per extra negative interaction, which is substantial when one considers that the underlying unit ranges between -120.98 and 96. While much of the analysis focused on finding the right measures to prove the points, the ultimate conclusion was that both returnees and interactions between natives and foreigners had clear effects on the ideological trajectories of their countries.

Finally, the main lesson of chapter 7 in the domain of politics is that policy interventions have complicated and potentially unintended consequences on the various components of a migration system. In that chapter, I explored the model's predictions for the consequences of raising the cost of crossing the border from Mexico to the U.S., which one may assume is an action with the relatively straightforward consequence of reducing immigration into the U.S. However, what I showed is that even as the goal was somewhat met, the way in which this occurred was not at all obvious and there were definitely side effects that likely would not have been predicted. Broadly speaking, foreign immigration was reduced at different levels across the cities in the U.S. in a non-obvious way (panel 1, figure 7.9). Los Angeles and Riverside showed lower numbers of foreign immigrants but San Diego did not. Interestingly, Riverside also showed lower numbers of migrants from other American cities (panel 2, figure

7.9) which lowered the overall population of Riverside compared to population if the policy intervention had not been made (panel 2, figure 7.8). To put it another way, Riverside suffered from stagnation of movement between American cities, which was also shown in the case of Los Angeles, which kept more of its native persons (panel 3, figure 7.9). Cities in Mexico were also affected, with Mexicali receiving more American migrants, and Tijuana receiving fewer. While I was not able to pinpoint the reasons for these outcomes, the patterns are suggestive of a sort of logjam effect. That is, by reducing the amount of movement across the border, less movement is made possible within borders as well because the space that would have been created by emigrants is no longer available for foreign or domestic migrants to move into. Stagnation begets more stagnation.

8.1.3 Findings: Economics

While the consequences of the economic aspects of MAPES were felt across all three scenarios, they were explicitly analyzed only in chapter 5. Consequently, my discussion of findings pertaining to economic issues will mainly derive from that chapter. Nonetheless, the general model used across all scenarios directly incorporated the micro-economic approach to migration, and indirectly permitted observation of the macro-economics of migration. That is to say, individual migrants were modeled as agents who made decisions based on micro-economic rationale (among other factors), and this permitted us to see whether the aggregate-level predictions made by macro-economic theory would be produced as a result. The answer was a qualified yes.

In chapter 5, I showed that given (1) individual agents with rational micro-economic decision making and (2) a Harris-Todaro style economic gradient between two countries, persons flow from one country to the other until an equilibrium is achieved. So, in terms of final outcome, the results of chapter 5 met theoretical expectations. However, the model from which those results derived diverged from classical theory in two important ways. First, persons in the classical theory are driven purely by wages (equation 5.2 from Zenou (2006, p.3), Equation 1.4), whereas for my model, each individual also incorporates several non-wage factors into their decision making. These factors include satisfaction deriving from social networks and financial costs deriving from political climate. Later chapters also introduced opportunity costs from multiple potential destinations, and tradeoffs between domestic and international movement. At the very least, a consequence of these additional factors was that the equilibrium level for Equation 5.2 was *not* at unity (figure 5.11), presumably because the other factors act as additional components in the utility calculation. Another finding that potentially resulted from this divergence was that the ratio between the left-hand side and right-hand side of equation 5.2 (which I referred to as the H/T ratio in section 5.3) actually increased toward the equilibrium rather than declining as predicted by the classical theory. This means that the other, non-wage factors must have been significant in their impact, sufficient to overcome the shortfall in utility implied by the unexpected functional form of the H/T ratio over time.

The second important difference is that each location in my model possesses a fully functioning job market whereas in the classical H/T setup, only the receiving location (the city) has a real job market and the sending location is assumed to always have full employment. This difference, combined with my model permitting migration of persons in both directions, meant that while my receiving location met the theoretical expectation of gaining persons in net, both locations were gross recipients of persons from the other and ended up with significant foreign populations – a distinct departure from the classical scenario. This may have also contributed to the unexpected progression of H/T ratio as job opportunities in the sending country counteracted the increased draw of those in the receiving one, which is what figure 5.11 implies.

Notwithstanding these two major divergences, the BEG scenario by and large honored the overall spirit of Harris and Todaro's macro-economic approach to migration. It is not clear that the same can be said for the TE scenario (chapter 6). In that chapter, we saw net movement occur even though there was initially no macro-economic gradient. To the contrary, all the locations were identical, and only over time did any macro-economic differences arise, as a result of rather than to cause net movement. This was possible because movement in MAPES is fundamentally a micro-economic decision, even though it can be observed at the macro-economic level of aggregation as we did for BEG. In the case of TE, even though there were no location-level differences at the outset of each simulation, for certain individuals, it was rational to move to another location for some combination of reasons whether wage-related, social, or political in nature.

Another economic lesson that we learned from MAPES is how migration affects different sections of the job market in both net sending and receiving locations. Figures 5.13 and 5.14 showed clearly that in the classic bilateral situation, the people who are most negatively impacted by migration in terms of job competition are those seeking jobs in the middle of the wage/education distribution. There is definitely competition for lower wage jobs in either country but it tends to be among immigrants rather than between them and natives. At the high end of the spectrum, persons from each country are technically in competition with each other but because there is a system-wide dearth of qualified (in terms of education-level) persons to fill those positions, there is little to no actual competition on a per-job basis. Where the pressure is felt is among middle-wage, middle-education jobs in both countries. Although the sending country has relatively fewer persons with mid-level education than the receiving, they are still joining a pool of labor that is already very densely populated at those levels. So, in effectively joining the job markets of the two locations, we are making competition for those mid-level jobs fiercer on average than before across both locations. On the other hand, for those willing to move, there are now more jobs to apply to, even though competition for each is higher.

