Essays in Economic Prehistory

by

Leanna May Mitchell

B.A., Simon Fraser University, 2009

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
Department of Economics
Faculty of Arts and Social Sciences

© Leanna May Mitchell 2021

SIMON FRASER UNIVERSITY
Spring 2021

Copyright in this work rests with the author. Please ensure that any reproduction or re-use is done in accordance with the relevant national copyright legislation.

Declaration of Committee

Name: **Leanna May Mitchell**

Degree: **Doctor of Philosophy**

Thesis title: **Essays in Economic Prehistory**

Committee: **Chair:** Alex Karaivanov
Professor, Economics

Arthur Robson
Supervisor
Professor, Economics

Greg Dow
Commitee Member
Professor, Economics

Bernie Crespi
Examiner
Professor, Biology

Mukesh Eswaran
External Examiner
Professor
Vancouver School of Economics
The University of British Columbia

Abstract

This thesis consists of three papers that explore early human organization. In the first paper I argue that the economic and social structure of early humans would have resulted in an especially difficult consanguinity problem. In particular, adverse selection in the exogamous marriage market would have resulted in high levels of consanguinity and resulting fitness depression. A partial solution to this problem was the evolution of aversion to endogamy, known as the Westermarck effect, and was essential for the survival of our species. The second paper (joint with Haiyun Chen) develops a model that explains linguistic diversity as the cumulative result of strategic incentives faced by linguistic groups. In this model, autonomous groups interact periodically in games that represent either cooperation, competition, or a lack of interaction. Common language facilitates cooperation such as trade, whereas language unique to one group affords that group an advantage in competitive interactions. The relative frequency of cooperation and conflict in a region provides incentives for each group to modify their own language, and therefore leads to changes in linguistic diversity over time. Our model predicts that higher frequency of conflict relative to cooperation will increase a region's linguistic diversity. The third paper (joint with Gregory K. Dow and Clyde G. Reed) investigates the incidence of early warfare among foragers and farmers in prehistory. Our focus is specifically on conflict over land. Food is produced using inputs of labor and land, and the probability of victory in a conflict depends on relative group sizes. The group sizes are determined by individual migration and Malthusian population dynamics. Both factors result in larger populations at better sites, which deters attack. There are two necessary conditions for warfare: high enough individual mobility costs and large enough shocks to the relative productivities of the sites. Together, these conditions are sufficient. In particular, technological or environmental shocks that alter the productivities of sites can trigger warfare, but only if individual agents do not change sites in response. These results are consistent with evidence from archaeology and anthropology.

Keywords: Prehistory, Origin of Preferences, Early Institutions

Dedication

For Tom

Acknowledgements

I would like to thank my family and friends, particularly my parents Bob and Karen Mitchell and my amazing coauthors, Haiyun Chen, Greg Dow, and Clyde Reed. Thank you to: my supervisors, Arthur Robson and Greg Dow; my Graduate Program Chair, Alex Karaivanov; my previous Graduate Program Chair, Simon Woodcock; and the Department of Economics at Simon Fraser University for you tremendous and sustained support. Finally to my husband Tom Cornwall, without whom this work would not have been possible.

I would like to acknowledge financial support from Human Evolutionary Studies Program at Simon Fraser University and from the Department of Economics at Simon Fraser University. Ideen Riahi, Haiyun Chen, and Zhifeng Sun provided valuable research assistance.

Table of Contents

List of Figures

Chapter 1

Endogamy, Exogamy, Fitness Depression and the Evolution of Endogamy Aversion in the Pleistocene

1.1 Introduction

Mating with close relatives, or repeated mating with medium distance relatives within a small population, causes genetic problems that result in harmful effects known as fitness depression (Charlesworth and Charlesworth, 1987; Lacy et al., 1993; Ralls et al., 1988; Charlesworth and Willis, 2009). To avoid this problem, animals have developed various methods for avoiding consanguinity, or for reducing its frequency (Pusey and Wolf, 1996). Humans, as well, appear to have evolved methods for avoiding consanguinity (Van den Berghe, 1983). The mechanism for this is believed not to be cultural (Van den Berghe, 1983), i.e., it is likely based in our genetics.

In this paper I argue that the consanguinity problem would have been especially difficult for early humans, due primarily to our species' unusual social organization. Specifically, the combination of monogamy and group living meant that marriage market candidates would have had excellent information on the quality of their relatives, and poor information on the quality of potential marriage partners from other groups. In this paper I show that this social structure can cause inefficiency and possibly breakdown of the exogamous marriage market, systemic marriage between related individuals, and fitness depression. I then ask the questions, how did humans overcome this problem? What were the consequences?

Monogamy and group living, where multiple pairs of reproducing adults live in close proximity, are possibly the oldest and most fundamental human institutions.[1] This combination is actually unique among primates, (Benshoof and Thornhill, 1979). Committed sexual and life partnership between two people was an important innovation as our ancestors became hunters (Benshoof and Thornhill, 1979). Such an ecological niche requires significant investment in offspring by the father, an arrangement known as "obligate monogamy" (Kleiman, 1977). The decision of whom to 'marry' became critically important for both sexes for the passing on of genes to future generations, and thus there would have been strong evolution by natural selection affecting all aspects of the decision problem.

There is very little direct empirical evidence regarding the marriage practices of our Pleistocene ancestors. Some stylized facts collected from the historical record and from

---

[1] Polygamy is also present in most human societies, but it is a relatively small fraction of reproductive unions (Benshoof and Thornhill, 1979). Polygamous arrangements are similar to monogamous ones in the key assumptions made in this paper - mate choice is very important, both parents are major contributors to the survival of children, and several stable unions of persons live together in groups. Many children are also born from non-monogamous unions, but monogamy is the major human mating system (Benshoof and Thornhill, 1979).

modern hunter-gatherer societies form the basis of the model in this paper. This paper draws on many literatures, but the main one that it contributes to is the evolution of economic behavior. For an introduction to this literature see Robson (2001). In addition, I will rely on biological research on genetics and animal behavior, as well as economic models of adverse selection and signaling.