As a result of these changes in the competitive landscape, persons in the receiving country generally suffered worse economic outcomes than they would have with no migration. Conversely, those in the sending country did better. We saw this in figure 5.15, which shows the difference in income per individual between running the simulation with migration enabled and running it without. The lifetime gains and losses experienced by persons from B and A, respectively, are more or less evenly distributed across the spectrum of education, as seen in figure 5.16. So one cannot say that any particular group in either country is disproportionately affected. And, the change arising from allowing migration is on the order of approximately 100 monetary units per tick (approx. \$10K / 100 ticks) – minor but not insignificant.

At this point, it is very important to qualify these findings lest they be used to erroneously justify some real-world political agenda. To be clear: the findings of chapter 5 are specific to the parameters defined in BEG. These parameters have some passing but limited resemblance to the real world. It would be utterly invalid for one to extrapolate the results of the economic analyses herein to argue that immigration has a simple negative economic impact on natives of the receiving country. For one thing, all of the agents in BEG (viz. locations, jobs, persons) are stylized to represent the classical, theoretical sending-receiving dyad and do not clearly represent any specific real-world scenario. When a scenario is crafted to imitate a real migration system as in chapter 7, interpretation of results becomes difficult and limitations of the model even more apparent, as we saw in that chapter. Given that the picture quickly becomes cloudy even between BEG and UMBC, attempting to overextend the lessons of BEG to the real world would certainly be a fallacy.

Another reason to be cautious against over-extending the economic lessons of BEG is that MAPES in its current form is limited in its ability to deal with many of the externalities we know to be caused by migration (e.g. job creation, economic growth) and certainly those we don't know about – the unknown unknowns. This may seem a contradiction of the claim that a primary benefit of ABMs as a research method is that it allows for emergence, but it is not. The seminal Sugarscape model obviously cannot model the rise of human civilization from its population of primitive cells, yet no one disputes that it displays emergence. This is because it is understood that emergence is limited in scope to those phenomena which are capable of being displayed within the limited mechanics of that model. Likewise for MAPES, we know that the model, by design, is simply incapable of creating new jobs so to critique the model's emergent capabilities on that basis would be to misunderstand how ABMs work. Nonetheless, this also means that the results that the model yields cannot and should not be used to make over-ambitious claims about migration in the real world, especially in regards to the negative economic impacts of migration.

8.1.4 Findings: Social Factors

While social factors are an integral part of the MAPES model, I did not devote any portion of this study to the explicit examination of the sociological aspects of the results. Nonetheless, I will summarize here how these factors were built into the model and make some general observations on their role in the results.

Of the many sociological theories discussed in subsection 2.1.1.2, MAPES was equipped to deal mainly with two (table 4.1). These are the ideas that social networks reduce the costs of movement (Hatton, Williamson, et al., 2005) and that ties with coethnics provide utility beyond financial concerns by promoting kinship and community, which D. S. Massey (1987) refers to as *paisanaje*. The cost-reduction mechanism was included in the model by discounting the cost of movement as a function of the number of social ties in the destination locale (subsection 4.1.4). Community was modeled by including utility derived from the number of social ties in a locale, with more ties leading to more utility up to a maximum value defined as a scenario parameter. I did not isolate this utility to only coethnic ties because there are other theories that suggest community to matter regardless of origin (e.g. Wells (2004), McLaren (2003)).

The most prominent demonstration of the importance of social factors was hinted at in the matter of population-level ideological trajectories over time. Because person-agents affect the ideology of those with whom they share social ties, the size, rapidity and composition of social networks had dramatic effects on the political outcomes we observed. Further, the lack of ties between two persons can also carry information, as implied by figure 6.16. The absence of a tie could just as well result from two individuals having come to loggerheads as from not having interacted at all. The former would have a conservatizing effect on the polity and enough of these can have real effects at the aggregate level as demonstrated in section 6.2. On the other hand, there are also reasons to discount the possible importance of social ties in affecting ideology. In my examination of the effect of returnees on the legislature's ideology, I did not find evidence that they cause others to liberalize via their social ties. Of course, this is not the same as saying that they actually do not, and the manner in which the effect was examined – by regressing the legislature's median ideology on lagged percent returnees – was ham-handed at best. A more focused examination of this specific effect would need to focus on the returnees' specific social networks, not just the country as a whole. The ambiguity of these hints aside, the broader point is that the role of social ties in propagating ideology definitely merits further examination.

The utility value of social ties has potential to matter a great deal as well. Though I did not examine them either in this study, I base this statement on the fact that social utility comprises no less than half of the utility calculation (figure 4.5)! Unfortunately, even secondarily, we did not see any evidence of how they affect individuals' decision making. But given what we know about the model, it is not implausible that some of the unexpected behavior we observed, such as the decline of migration despite increasing expected wage in BEG, is explained by the utility gains from establishing social ties.[2] Given how important such mechanisms may potentially be, future studies should definitely examine these and other social factors using the MAPES framework.

8.2 Limitations & Future Directions

Although I have shown that MAPES can facilitate discovery in the field of IM, it is clearly not without shortcomings, especially in its debut form as presented in this study. Many of these shortcomings were known a priori, when the model was constructed, but were accepted for the sake of parsimony and interpretability in exploring a method that is clearly in its infancy. Some became apparent only once interpretation of results had begun through the presentation of clearly erroneous or pathologic but not fully understood behavior. Yet others concern improvements to the methodologies used to analyze model output in the results chapters.

Among the shortcomings we knew a priori, perhaps the most glaring is the static nature

[2]In this example, the explanation would be that as the simulation progresses, people grow more social ties wherever they are currently located. So their utility from being in the current location increases, which offsets the wage-based utility gains to be had by moving.

of many agent characteristics and the agents themselves. Among persons, this includes the lack of mechanisms to account for birth and death – events that impact the total number of persons in the simulation. The biggest consequence of these omissions is that it constrains the definition of "winning" and "losing" to a purely economic/utility-based perspective. The philosophical arguments against such a view aside, a better reflection of reality would be to consider these not just as ends in themselves but also as means to the ends of procreation and avoiding death. Allowing persons to procreate and die would also have enabled us to think about winning and losing in terms of groups of individuals – winning groups would increase in size while losing ones would shrink. A further refinement would have been to include a mechanism such as marriage that could also accelerate naturalization, thereby affecting political aspects of the simulation as well.

Another limitation to person-agents in MAPES is that their age and sex are not modeled. We know from the extant literature and as well as raw statistics that both factor heavily into decision making among migrants as well as the institutions regulating movement. Whether considering gender-based job preferences or child-bearing age among female marriage migrants, age and sex can have substantial effects on the size and composition of migrant flows. While observing such mechanics would have been fascinating, these factors were ultimately excluded from the model because they were deemed too complicated for this introductory version of MAPES.

As with people, jobs are another agent-class that had well-known shortcomings because in the current model, they may not be created or destroyed, and wages remain static. Obviously, this is highly unrealistic and poses clear limitations on the lessons we may learn. For example, with the current model, we can only measure economic health through metrics relating to employment – total (un)employment, vacancy duration, and competition among job seekers. If jobs were more dynamic, we also would have been able to examine phenomena such as changes in wages and job creation/loss.

Further, though jobs are differentiated on the basis of required education (which correlates with provided compensation), in the current model there is no concept of market sector (e.g. manufacturing, white-collar, agricultural). Assigning jobs to specific industries would have allowed us to see the effects of migration on each. The cost of doing so, of course, would have been that each industry would have needed to be modeled along relevant axes beyond just required education – a difficult task. As an example, including agriculture as a meaningful, unique industry would potentially require modeling characteristics as far afield such as geography, climate and growing seasons. Clearly such an endeavor would have been outside the purview of this study.

One other problem with the jobs in this model (albeit a relatively smaller one) is that they exist in vacuo with nothing owning them. A more accurate depiction of the real world would have firms as the first tier economic agents that own, create and destroy bundles of jobs. This did not pose a big problem for the current model because jobs were static, but one could imagine the lack of firms to pose a problem were this not the case. If firms were added to the model, jobs could be added or removed from individual firms as a function of their profitability, which seems a lot more realistic than spontaneously creating or destroying jobs without that mediating layer.

In the case of politicians, one major shortcoming in their implementation is that each one's ideal point on immigration is set in stone and no new politicians may join the fray, though they may be voted in and out of power. Although politicians are seeded to represent the entire spectrum of ideology, that their composition is static means that they can represent their constituency only insofar as their initial distribution (normally distributed about the mean of possible ideology) allows. So, if a simulation were to lead to a extremely polarized polity, for example, the pool of available politicians would not be able to represent the polity in the legislature accurately.

An even bigger limitation, of course, is that both politicians and normal persons are modeled as single issue voters. What I have referred to as ideology throughout this study is really just each entity's sentiment on migration. The oft-quoted dictum that "immigration politics makes for strange bedfellows" cannot manifest in this implementation of MAPES because that statement requires at least one other issue to cross-cut the political cleavages caused by migration. By extension, any other political phenomenon that requires more than one political dimension will simply not arise in this model because it is mechanically incapable of representing it.

Finally, the locations in this model have a major limitation in that they are relatively limited in number. Obviously, even in the real world there are only a finite number of locations, but relative to a migration system comprised of only a handful of locations, they still function as essentially unlimited sinks (i.e. destinations) and sources for migration. For example, in chapter 7, I modeled the migration system comprised of 5 cities on either side of the US-Mexico border in Southern California. In total, 800 person-agents were included in the simulation, which is approximately 1/10,000 of the real-world population of that area. At that scale, if the entire world were modeled, there would be 749,200 other person-agents and tens of thousands of other locations at the city scale (depending on how one defines cities). Each of those persons would be a potential migrant into any of the locations in UMBC and each of the locations would be a potential destination. Given the relative scales (800 versus 749,200; 5 versus tens of thousands), they present effectively unlimited sources and sinks, respectively. We have seen in the preceding chapters how pools of potential immigrants and alternative destinations alter individuals' decision making and aggregate patterns. So even though many of those other persons and locations would be largely irrelevant to any particular system (e.g. we really don't care about a person in Tibet and or a city in Norway when considering the US-Mexico border), even at a fraction, there may still be enough others who do matter that enabling the model to account for those tertiary agents could dramatically alter model results.

In each of these cases, the lack of a particular mechanism or feature can be thought of as a lost degree of freedom, which renders the model less dynamic than it would have been with its inclusion. The decreased dynamism makes the model easier to understand, but also reduces the potential for emergent behavior. So while excluding them in this study was a strategic choice made for the sake of tractability, they should be investigated in future work to see how they alter or enhance the findings presented here.

Thus far we have discussed limitations in model scope. The other broad category of shortcomings concerns limitations in methodology, first in the underlying ABM approach and second, in the means by which model results were analyzed. The objection in the former regard is that the ABM approach is self confirmatory to a large extent. Another way to state this critique is that the model is just yielding the results it is programmed to yield. Should it be any surprise that persons flow from low to high wage countries when all the conditions were so clearly programmed to result in this outcome?