I begin by collecting some stylized facts from diverse lines of scholarly work regarding the relevant environment. These include pre-Holocene human economics and social organization, fitness depression due to inbreeding in general, and consanguinity avoidance in other animals. I draw some hypotheses from this set of stylized facts. Next, I develop an economic model of a two-sided matching market with endogamy costs. An asymmetric information problem results in the collapse or partial collapse of the exogamous market. I develop a principal-agent model where evolution, or nature, is the principal and individuals are the agents.[2] I show that the evolution of aversion to endogamy would have supported the exogamous marriage market, alleviating the fitness depression problem to some degree. I provide evidence that this aversion is still in evidence today, known as the Westermarck Effect (Westermarck, 1921).

The main purpose of this paper is to explain the Westermarck Effect, (Westermarck, 1921), also known as reverse sexual imprinting. Adolescents and adults are not (or much less) sexually attracted to their childhood intimate associates. Van den Berghe (1983), for example, found that children ages 2-6 seem to imprint negatively on all of their intimate associates, even if they are not closely related. In the model I refer to this aversion to sexual relations with intimate associates, or group mates, as "endogamy aversion" and use an evolutionary framework to derive its fitness-maximizing level, as well as the corresponding equilibrium level of fitness depression.

The theoretical model is a discrete time, heterogenous agents model in an evolutionary setting. Agents, or groups, are heterogenous in their level of fitness depression, and this fitness depression is persistent over time. Each period has five phases. In the first phase, groups draw a level of quality from an i.i.d. distribution. This quality is transient, applicable to the group for only one period. It represents characteristics that tend to be mean-reverting such as skill, private information regarding resources, and the fit of genetic traits with the current environment.

---

[2] For an overview of this approach see Robson (2002).

Each group is randomly paired with another group in a marriage market. Candidates may choose exogamy or endogamy, and exogamy occurs if both parties choose it. Exogamy reduces both groups' levels of fitness depression in the subsequent period, whereas endogamy increases it.

Next, groups produce output. This output can be thought of as children that will form the next generation of descendant groups. If the adults marry endogamously, this output is a function of their group's quality and its degree of fitness depression, which is determined by the history of the groups endogamous and exogamous matches. If they marry exogamously, the output is a function of their own quality and the quality of the group they are paired with, because individuals from that group contribute to group i's production. All groups are assumed to be the same size. High output groups are more likely to become two or more groups, analogous to cell growth, while low output groups are more likely to die out. Next, groups receive utility payoffs. Finally, the fitness depression of the resulting groups updates, and the next period begins.

The main results of the model are as follows. If the level of exogamy aversion is too low, or if the population settles on a bad equilibrium, the level of fitness depression converges to 1 rapidly, and the entire population dies. For mid-range levels of endogamy aversion there will be a threshold quality value; types above this threshold will choose endogamy, and those below it choose exogamy. This threshold, and thus the exogamy rate, is increasing in the level of endogamy aversion. If the level of endogamy aversion is sufficiently high, the entire population marries exogamously and after sufficient time fitness depression disappears.

Next, I endogenize the level of fitness depression and inbreeding aversion and solve for a steady state. There exists steady state equilibrium levels of fitness depression and inbreeding aversion that are robust to invasion from other types. These levels are substantially larger than zero.

Section 2 reviews the most relevant literature, establishes stylized facts, and conducts a thought experiment. Section 3 introduces the model. Section 4 establishes results regarding the relationship between inbreeding aversion and marriage market outcomes, and then endogenizes the inbreeding aversion to establish a steady state. Section 5 concludes.

1.2 Background, Stylized Facts and Thought Experiment

In this section I review relevant literatures and conduct a thought experiment. First, I review the economic literature on incentives for endogamy vs exogamy, then the literature on the genetics of consanguinity for sexual organisms, the costs of consanguinity in humans, the methods other animals use for avoiding consanguinity, why the consanguinity problem would have been especially hard for humans, and how humans avoid consanguinity.

There is a small literature in economics on the costs and benefits of exogamy as opposed to endogamy in small scale societies. Chief among these is Dow et al. (2016), who explain exogamy levels across societies that vary along a number of dimensions. They posit that the main benefits of exogamy are a) exogamous couples are more mobile because they can share knowledge about local environments and access to each other's kin network, and b) the pool of potential partners is larger, allowing heterogenous agents to find better matches than they could in their own community. Using both modeling and empirical techniques, they find that exogamy rate is an increasing function of the productivity gap between communities and a decreasing function of community size.

The focus and conclusions of this paper are different from that of Dow et al. (2016) primarily because I restrict attention to sparsely populated, Pleistocene era mobile foragers, which are only a small subset of the societies that Dow et al. (2016) consider. The limited technology and less productive environmental conditions of the Pleistocene result in smaller community sizes. These communities would therefore have been the most vulnerable to asymmetric information problems and fitness depression. The mobility of households is less important because there are fewer or no property rights to land in this context, so social barriers to accessing land are not a significant issue. If all groups can access all land, exogamy does not increase access to resources unless the group married into has unusually high human capital. In this paper the cost of exogamy is in the uncertainty of partner quality, instead of in isolation from family and familiar territory. Exogamy increases production only if the potential exogamous partner is of higher quality than the endogamous option, after accounting for fitness depression.

A second explanation for exogamy vs endogamy rates in the economics literature is risk aversion and the use of marriage to smooth consumption (Rosenzweig and Stark, 1989; Rosenzweig, 1988b,a). These authors use longitudinal data from surveys of up to 400 households in six agricultural villages in rural India during 1975/1976 and 1984-1985. They find that households with higher profit variance have marriages where the partners come from locations that are further apart geographically, consistent with a consumption smoothing story.

Unfortunately, it is difficult to directly study the role that exogamy and endogamy would have had on fitness depression in Pleistocene-era populations. It is therefore helpful to examine the biological mechanism that causes inbreeding depression and its occurrence in non-human populations.