One way to respond to this critique begins by considering how ABMs work and their role in social scientific research. As mentioned in section 8.1 of this chapter, it was never a foregone conclusion that many of the behaviors we observed at aggregate levels would actually manifest in the results as predicted by theory. This is for at least two major reasons. First, for theories that weren't explicitly defined at the level of individual agents, agents' decision making rules were inferred from the aggregate level theory and reasonable guesses were made where no explicit theory existed. For example, we know from the literature that social networks help derisk migration – an aggregate level theory. But the way in which those networks form at the level of individual ties is not defined to the point where one can simply translate the theory into a series of modeling statements; so, guesses (albeit reasonable ones) had to be made. It would have been very possible for the mechanics of such theories to have broken in that process of translation.

Second, because the theories that were mixed together in MAPES originated from different literatures, they generally did not take each other into account in their fundamental formulation. For example, while Peter Andreas may have been keenly aware of Harris and Todaro, it is hard to imagine how one might analytically combine his work on smuggling and illegal border crossings with the latters' macro-economic framework. MAPES avoids the issue by joining them in a simulation context, but this potentially introduces its own problems. Namely, with no prior work examining the interactions between theories from such differing corpora, they could have easily broken each other and the simulation as a whole when thrown into the mix together for the first time. Fortunately neither the process of theory translation nor amalgamation led to insurmountable problems for my model, and that is a victory that should not be understated.

The other reason why self-confirmation is not actually a problem is that the role of ABMs is more in the realm of theory testing than theory formulation. Often, persons who are inexperienced in simulation or machine-learning methodologies think that these tools are intended to make new discoveries in an automated fashion. To a certain extent, this is true, especially in endeavors that require operating on large datasets with the potential for complicated multi-part relationships. There are many reported cases in the literature where, in the analysis of some potentially complex dataset, a neural network ML algorithm will uncover some hidden, multi-stage, non-linear relationship that humans are not even capable of conceiving. At the same time, for each such reported victory for these methods, there are almost certainly countless other "findings" that overfit to the data and though they may be "good" models in a mathematical sense, are bunk in a substantive one. It is the role of the human operator to make this determination, and to do so on the basis of theory. Even when the theory ends up being a novel one, it almost always stands on the shoulders of existing theory, building and extending rather than materializing out of thin air. In other words, simulation and ML methods are tools to aid in human-directed theory testing that may, at times, lead to new theories; they should not be viewed as theory generating machines.

Accordingly, the approach I used in this study was to use MAPES as a means of testing existing theories and making incremental augmentations to those theories. With such an approach, self-confirmation is actually a desired outcome since it is a clear answer to the test of a hypothesis. However, even if the model does not yield the expected output, that too may be desirable outcome if it can be established that it occurred due to a substantive reason rather than as an artifact of the method. Either way, a priori theory-building and research design should and did play the central roles in my use of the ABM method, therefore confirming that theories function in the way that they are supposed to was the intended and generally achieved goal.

As to the other major methodological shortcoming of this study, it is that I severely underutilized the experimental capabilities of the ABM method. Indeed, almost all the analyses performed using statistical regression would have been better performed through the experimental method. This is because regression analysis suffers from the fundamental weakness that it is just an indirect means of establishing causality. A wiser use of the capabilities of MAPES would have been to set up mini experiments to directly show causality.

Within an ABM setting, experiments can be time-consuming to set up but are basically unlimited in scope since the modeler controls everything.

As an example, in chapter 6, I analyzed a series of regression tables to show that returnees and native-foreigner interactions affect legislative ideology. To turn that examination into an experiment, theoretically all that would have been required would have been to pause the model at certain points and vary the numbers of returnees and interactions while holding everything else fixed. In practice, this would have been somewhat painful because of the way the simulation was set up, but it was certainly not an impossible task. On the other hand, the challenges encountered in doing so would have purely been computer programming ones, whereas they would have been far more difficult to overcome in a real-world experimental setting and certainly impossible in an observational one.

Another example where the experimental method could have been used is from chapter 7, where I showed *how* raising the cost of crossing the border affected the cities in a particular migration system but not really *why*. To show why, I could have used the experimental method to run simulations with differing levels of border closure, and with many more cities on either side of the border with sufficient variation in characteristics such as population size and composition, number and types of jobs, and distance from the border and other cities, just to name a few. By doing so, I could have increased the degrees of freedom available enough to answer the "why" question, which the single UMBC scenario simply was not able to do with its one effective variable.

Other portions of this study that would have benefitted from such an approach include:

- Measuring the impact of different wage gradients on net flows in the classical migration dyad (section 5)
- Establishing causal direction in the relationship between changes in population composition and population/legislature ideology (section 6.2)
- Showing the direction of the logjam effect when the cost of border crossings in increased (section 7.2)

- Determining the effect of different city and dyad characteristics on the size of flows between two cities (section 7.1).

That I did not use this capability more is a missed opportunity and highlights a major limitation of the current study. Nonetheless, highlighting this weakness is useful because it presents clear paths for future research using MAPES.

Despite all of the shortcomings I listed here, I believe this study was a success, at the very least because of the equally long list of findings described in the previous section. This study was conceived in the belief that the ABM approach (which, at the time of this writing, had not been commonly used to study immigration) could be harnessed to help bring together the disparate literatures that treated the topic of immigration studies. Beyond its discrete findings, this study serves the higher purpose of demonstrating the usefulness of ABMs to that field. And while MAPES itself may be of limited utility, it is my hope that extensions or derivatives thereof, or other models inspired by it will continue to expand our knowledge of this important topic.

Andreas, P. (2000). *Border games: Policing the US-Mexico divide.* Cornell University Press.

Apodaca, C. (1998). Human rights abuses: Precursor to refugee flight? *J. Refugee Stud. 11*, 80.

Bankes, S. C. (2002). Agent-based modeling: A revolution? *Proceedings of the National Academy of Sciences*, *99*(suppl 3), 7199–7200.

Barbosa Filho, H. S., de Lima Neto, F. B., & Fusco, W. (2011). Migration and social networks—an explanatory multi-evolutionary agent-based model. In *Intelligent Agent (IA), 2011 IEEE Symposium on* (pp. 1–7). IEEE.

Barbosa Filho, H. S., Neto, F. B. L., & Fusco, W. (2013). Migration, communication and social networks–an agent-based social simulation. In *Complex Networks* (pp. 67–74). Springer.

Berman, M., Nicolson, C., Kofinas, G., Tetlichi, J., & Martin, S. (2004). Adaptation and sustainability in a small arctic community: Results of an agent-based simulation model. *Arctic*, 401–414.

Billari, F. C., Prskawetz, A., Diaz, B. A., & Fent, T. (2008). The "Wedding-Ring": An agent-based marriage model based on social interaction. *Demographic Research*, *17*(3), 59–82.

Biondo, A. E., Pluchino, A., & Rapisarda, A. (2012). Return migration after brain drain: A simulation approach. *arXiv preprint arXiv:1206.4280.*

Blinder, S. (2011). UK public opinion toward migration: determinants of attitudes. *Migration Observatory Briefing, University of Oxford.*

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*, *99*(suppl 3), 7280–7287.

Borjas, G. J., Freeman, R. B., Katz, L. F., DiNardo, J., & Abowd, J. M. (1997). How much do immigration and trade affect labor market outcomes? *Brookings papers on economic activity*, *1997*(1), 1–90.

Cai, R., Oppenheimer, M. et al. (2013). An Agent-Based Model of Climate-Induced Agricultural Labor Migration. In *Agricultural and Applied Economics Association's 2013 AAEA Annual Meeting, Washington DC*.

Castles, S. (2004). Why migration policies fail. *Ethnic and racial studies*, *27*(2), 205–227.

Castles, S. (2010). Understanding global migration: A social transformation perspective. *Journal of ethnic and migration studies*, *36*(10), 1565–1586.

Cederman, L.-E. (1997). *Emergent actors in world politics: how states and nations develop and dissolve*. Princeton University Press.

Clark, L. (1989). *Early warning of refugee flows*. Refugee Policy Group.

Cohen, R. (1996). Introduction to Theories of migration, by Cohen, R., ed. In *Theories of migration*. Cheltenham: Edward Elgar.

Cornelius, W. A. (2001). Death at the border: Efficacy and unintended consequences of US immigration control policy. *Population and development review*, *27*(4), 661–685.

Cornelius, W. A. & Rosenblum, M. R. (2005). Immigration and politics. *Annu. Rev. Polit. Sci. 8*, 99–119.

Cornelius, W. A. & Tsuda, T. (2004). Controlling immigration: The limits of government intervention. *Controlling immigration: A global perspective*, *3*, 7–15.

Curran, S. R. & Rivero-Fuentes, E. (2003). Engendering migrant networks: The case of Mexican migration. *Demography*, *40*(2), 289–307.

Davenport, C., Moore, W., & Poe, S. (2003). Sometimes you just have to leave: Domestic threats and forced migration, 1964-1989. *International Interactions*, *29*(1), 27–55.

De Haas, H. (2007). Turning the tide? Why development will not stop migration. *Development and Change*, *38*(5), 819–841.

DeVanzo, J. (1980). *Microeconomic Approaches to Studying Migration Decisions*. RAND Corporation. 1776 Main Street, Santa Monica, CA 90401-3208. Retrieved from https://www.rand.org/pubs/notes/N1201.html

Dustmann, C., Preston, I. et al. (2000). *Racial and economic factors in attitudes to immigration*. Centre for Economic Policy Research.

Edmonston, B. & Lee, S. (1992). Why refugees flee: an analysis of Refugee Emigration Data. In *Annual Meeting of the Social Science History Association, at Chicago*.

Ekins, E. (2013). *Support for Immigration Reform Correlated with Perception of Immigrants' Economic Impact*. Reason. 1747 Connecticut Avenue NW, Washington DC 20009. Retrieved from https://reason.com/assets/db/13648288657856.pdf

Ellermann, A. (2005). Coercive capacity and the politics of implementation: Deportation in Germany and the United States. *Comparative Political Studies*, *38*(10), 1219–1244.

Epstein, J. M. (1999). Agent-based computational models and generative social science. *Complexity*, *4*(5), 41–60.

Epstein, J. M. & Axtell, R. (1996). *Growing artificial societies: social science from the bottom up*. Brookings Institution Press.

Espíndola, A. L., Silveira, J. J., & Penna, T. (2006). A Harris-Todaro agent-based model to rural-urban migration. *Brazilian journal of physics*, *36*(3A), 603–609.

Esses, V. M., Dovidio, J. F., Jackson, L. M., & Armstrong, T. L. (2001). The immigration dilemma: The role of perceived group competition, ethnic prejudice, and national identity. *Journal of Social issues*, *57*(3), 389–412.

Fadloullah, A., Berrada, A., & Khachani, M. (2000). *Facteurs D'attraction Et de Répulsion À L'origine Des Flux Migratoires Internationaux: Rapport National, Le Maroc*. Commission européenne, Office statistique des communautés européenne.

Fitzgerald, D. (2008). *A nation of emigrants: How Mexico manages its migration*. Univ of California Press.

Frasure-Yokley, L. & Greene, S. (2013). Black Views toward Proposed Undocumented Immigration Policies: The Role of Racial Stereotypes and Economic Competition. In *Black and Brown in Los Angeles: Beyond Conflict and Coalition* (Chap. 3, pp. 90–111). Berkeley, CA: University of California Press.

Freeman, G. P. (1995). Modes of immigration politics in liberal democratic states. *International migration review*, 881–902.