The offspring of closely related individuals, or closely inbred strains, are generally subject to reduced survival and fertility. This phenomenon is mainly due to an increase in homozygosity, the frequency of identical pairs of alleles (Charlesworth and Willis, 2009). Recessive deleterious alleles that are most often masked in outbred populations occur more often in pairs in inbred populations (Leberg and Firmin, 2008), resulting in infertility, miscarriage, and illness. There is also evidence for over-dominance - being heterozygotic for a pair of alleles results in higher fitness than being homozygotic for either one. Intercrossing (outbreeding) of inbred strains, known as heterosis, produces offspring of increased health and vigour (Charlesworth and Willis, 2009).

Two related measures of inbreeding are used in the biological literature. The first is Wright's coefficient, f, which is the probability for a given individual that the pair of alleles at any base pair on the DNA sequence are identical by descent. Wright's coefficient is also known as the coefficient of inbreeding. If a couple are only related in one way (for example they share one great grandparent and no others), then their f is half the value of their R, the more commonly used relatedness coefficient. The advantage of f is that it can measure complex relationships between individuals, whereas R assumes that apart from the relationship it measures, a pair is outbred.

The DNA at a loci codes for a specific product which is used in some biological process. A deleterious allele produces a mildly or severely dysfunctional product. If a high frequency of the loci contain pairs of alleles that are identical by descent, then there is also a high frequency of pairs of deleterious alleles. A high percentage of the loci will therefore not produce the required product and the organism will suffer.

An individual with a high f will, on expectation, have lower fitness and lower fertility. Genetic disease will be more common, as well as generally less vigour. If she has offspring, however, these will not be subject to the same challenges, so long as their father is not closely related to their mother.

There are three related ways for an individual to have a high Wright's coefficient. One is if its parents have a high R value but are otherwise unrelated. For example, the parents could be half siblings ($R = 1/4$) but otherwise related at the average level of a large population. A second way is if the parents are related in several ways, but are otherwise unrelated. For example, the parents could be double first cousins. A third way is if there is a small population that is sexually isolated for many generations. Over time individuals would become related in many ways. In this third case the population will also lose functional alleles due to genetic drift (Leberg and Firmin, 2008), resulting in a higher f value. Some alleles become fixed - no population diversity at a given locus - and the population thereby loses resilience.

An inbred population may succeed in purging their genetic load. That is, they produce many offspring and those homozygous for the deleterious alleles are less likely to reproduce, thus reducing the frequency of these alleles in the population. In a successful purge, the individuals in the population become less and less fit for several generations, and then fitness increases back to, or close to, initial levels. Many populations, however, fail to purge their genetic load and either die out or remain at a lower level of fitness (Leberg and Firmin, 2008).

In circumstances where competition for resources with other populations is reasonably intense, a small endogamous population is likely to die out before they become healthy again. Purging of minor deleterious alleles works best in populations of several hundred individuals, and not very well in smaller populations, due to drift and fixation of the alleles (Leberg and Firmin, 2008). Furthermore, any effects of inbreeding depression due to over-dominance (cases where a heterozygote has higher fitness than any homozygote) cannot be purged, as inbreeding and drift causes only the loss of useful alleles. Even if the population survives a purge, fixation of deleterious alleles will reduce their fitness relative to outbred populations. At this point, any outbreeding would be beneficial.

The vast majority of our Pleistocene ancestors lived on large continents where competition for resources with other homo sapiens, hominid species and predators would have been significant. Group size would typically have been between 6 and 30. For these reasons, in the model I will assume that endogamy meaningfully reduces fitness at any level of relatedness.

The health costs of a single instance of consanguinity decrease quickly as the f value becomes smaller. (Helgason et al., 2008) study Icelandic couples that have varying degrees of relatedness. They find that couples which are third or fourth cousins have the highest reproduction rate. Third cousins that are otherwise unrelated have an R value of $\frac{1}{128}$, and fourth cousins $\frac{1}{512}$. Consequently, for third cousins, the probability of an offspring having alleles identical by descent at any one loci is less than 0.4%. It is unsurprising then that the offspring of third cousins and fourth cousins are not significantly affected by fitness depression.

Humans face significant consanguinity costs (Charlesworth and Willis, 2009). Bennett et al. (2002) collected measures of risk to a child of a first cousin union, relative to the same statistic for the rest of the population in which the consanguineous unions occurred. One instance of first cousin marriage results in the following observed increases: risk of major malformations increased from 2.1% to 4.1%, risk of congenital anomaly from 0.8% to 2.5%, risk of child mortality from 6.4% to 9.2% in one population and 7.7% to 8.7% in another; in a third population (comprised of 38 subpopulations) there was a 4.4% increase, and in a fourth there was an 8.8% increase of child mortality. McQuillan et al. (2012) found that children of consanguineous marriages are up to 3 cm shorter than those of parents related at the average population level. Probability of a stillbirth or infant death is about twice as high when the parents are consanguineous (Stoltenberg et al., 1999; Kuntla et al., 2013).

It is important to note that socioeconomic factors may be responsible for some of the increases in risk to children from cousin union, and that these can only partially be controlled for by comparing unions from within the same population. When taken cumulatively, however, these health statistics appear to represent a significant biological cost.[3][4]

Animals have various methods for avoiding consanguinity. According to Pusey and Wolf (1996), young are born geographically close to their parents and siblings, so if they were to mate randomly with any opposite gender individual, there would be a very high chance of consanguinity. They argue that a common incest avoidance mechanism is for the young to disperse further than they otherwise would at puberty. Depending on the species the dispersal can be primarily females, primarily males or all individuals. If only the males disperse, for example, when they become sexually mature they would not commonly encounter their mothers, sisters, and most of their aunts and female cousins.

A second method is extra-pair copulation (Pusey and Wolf, 1996). Females mate with several males and the fetuses and infants that are unhealthy die, letting the mother concentrate her resources on the remaining young. Another method is to have very numerous offspring and expend few resources on each one. The weaker ones die, thus purging the genetic load.