Fussell, E. & Massey, D. S. (2004). The limits to cumulative causation: International migration from Mexican urban areas. *Demography*, *41*(1), 151–171.

Gardner, M. (1970). Mathematical games: The fantastic combinations of John Conway's new solitaire game "life". *Scientific American*, *223*(4), 120–123.

Gibney, M., Apodaca, C., & McCann, J. (1996). Refugee Flows, the Internally Displaced, and Political Violence (1980-1993): An Exploratory Analysis. *Whither refugee*, 45–66.

Gibson, J. & McKenzie, D. (2009). *The microeconomic determinants of emigration and return migration of the best and brightest: Evidence from the Pacific*. The World Bank.

Givens, T., Freeman, G. P., Leal, D. L., et al. (2008). *Immigration policy and security: US, European, and commonwealth perspectives*. Routledge.

Glazer, N. (1998). *We are all multiculturalists now*. Harvard University Press.

Goldenziel, J. I. (2010). Refugees and International Security. In *On the Move: Migration Challenges in the Indian Ocean Littoral* (pp. 29–42).

Goldsmith, M. M. & Holzner, C. A. (2015). Foreign-born voting behavior in local elections: Evidence from new immigrant destinations. *American Politics Research*, *43*(1), 27–58.

Graves, P. E. (1979). A life-cycle empirical analysis of migration and climate, by race. *Journal of Urban Economics*, *6*, 135–147.

Hainmueller, J., Hiscox, M. J., & Margalit, Y. (2015). Do concerns about labor market competition shape attitudes toward immigration? New evidence. *Journal of International Economics*, *97*(1), 193–207.

Hainmueller, J. & Hopkins, D. J. (2014). Public attitudes toward immigration. *Annual Review of Political Science*, *17*, 225–249.

Hainmueller, J. & Hopkins, D. J. (2015). The hidden American immigration consensus: A conjoint analysis of attitudes toward immigrants. *American Journal of Political Science*, *59*(3), 529–548.

Hansen, R. (2002). Globalization, embedded realism, and path dependence: The other immigrants to Europe. *Comparative Political Studies*, *35*(3), 259–283.

Hansen, R. & Weil, P. (2001). *Towards a European nationality: Citizenship, immigration and nationality law in the EU*. Macmillan.

Hao, L. & Mitchell, C. (2013). Agent-Based Modeling for Rural Migration Decision and Action in China. Paper presented at the 2013 Annual Meetings of Population Association of America, New Orleans, Louisiana, April 2013.

Harris, J. R. & Todaro, M. P. (1970). Migration, unemployment and development: a two-sector analysis. *The American economic review*, *60*(1), 126–142.

Hassani-Mahmooei, B. & Parris, B. W. (2012). Climate change and internal migration patterns in Bangladesh: an agent-based model. *Environment and Development Economics*, *17*(06), 763–780.

Hatton, T. J., Williamson, J. G. et al. (2005). *Global migration and the world economy.* MIT Press, Cambridge, Massachusetts.

Heppenstall, A. J., Crooks, A. T., See, L. M., & Batty, M. (2011). *Agent-based models of geographical systems.* Springer Science & Business Media.

Hinojosa-Ojeda, D. R. & Center, U. N. (2010). Raising the Floor for American Workers. *Center for American Progress, American Immigration Council, January.*

Iyengar, S., Jackman, S., Messing, S., Valentino, N., Aalberg, T., Duch, R., ... Kobayashi, T. (2013). Do attitudes about immigration predict willingness to admit individual immigrants? A cross-national test of the person-positivity bias. *Public opinion quarterly*, nft024.

Jacobs, D., Martiniello, M., & Rea, A. (2002). Changing patterns of political participation of citizens of immigrant origin in the Brussels Capital Region: the October 2000 elections. *Journal of International Migration and Integration/Revue de l'intégration et de la migration internationale*, *3*(2), 201–221.

Jiménez, M. (2008). Parties get sophisticated in bid for immigrant vote. *Globe and Mail.*

Just, A. & Anderson, C. J. (2014). Opinion Climates and Immigrant Political Action: A Cross-National Study of 25 European Democracies. *Comparative Political Studies*, *47*(7), 935–965.

Kalipeni, E. & Oppong, J. (1998). The refugee crisis in Africa and implications for health and disease: a political ecology approach. *Social Science & Medicine*, *46*(12), 1637–1653.

Kandel, W. & Massey, D. S. (2002). The culture of Mexican migration: A theoretical and empirical analysis. *Social forces*, *80*(3), 981–1004.

Kiel, L. D. (2005). A Primer for Agent-Based Modeling in Public Administration: Exploring Complexity in" Would-Be" Administrative Worlds. *Public Administration Quarterly*, 268–296.

King, R. (2002). Towards a new map of European migration. *International journal of population geography*, *8*(2), 89–106.

King, R. (2012a). Geography and migration studies: Retrospect and prospect. *Population, space and place*, *18*(2), 134–153.

King, R. (2012b). *Theories and typologies of migration: An overview and a primer*. Malmö Universtity, Malmö Institute for Studies of Immigration, Diversity and Welfare (MIM).

Klabunde, A. & Willekens, F. [Frans]. (2016). Decision-making in agent-based models of migration: state of the art and challenges. *European Journal of Population*, *32*(1), 73–97.

Koslowski, R. (2004). Possible steps towards an international regime for mobility and security. *Global Migration Perspectives*, *8*.

Kunz, E. F. (1973). The Refugee in Flight: Kinetic Models and Forms of Displacement. *The International Migration Review*, *7*(2), 125–146.

Kunz, E. F. (1981). Exile and resettlement: refugee theory. *International Migration Review*, 42–51.

Langton, C. G. et al. (1989). Artificial life.