Third, some animals have methods for direct recognition of kin so that they can avoid sexual intercourse with them (Pusey and Wolf, 1996); these methods can involve memory and avoidance of mother and litter-mates, which reduce the probability of consanguinity but do not prevent copulation with a father, aunts, cousins, etc.. Some animals with a keen sense of smell are able to sense directly who is a relative and, as a consequence, avoid mating with them. Maturation is sometimes repressed in the presence of an opposite sex parent.

---

[3] In some societies, there may be non-health related benefits to marrying relatives. These generally fall under the categories of control of economic resources or increasing potential for altruism.

[4] If potential partners are sufficiently genetically different, then fitness may start decreasing with decreasing relatedness. This usually only occurs where two populations of a species have adapted to different ecological niches to the point where a hybrid no longer has high fitness in either niche, or is considered less attractive to members of either population, or is less fertile. An example of the latter is the mule, a hybrid of a horse and donkey, which is not fertile.

An important aim of this paper is to establish that the consanguinity problem would have posed a significant problem for our early human ancestors. In the following paragraphs, I pose the hypothesis that the hunting and broad diet ecological niche / economy and its impact on the social structure and life history of our early ancestors would have resulted in a difficult consanguinity problem.

Humans are unusual in that they are the only contemporary primate species to live in groups containing multiple mated pairs and have monogamy as the major mating system (Benshoof and Thornhill, 1979).

Early hominids, such as Australopithecus, are thought to have lived in small multi-male polygynous groups (Benshoof and Thornhill, 1979). During the time of Homo erectus, brain size increased and hunting became more important (Benshoof and Thornhill, 1979). Hunting large game and broad spectrum foraging takes great intelligence and much training and practice (Walker et al., 2002). Human young are born with very large brains and are completely helpless for many years, taking decades to become productive adults (Gurven et al., 2006). Unlike most other primates, a single female would be unable to provide an infant with adequate resources, necessitating significant involvement from her children's father(s) (or other adults) in providing sustenance, care, and education. Relative monogamy would have been beneficial in that it assures relative certainty of paternity. Evolution cannot sustain a system of significant paternal investment in children without fairly high certainty of paternity (Kleiman, 1977).

From this set of circumstances the norm of marriage emerged. From the perspective of evolution, choice of one's marriage partner was vitally important for both sexes. Monogamy and extensive nurturing of young does not allow for such mechanisms as extra-pair copulation as a significant defense against inbreeding.

Due to our ancestral broad diet and hunting niche, human capital, i.e., intelligence, education and experience, was extremely important. Productivity among Ache hunters, for example, peaks ten to twenty years after individuals achieve adult body size (Gurven et al., 2006). Hunting productivity also has high variance. Kaplan et al. (1985), for example, found as much as a two-fold difference in productivity between the best and worst hunters. While the members of any hunting species likely vary in their level of skill, due to differences in size, muscle mass, and historical injuries, these traits are easily observable in a way that human capital is not. This unobservability would compound the asymmetric information problem in the human marriage market.

The methods of hunting that humans adopted were not dependent on a keen sense of smell. It would have been difficult for the relatively undeveloped olfactory system to gain the ability to distinguish genetic relatives from non-relatives. The identity of siblings, aunts, cousins, etc., would be somewhat observable from direct observation and verbal communication, but not completely so. Potential Wright's coefficient would be difficult to predict because an individual can not know directly how long and to what degree the group as a whole has been endogamous, and what the consequences are likely to be of further endogamy. In any case, degree of relatedness is not directly observable to humans, apart from mother-child relatedness.

Relative to many later human societies, early human society had few non-genetic incentives for exogamy. Low resource density meant low population density and relatively open access to land. Therefore, it was not necessary to marry into another group in order to reduce risk by gaining access to their resources.[5] There was likely little or no warfare (Dow et al., 2017). Therefore, exogamous marriages for the purpose of political alliance - to help with offense, defence, or to establish peace, would also have been relatively unimportant.

There are significant advantages to humans of living with relatives, foremost being that evolution can support much higher levels of cooperation between relatives. Sharing of food, assisting with child-rearing, and cooperation on hunts were all important aspects of the early human economy and are more feasible with high within-group relatedness. Considering many such factors, there may be an optimal level of relatedness for spouses, when taking into account co-operative and other possible benefits of marrying a relative. These would, however, only compound the fitness depression problem and enhance the evolutionary importance of a preference for endogamy aversion. In general for humans, the work of both sexes benefits from living near same-gender close relatives, so a system of female or male dispersal would be costly. In practice however many societies are patrilocal or matrilocal, meaning that when marriage takes place one sex tends to remain with their group while the other changes groups.

A tendency for both sexes to remain with one's natal group would result in most potential endogamous marriage partners having a high Wright's coefficient, relative to the average of the population at large. They would likely be aunts/uncles/nieces/nephews or first cousins. This is in contrast to our closest relatives; chimpanzees, bonobos, and gorillas, where individuals also live in social groups but much of the population disperses to other groups at puberty (Van Vuren and Isbell, 1996), for the most part prior to engaging in wooing or competition for mates.

---

[5] Having said this, land ownership also can bring incentives for endogamy due to the efforts of elites to keep the control of land concentrated within a small upper class. Evidence for these efforts is apparent even in prehistoric times (Alt et al., 2013).

The mechanism for human consanguinity avoidance occurs during childhood. Van den Berghe (1983) finds that 2-6 year old humans seem to negatively imprint on their intimate associates. Wolf (1995) studied sim-pua marriages, where the future bride is adopted by the groom's family and they are raised together, and found that intimate childhood association in the first 30 months of life inhibits later sexual attraction. He suggests that "the attachment-caregiving complex and the inbreeding avoidance complex coevolved" in primates.[6]

Schneider and Hendrix (2000) suggest, that olfaction (smelling) is likely involved in the reverse sexual imprinting mechanism. People have some ability to identify relatives and acquaintances by smell and this is stronger for opposite gender relatives (Weisfeld et al., 2003). Humans, however, have notoriously poor senses of smell relative to other mammals and would be unlikely to identify a close relative by smell if that person was previously unknown to them.

Pleistocene human populations were very low. In Europe, for example, (Tallavaara et al., 2015) estimate that between 30,000 years ago and 13,000 years ago population density in the habitable area varied between 2.8 and 5.1 persons per 100 km^2.

In populations with less than one person per square mile (2.59 km^2), the dominant form of organization is the Family-Level group (Johnson and Earle, 2000). Membership of these groups and camps is more fluid and flexible than groups as they are modeled in this paper. They can be as small as 5-8 closely related persons, and aggregate into as many as 25-50 in order to harvest special resources at certain times of year (Johnson and Earle, 2000). Much of the time, the population of a group would be somewhere in between, composed of several related nuclear families.

To develop intuition, consider the following thought experiment, comparing the marriage situation in a mobile society organized at the family/camp level with that of a sedentary village of one thousand people.

In both societies, children and adolescents live with their nuclear family. In the society organized at the family/camp level, nuclear families live for the most part on a home range (Johnson and Earle, 2000). They are sometimes alone, sometimes assembled with one or more closely related nuclear families, who live on ranges close by or overlapping. Once or twice a year, all the people in the area assemble into a group of as many as 50 people.

Children of ages 2-6 years old would become intimate with all of their first cousins, aunts, uncles, and probably first cousins once-removed. All of these people would have significant

---

[6] As it is a proxy for genetic relatedness, it seems likely that this same mechanism may also be at work in some forms of altruistic behavior.

fitness depression if they produced children together later, and thus the Westermark effect, reducing later sexual attraction to these people, is evolutionarily justified. When the child later becomes a marriage market candidate, however, the situation would become problematic. Perhaps 1/5 of the people in the larger group, 10 people, are on the marriage market at any one time. Only 5 of these are of the appropriate gender for our individual, and most are likely first cousins or aunt/uncles, with whom the individual was intimately associated with as a child. They are almost all likely to be related with an f of at least $^1/_{16}$, and thus subject to significant fitness depression upon sexual union.

Consider now that our individual encounters a marriage market candidate in a group from a neighbouring home range which is not at all related to her family. Setting aside sexually selected traits that signal quality, the quality of this person is not known at all to our individual. She might want her nuclear family to go and live with his nuclear family for 6 months to get to know each other, but this is very costly. During this time, one of the groups would be on unfamiliar territory, and their foraging productivity would be reduced. They would be living with unrelated people, and thus unable to benefit from the cooperation benefits that come with relatedness. They may speak a different language, and misunderstandings might arise. The resources available to support 16 people at a time only last a few weeks. The two young people might reproduce before quality had been determined and a ceremony had taken place. In sum, there is considerable pressure to come to a decision quickly, before quality is completely observed.

Next, consider a community of 1000 sedentary people. A young child spends a great deal of time with his extended family, becoming intimate with all of his cousins, aunts, uncles, perhaps 20-30 people. He plays regularly with all the local children, getting to personally know his entire ten year cohort of perhaps 200 people. He learns which families never have enough to eat, which are violent, sick, well fed, highly respected etc. He learns which children are strong, clever, kind, reliable. When he becomes a marriage market candidate, he is not attracted to perhaps 10 intimate associates (many of them relatives) of the opposite sex in his cohort, but there are still 90 possible mates. Signaling of quality is not particularly necessary because he knows their quality directly. Of these he tries to choose the best one possible. With perfect information, the market clears from the top and in the end he settles for someone that is unrelated to him and approximately his own quality. They are efficiently matched as neither can do better on this market, and their biological relationship is distant enough that their children are not afflicted by inbreeding depression. Furthermore, neither family had to make sacrifices to gain information.

1.3 Model

1.3.1 Summary

The model is a discrete time, heterogenous agents model in an evolutionary setting. Agents, or groups, are heterogenous in two variables: their level of fitness depression and their environmental or cultural quality. Fitness depression is persistent over time whereas quality is not. All groups in the population are endowed with a common level of endogamy aversion.

Each period has five phases. In the first phase, groups draw a level of quality from an i.i.d. distribution. This quality is environmental or cultural, but not genetic. In the second phase, each group is randomly paired with another group in a marriage market. This pairing provides each group with an opportunity to marry exogamously. In the third phase, groups produce output as a function of: their quality; their fitness depression; whether they matched with another group; and, if they matched with another group, the other group's quality. High output groups are more likely to succeed in reproducing into two or more groups, while low output groups are more likely to die. In the fourth phase, groups receive payoffs. Finally, the level of fitness depression of the resulting groups updates. If there was exogamy, both groups' levels of fitness depression is reduced for the subsequent period, whereas if there was endogamy it is increased.

The main results of the model concern the the exogamy rate, which individuals will be exogamous, and how these depend on inbreeding preferences. The results can be divided into four parts.

First, I show that the population growth rate is increasing in the degree of exogamy. This is a fairly simple result that follows immediately from the production function, in which the product of endogamous marriages is depressed and that of exogamous ones are not. As a consequence, total exogamy maximizes the growth rate of the population.

Second, I examine the case of a population with a fixed level of endogamy aversion, common to all its members. I show that in any BNE, agents will follow a cutoff strategy, with types above the cutoff marrying endogamously and types below the cutoff marrying exogamously. Expected utility from remaining home is increasing in one's quality, whereas expected utility on the exogamous market is constant in one's quality. Thus, if it is optimal for a certain type to prefer exogamy, it is also optimal for any lower type to prefer it as well. In a PBNE, the cutoff will be the same for all groups. If inbreeding aversion is too low, the cutoff and the exogamy rate will be zero, and the population will die out. The intuition here is that of Akerlof's market for lemons (Akerlof, 1970), in which the market unravels from the top. The type space is continuous, allowing the market to unravel completely until all individuals expect that anyone willing to marry exogamously has quality zero.

After making some additional assumptions on the type distribution, $F(\theta)$, I show that if inbreeding aversion is below $^1/_4$, the exogamous marriage market collapses and the population dies out. The inbreeding aversion in this case is not strong enough to overcome the asymmetric information problem, not even in part. If aversion is above $^1/_2$, exogamy will be universal. In this case, every individual has so much aversion to endogamy that she prefers exogamy, even if the partner has quality zero.