Layton-Henry, Z. (1992). *The politics of immigration: immigration, "race" and "race" relations in post-war Britain.* Cambridge Massachusetts/Oxford England Blackwell 1992.

Levin, I. (2013). Political inclusion of Latino immigrants: Becoming a citizen and political participation. *American Politics Research*, *41*(4), 535–568.

Mabogunje, A. L. (1970). Systems approach to a theory of rural-urban migration. *Geographical analysis*, *2*(1), 1–18.

Makowsky, M. (2006). An agent-based model of mortality shocks, intergenerational effects, and urban crime. *Journal of Artificial Societies and Social Simulation*, *9*(2).

Makowsky, M., Tavares, J., Makany, T., & Meier, P. (2006). An agent-based model of crisis-driven migration. *Proc. of the Complex Systems Summer School. New Mexico.*

Malhotra, N., Margalit, Y., & Mo, C. H. (2013). Economic explanations for opposition to immigration: Distinguishing between prevalence and conditional impact. *American Journal of Political Science*, *57*(2), 391–410.

Massey, D. (1999). Space-time,'science' and the relationship between physical geography and human geography. *Transactions of the Institute of British Geographers*, *24*(3), 261–276.

Massey, D. S. (1987). Understanding Mexican Migration to the United States. *American Journal of Sociology*, *92*(6), 1372–1403.

Massey, D. S., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A., & Taylor, J. E. (1993). Theories of international migration: A review and appraisal. *Population and Development Review*, 431–466.

Massey, D. S. & Espinosa, K. E. (1997). What's driving Mexico-US migration? A theoretical, empirical, and policy analysis. *American journal of sociology*, *102*(4), 939–999.

Massey, D. S. & Zenteno, R. M. (1999). The dynamics of mass migration. *Proceedings of the National Academy of Sciences*, *96*(9), 5328–5335.

McLaren, L. M. (2003). Anti-immigrant prejudice in Europe: Contact, threat perception, and preferences for the exclusion of migrants. *Social forces*, *81*(3), 909–936.

Melander, E. & Öberg, M. (2006). Time to go? Duration dependence in forced migration. *International Interactions*, *32*(2), 129–152.

Merolla, J. L., Pantoja, A. D., Cargile, I. A., & Mora, J. (2013). From coverage to action: The immigration debate and its effects on participation. *Political Research Quarterly*, *66*(2), 322–335.

Merton, R. K. (1957). *Social theory and social structure*. Free Press of Glencoe.

Milton, D., Spencer, M., & Findley, M. (2013). Radicalism of the hopeless: Refugee flows and transnational terrorism. *International Interactions*, *39*(5), 621–645.

Mines, R. & Martin, P. L. (1984). Immigrant workers and the California citrus industry. *Industrial Relations: A Journal of Economy and Society*, *23*(1), 139–149.

Moore, W. H. & Shellman, S. M. (2004). Fear of persecution: Forced migration, 1952-1995. *Journal of Conflict Resolution*, *48*(5), 723–745.

OECD. (2010). Public Opinions and Immigration. In *International Migration Outlook 2010*. Paris: OECD Publishing. doi:http://dx.doi.org/10.1787/migr_outlook-2010-6-en

Pantoja, A. D., Ramirez, R., & Segura, G. M. (2001). Citizens by choice, voters by necessity: Patterns in political mobilization by naturalized Latinos. *Political Research Quarterly*, *54*(4), 729–750.

Parunak, H. V. D., Savit, R., & Riolo, R. L. (1998). Agent-based modeling vs. equation-based modeling: A case study and users' guide. In *International Workshop on Multi-Agent Systems and Agent-Based Simulation* (pp. 10–25). Springer.

Patrick, S., Dorman, P. M., Marsh, R. L., et al. (1999). Simulating correctional disturbances: the application of organization control theory to correctional organizations via computer simulation. *Journal of Artificial Societies and Social Simulation*, *2*(1), 1–21.

Peri, G. (2011). *The impact of immigration on native poverty through labor market competition*. National Bureau of Economic Research.

Petersen, W. (1958). A general typology of migration. *American Sociological Review*, *23*(3), 256–266.

Piguet, E. (2008). Climate change and forced migration.

Piore, M. J. (1979). Birds of passage: migrant labor and industrial societies.

Portes, A. (1997). Immigration theory for a new century: Some problems and opportunities. *International migration review*, 799–825.

Price, C. (1969). The study of assimilation. *Migration*, 181–237.

Prietula, M., Carley, K., & Gasser, L. (1998). *Simulating organizations: Computational models of institutions and groups*. The MIT Press.

Pritchett, L. (2006). *Let their people come: breaking the gridlock on international labor mobility*. Ctr for Global Development.

Ramakrishnan, K., Esterling, K., & Neblo, M. (2014). *Illegality, National Origin Cues, and Public Opinion on Immigration*. Working Paper. Ohio State University, Columbus.

Raphael, S. & Ronconi, L. (2007). The effects of labor market competition with immigrants on the wages and employment of natives. *Du Bois Review: Social Science and Research on Race*, *4*(02), 413–432.

Reichert, J. (1981). The migrant syndrome: Seasonal US wage labor and rural development in central Mexico. *Human Organization*, *40*(1), 56–66.

Richmond, A. H. (1988). Sociological theories of international migration: the case of refugees. *Current Sociology*, *36*(2), 7–25.

Rottman, A. J., Fariss, C. J., & Poe, S. C. (2009). The Path to Asylum in the US and the Determinants for Who gets in and Why. *International Migration Review*, *43*(1), 3–34.

Rudolph, C. (2003). Security and the political economy of international migration. *American Political Science Review*, *97*(04), 603–620.

Rudolph, C. (2006). National security and immigration. Stanford: Stanford University Press.