If inbreeding aversion is in an intermediate range, the exogamy cutoff (and therefore rate) is increasing in the degree of endogamy aversion. There is still substantial inefficiency due to asymmetric information, but the solution is interior. Higher levels of aversion to inbreeding mean that a higher quality type is indifferent between endogamy and exogamy, and thus a larger fraction of the population marries exogamously.

Third, I allow for invading groups with different levels of inbreeding aversion. I show that there is a unique equilibrium with positive exogamy that is robust to invaders. As the solution depends on the specific functional form of the distribution of types, no explicit closed-form solution exists, but it is still possible to describe some characteristics of the equilibrium.

1.3.2 Setup

Initial Conditions

Time is discrete and indexed by t. In period 0, there exists a measure of groups, indexed by $i \in [0,1]$. Each group has size $s > 2$, of which two are fertile in any period, and is a measure 0 share of the total population. Each group is endowed with a initial latent level of fitness depression, $\delta_{i0} \in \{0, \epsilon, 2\epsilon ..., 1\}$ (where $\epsilon^{-1} = D \in \mathbb{N}$), which is unobservable and common across all group members.[7] It represents the percentage of alleles in the group that are identical by descent, and its value is determined by the previous instances of exogamy and endogamy of the group, as will be specified in section 1.3.2. Although groups do not know their own level of fitness depression, they are endowed by nature with a degree of endogamy aversion $\lambda \in [0,1]$, which is common to all groups and common knowledge. Each period can be divided into five phases.

---

[7] None of the results presented below depend directly on the size of the group, s. However, it seems reasonable to believe that s and ϵ are linked; the smaller the group the larger the impact of one instance of endogamy on fitness depression.

Draw Quality

In the first phase, each group draws and observes a level of quality, $\theta_{it} \in \mathbb{R}_{+}$, common to all group members. An individual's quality can be interpreted as the quantity of food they can produce. $\theta_{it} \sim F(\cdot)$ is an absolutely continuous random variable, and draws are independent with respect to previous quality draws and the full history of levels of fitness depression. Thus, a high θ_{it} is a transitory advantage for group i.

A group with a high θ_{it} is not fitter in the sense that their genes are absolutely better, nor that the genes are expected to dominate the population over time. Instead, they have a transitory advantage. This could be because they temporarily accumulated more human capital, either technological skill or information regarding resources; were lucky in terms of the amount of food that they were able to collect recently and so are healthier; or their genes are better adapted to environmental conditions that are not expected to persist (for example, it is a cold decade and they are shorter and stockier than the average group).

Marriage

In the second phase, groups have the opportunity to participate in a marriage market. In this market, each group i is randomly paired with another group j. Groups observe neither the quality nor the level of fitness depression of the other group. $y_{it} \in \{0, 1\}$ represents group i's decision on whether to offer to exchange one member from group i marrying exogamously into group j while one member of group j marries exogamously into group i. Consequently, exogamous marriage does not immediately change the size of either group.While groups may have more than two members, I assume that within a given period, only two people from that group (those who were on the marriage market that period) produce offspring. Other group members are assumed to be too young, too old, infertile, repressing fertility etc.

Exogamous marriage, $m_{it} = m_{jt} = 1$, occurs only if both groups agree. $m_{it} = y_{it}y_{jt}$.[8] If one or both groups do not agree, all candidates marry endogamously.

---

[8] As a result of the payoffs introduced in the next paragraph, if one marriage candidate from group i wishes to marry exogamously so will the other one, and if group j accepts one the other will be accepted as well. Assuming one couple will reside with i and the other with j, the marriage phase will not change the size of either group.

Production

In the third phase children, and thereby daughter groups, are produced, in a process analogous to how a cell grows and then splits into daughter cells.[9] The expected number of groups in period $t+1$ descending from group i in period t is[10]:

$$\pi_{it} = \begin{cases} \theta_{it} + \theta_{jt} & \text{if } m_{it} = 1 \\ (\theta_{it} + \theta_{it})(1 - \delta_{it}) & \text{otherwise} \end{cases} \tag{1.1}$$

Thus, there are two inputs in the production of daughter groups: the summation of parental qualities (food produced) and the genetic compatibility of the parents. The expected number of daughter groups is equal to one's own quality plus one's partner's quality, depreciated by the level of fitness depression, δ_{it}, if and only if the marriage is endogamous. The quality of each parent is important because it is represents their ability to produce food, with which to feed themselves and their children. If exogamy occurs, production is increasing in the quality of the partner group j because an individual from that group contributes to the success of the next generation, and thus the reproduction of the group into multiple groups, or of extinction.

The fitness depression coefficient represents the death of some of generation t's children due to genetic challenges, either in utero or as children. Children that do not survive cannot contribute to the formation of daughter groups. The period t children that survive to be the adults in period $t+1$ are those that do not have significant genetic challenges, as evidenced by their survival to adulthood. The higher δ_{it} is, the fewer the surviving number of children that are available to form daughter groups.

The expected number of daughter groups is continuous, but the realization is an integer. This is similar to how expected number of children is commonly expressed as a continuous variable, but its realization is an integer. There is stochasticity in the process by which daughter groups are produced, but the expectation is assumed to be increasing linearly in parental quality so that the distribution of this variable is adequately described by its mean. Although groups are all the same size, successful groups are more likely to have many healthy children and thereby split and multiply into several groups, and unsuccessful groups face a

---

[9] In this way, characteristics that increase group fitness can still come to dominate the population over time.

[10] Suppose the total number of offspring is s times the expression in equation (1.1). Then the number of daughter groups will be π_{it}, provided that the parents die before the next period. If instead, the parents live forever, the number of offspring per unit of θ_{it} implied by equation (1.1) is $s-1$.

higher probability of dying out. I do not explicitly model this process except to specify the expected number of period $t+1$ descendent groups.[11][12]

If endogamy occurs, production of food is heavily dependent on the quality of group i, and the fitness depression factor δ_{it} associated with group i. Fitness depression affects the expected number of daughter groups because it affects the number of offspring that survive to become adults and form new groups. The survivors are assumed not to have any lingering genetic impairment, just the potential for their children to have fitness depression problems if they marry endogamously (but not exogamously). These payoffs are consistent with the empirical finding (discussed in section 2) that a single instance of heterosis (exogamy) will fully restore vigour in the offspring.[13] They are also consistent with the finding that a single instance of endogamy between related persons will result in fitness depression, which is increasing in the degree of relatedness of the parents.

The assumption of a strong positive correlation between one's own quality, θ_{it} and that of an endogamous partner is made because the life histories of both persons would be very similar. They have experienced similar history, technological exposure and environmental knowledge. Their genes are similar, rendering them a similar fit to the current environment. Strong sharing norms in hunter gatherer societies (Kaplan et al., 1985, 2012) ensure that they have been eating approximately equally. Remaining in one's natal group at marriage is also remaining in community with a group of individuals whose quality is highly correlated with one's own quality.

---

[11] One could precisely specify how groups split/die out as follows. If the expected number of daughter groups is π_{it}, the realized number of daughter groups is $\lfloor \pi_{it} \rfloor$ (floor(π_{it})) with probability $(1-(\pi_{it}-\lfloor \pi_{it} \rfloor))$, and $\lfloor \pi_{it} \rfloor + 1$ with probability $(\pi_{it}-\lfloor \pi_{it} \rfloor)$, so that $\lfloor \pi_{it} \rfloor (1-(\pi_{it}-\lfloor \pi_{it} \rfloor)) + (\lfloor \pi_{it} \rfloor + 1)(\pi_{it}-\lfloor \pi_{it} \rfloor) =$
$\lfloor \pi_{it} \rfloor - \pi_{it} \lfloor \pi_{it} \rfloor + \lfloor \pi_{it} \rfloor^2 + \pi_{it} \lfloor \pi_{it} \rfloor - \lfloor \pi_{it} \rfloor^2 + \pi_{it} - \lfloor \pi_{it} \rfloor = \pi_{it}$.
However, for all results in this paper only the expected number of daughter groups is relevant, so other ways of allocating realized numbers of daughter groups to parents could be allowed provided they resulted in the same number of groups on expectation.

[12] In principle, I could create notation to explicitly track the level of fitness depression of each group i, the parents of group i, the period in which i split from its parent, and whether i is still alive at period t. However, the model has only one state variable for each group, δ_{it}, and so I only need to track the fraction of the population at each of a finite number of levels of fitness depression. Consequently, extending the notation would involve considerable work without providing any additional benefit.

[13] Production resulting from exogamy is not subject to fitness depression because the minimum possible number of alleles will be identical by descent, and therefore the offspring will be subject only to the minimum level of genetic disease possible in the population. There will still be exogamous offspring that are homozygotic for deleterious alleles, but for the population in question, these are minimized with a single instance of exogamy. Furthermore, any potential benefits of heterosis will be fully realized.

There is a stark difference in the relevance of fitness depression for endogamous and exogamous marriages, which captures the fitness cost of homozygosity. An individual from a highly-inbred group will produce fewer offspring when matched to a member of their own group, because of the relatively high probability of also matching recessive genes for deleterious traits. However, when matched with a member of another group (in the model, a randomly drawn member of the population), the probability of sharing recessive traits reverts to the lowest possible level.

Surviving to adulthood is a process that screens for major genetic problems in any particular individual, but not for potential problems in endogamous reproduction. For this reason, the fitness depression coefficient is assumed not to apply directly to one's own quality, only to production of offspring if one is endogamous.

(1.1) can equivalently be expressed as:

$$\pi_{it} = m_{it}\left(\theta_{it} + \theta_{jt}\right) + 2\left(1 - m_{it}\right)\left(1 - \delta_{it}\right)\theta_{it} \tag{1.2}$$

Payoffs:

In the fourth phase, groups receive utility payoffs. Payoffs cannot exactly match the level of production of daughter groups because the group's level of fitness depression, δ_{it} is unobservable and varies over time.[14] To address this, I assume that nature assigns a level of inbreeding aversion, λ to the population.

Group i's payoff can be written as:

$$U_{it} = \begin{cases} \theta_{it} + \theta_{jt} & \text{if } m_{it} = 1 \\ 2\left(1 - \lambda\right)\theta_{it} & \text{otherwise} \end{cases} \tag{1.3}$$

$$U_{it} = m_{it}\left(\theta_{it} + \theta_{jt}\right) + 2\left(1 - m_{it}\right)\left(1 - \lambda\right)\theta_{it} \tag{1.4}$$

Where λ is the level of consanguinity or inbreeding aversion, common to all groups in the population. Notice that the utility function is very close to the production function, differing only by having the variable λ instead of δ_{it}. Groups cannot observe the level of δ_{it}, and individuals cannot be expected to understand the genetic mechanism underlying the

---

[14] This may seem like a strong assumption, but I will show below that nature can endow individuals with a level of endogamy aversion such that they behave "as if" they approximately understand the impact of endogamy on fitness depression. Practically speaking, it seems unlikely that early humans had a deep understanding of the mechanics of fitness depression, and even if they did it would be difficult for them to accurately predict the degree of harm it would cause in the case of any particular endogamous match.

impact that endogamy will have on their offspring, and their descendants more generally. Nature endows individuals with aversion to partnering with their group members. The level of aversion would evolve to be a proxy for the cost of partnering with a person that is likely a relative.

There is no conflict of interest between individual marriage candidates and their groups. In what follows I refer to the group as the the decision making agent.