Sassen, S. (1996). Losing Control? Sovereignty in an Age of Globalization. The 1995 Columbia University Leonard Hastings Schoff Memorial Lectures. New York: Columbia University Press.

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of mathematical sociology*, *1*(2), 143–186.

Scheve, K. F. & Slaughter, M. J. (2001). Labor market competition and individual preferences over immigration policy. *Review of Economics and Statistics*, *83*(1), 133–145.

Schmeidl, S. (1997). Exploring the causes of forced migration: A pooled time-series analysis, 1971-1990. *Social Science Quarterly*, 284–308.

Schweitzer, F. (1998). Modelling migration and economic agglomeration with active brownian particles. *Advances in Complex Systems*, *1*(01), 11–37.

Shain, Y. (2010). *The frontier of loyalty: political exiles in the age of the nation-state*. University of Michigan Press.

Shellman, S. M. & Stewart, B. M. (2007). Predicting risk factors associated with forced migration: An early warning model of Haitian flight. *Civil Wars*, *9*(2), 174–199.

Silveira, J. J., Espíndola, A. L., & Penna, T. (2006). Agent-based model to rural–urban migration analysis. *Physica A: Statistical Mechanics and its Applications*, *364*, 445–456.

Singer, A. & Massey, D. S. (1998). The social process of undocumented border crossing among Mexican migrants. *International Migration Review*, 561–592.

Smith, C., Wood, S., & Kniveton, D. (2010). Agent based modelling of migration decision-making. In *Proceedings of the European workshop on multi-agent systems (EUMAS-2010)*.

Soysal, Y. N. (1994). *Limits of citizenship: Migrants and postnational membership in Europe*. University of Chicago Press.

Stanley, W. D. (1987). Economic migrants or refugees from violence? A time-series analysis of Salvadoran migration to the United States. *Latin American Research Review*, *22*(1), 132–154.

Stark, O. & Lucas, R. E. B. (1988). Migration, Remittances, and the Family. *Economic Development and Cultural Change*, *36*(3), 465–481. Retrieved from http://www.jstor.org/stable/1153807

Stewart, J., Caren, N., Quinn, T., & Young, Y. (2013). *An Exploration of Immigrant Political Participation: Placing a Life Course Perspective in Context*. National Center for Border Security and Immigration.

Sun, S. & Manson, S. M. (2010). *An agent-based model of housing search and intraurban migration in the Twin Cities of Minnesota* (Doctoral dissertation, International Environmental Modelling and Software Society).

Takahashi, S. & Goto, Y. (2005). Agent-based simulation of adaptive organizational structures to environmental change. In *Agent-Based Simulation: From Modeling Methodologies to Real-World Applications: Post Proceedings of the Third International Workshop on Agent-Based Approaches in Economic and Social Complex Systems 2004* (pp. 99–110). Springer Japan.

Teitelbaum, M. S. (1995). *Threatened peoples, threatened borders: World migration and US policy*. WW Norton & Company.

Tichenor, D. J. (2002). *Dividing lines: The politics of immigration control in America*. Princeton University Press.

Togeby, L. (1999). Migrants at the polls: an analysis of immigrant and refugee participation in Danish local elections. *Journal of Ethnic and Migration Studies*, *25*(4), 665–684.

Topa, G. (2001). Social interactions, local spillovers and unemployment. *The Review of Economic Studies*, *68*(2), 261–295.

Tumlin, K. C. (2004). Suspect first: How terrorism policy is reshaping immigration policy. *California Law Review*, 1173–1239.

Waldinger, R. & Lichter, M. I. (2003). *How the other half works: Immigration and the social organization of labor*. Univ of California Press.

Wallerstein, I. (1974). The rise and future demise of the world capitalist system: concepts for comparative analysis. *Comparative studies in society and history*, *16*(04), 387–415.

Weiner, M. (1996). Bad neighbors, bad neighborhoods: An inquiry into the causes of refugee flows. *International Security*, *21*(1), 5–42.

Wells, M. J. (2004). The grassroots reconfiguration of US immigration policy. *International Migration Review*, *38*(4), 1308–1347.

Willekens, F. (2013). Agent-based modeling of international migration. *Research plan for independent research group, Max Planck Institute for Demographic Research, Rostock*.

Williams, N. & Pradhan, M. S. (2009). Political conflict and migration: how has violence and political instability affected migration patterns in Nepal. *Population Studies Center Research Report*, (09-677).

Willis, H. H., Predd, J. B., Davis, P. K., & Brown, W. P. (2010). *Measuring the Effectiveness of Border Security Between Ports of Entry*. DTIC Document.

Zenou, Y. (2006). The Harris-Todaro Model.

Zepeda-Millán, C. (2014). Perceptions of Threat, Demographic Diversity, and the Framing of Illegality: Explaining (Non) Participation in New York's 2006 Immigrant Protests. *Political Research Quarterly*, *67*(4), 880–888.

Zhang, Q. & Jager, W. (2011). Agent based modeling of population dynamics in municipalities: Migration in the Derbyshire & Nottinghamshire cases in the UK. *Report for* *the EU Prima project, WP3. University of Groningen, Center for Social Complexity Studies*.

Zinn, S. et al. (2012). A mate-matching algorithm for continuous-time microsimulation models. *International Journal of Microsimulation*, *5*(1), 31–51.

Zolberg, A. R. (1999). *Matters of state: Theorizing immigration policy*. na.

Zolberg, A. R., Suhrke, A., & Aguayo, S. (1992). *Escape from violence: Conflict and the refugee crisis in the developing world*. Oxford University Press on Demand.

Zolberg, A. R. & Woon, L. L. (1999). Why Islam is like Spanish: cultural incorporation in Europe and the United States. *Politics & Society*, *27*(1), 5–38.