Law of Motion:

In the fifth phase, the fitness depression of groups updates, and the period ends. At each time t, the degree of fitness depression in each period is a discrete state variable. $\delta_{it} \in \{0, \epsilon, 2\epsilon, ..., 1\}$, where $\epsilon = D^{-1}$ for some $D \in \mathbb{N}$. Furthermore, δ_{it} evolves as follows:

$$\delta_{it+1} = \begin{cases} 1 & \text{if } m_{it} = 0 \text{ and } \delta_{it} \geq 1 - \epsilon \\ \delta_{it} + \epsilon & \text{if } m_{it} = 0 \text{ and } \delta_{it} < 1 - \epsilon \\ \delta_{it} - \epsilon & \text{if } m_{it} = 1 \text{ and } \delta_{it} \geq \epsilon \\ 0 & \text{if } m_{it} = 1 \text{ and } \delta_{it} < \epsilon \end{cases} \tag{1.5}$$

Exogamy decreases the level of fitness depression of the group, while endogamy increases it. This definition normalizes the minimum level of fitness depression to 0.[15] Fitness depression also cannot reasonably be greater than one. To see this, suppose $\delta_{it} = 1$ and i marries endogamously in period t. By (1.5), i's expected number of daughter groups is $\pi_{it} = (\theta_{it} + \theta_{it})(1 - 1) = 0$, which means that the group dies out almost surely. If $\delta_{it} = 1 - \epsilon$ however, and group i marries endogamously, the group's expected offspring is $(\theta_{it} + \theta_{it})\,\epsilon$, which is positive and therefore they have a positive probability of producing a daughter group.

Daughter groups inherit their parent groups' δ_{it}, plus or minus ϵ depending on their marriage market outcome in period t.[16] In subsequent periods, sister daughter groups' levels of fitness depression will diverge as they face different outcomes on the marriage market.

---

[15] One interpretion of this minimal level of fitness depression is that it occurs where group i already have the lowest possible frequency of pairs of alleles that are identical by descent, at least in the population, and therefore cannot lower this frequency by further exogamy. This is discussed in more detail in section 2

[16] If there is endogamy, this inheritance is straightforward. If there is exogamy, I assume that one of i's candidates will go to group j and one of j's will go to group i. To simplify the model, we assume that group i passes on its δ_{it}, plus or minus epsilon, to its daughter groups, despite the presence of a member of group j.

Denote the period t distribution of fitness depression coefficients by $\delta_{it} \sim H_t(\cdot)$. In each period, this distribution depends on λ, the distribution of quality, $F(\theta)$, and the initial distribution of fitness depression in the population, $H_0(\cdot)$. In later sections when we introduce the principal-agent metaphor, we assume Nature observes a stable long-run equilibrium population distribution of inbreeding depression coefficients, $H^{\lambda}(\cdot) = \lim_{t\to\infty} H_t(\cdot)$, but not an out of equilibrium period t distribution, $H_t(\cdot)$, or the individual realizations of fitness depression, δ_{it}.

1.4 Results

1.4.1 Population-Maximizing Growth:

I first show a minor result regarding the overall higher productivity of exogamy relative to endogamy. It follows almost directly from the production function defined in equation (1.2).

Proposition 1.1. *The growth rate of the population is weakly increasing in the amount of exogamy.*

Proof. For this proof, I define that the amount of exogamy has increased from one state to another if a) all the groups that were marrying exogamously in the first state continue to do so in the second state, and b) a positive measure of groups that were marrying endogamously in the first state switch to exogamy in the second state.

Consider any two groups that are paired on the marriage market, i and j. Obtaining π_{it} and π_{jt} from equation (1.2), and averaging, obtains the two groups' average growth rate:

$$\frac{\pi_{it}+\pi_{jt}}{2} = \begin{cases} \theta_{it}+\theta_{jt} & \text{if } m_{it}=1 \\ (1-\delta_{it})\,\theta_{it}+(1-\delta_{jt})\,\theta_{jt} & \text{otherwise} \end{cases}$$

The first term is the joint expected offspring from exogamy, while the second term is the joint expected offspring from endogamy. Comparing these, exogamy results in weakly higher growth if:

$$\theta_{it}+\theta_{jt} \geq (1-\delta_{it})\,\theta_{it}+(1-\delta_{jt})\,\theta_{jt}$$

Rearranging, I get:

$$\delta_{it}\theta_{it}+\delta_{jt}\theta_{jt} \geq 0 \tag{1.6}$$

Exogamy therefore gives a weakly higher growth rate for any pair of groups. A positive measure of groups switching from endogamy to exogamy, (and a measure zero switching from exogamy to endogamy), will therefore give a weakly higher population growth rate.[17] □

---

[17] It is possible to prove a stronger version of the above result; that the inequality holds strictly in most cases. Clearly, if both groups have zero quality, $\theta_{it}=\theta_{jt}=0$, equation (1.6) is satisfied with equality. However, this expression holds only for a measure zero fraction of groups. If both $\theta_{it}>0$ and $\theta_{jt}>0$ we can apply the fact that δ_{it} and θ_{it} are independent to sum the above equation over all possible levels of fitness depression, and obtain:

$$\delta_t\left(\theta_{it}+\theta_{jt}\right) \geq 0$$

Here, and in the rest of the paper, I restrict my attention to strategies that do not depend on δ_{t-1} or t.

1.4.2 Inbreeding Aversion and the Exogamy Cutoff

In this subsection I take the population level of endogamy aversion, λ, to be exogenous and establish the PBNE of marriage choices for different ranges of λ. In subsection (1.4.4), the level of endogamy aversion will be endogenized.

In principle, the strategy space for this game is quite large. Each type $\theta \in \mathbb{R}_+$ could offer to marry exogamously with any probability between zero and one. Together, the next two Lemmas show that, under reasonably general conditions, I need only consider a tiny subset of the possible strategies. Specifically, each group will play a strategy defined by a cutoff that is common across groups, where types above the cutoff choose endogamy, while those below it choose exogamy. In the results in this section, I make reference to the following assumption:

Assumption 1. The support of the distribution of θ is an interval with $\inf \theta \geq 0$

Lemma (1.1) shows that all groups will play a cutoff strategy, and Lemma (1.2) shows that the cutoff is the same for all groups.

Lemma 1.1. *For almost all combinations of* $\lambda \in [0,1]$ *and beliefs over the quality distribution of marriage market participants for each group,* $G_{it}(\theta)$*, a BNE requires agents to follow a cutoff strategy* $\bar{\theta}_{it} \in \overline{\mathbb{R}}_+$*. If* i*'s type is realized below the cutoff* i *will prefer exogamy; if above,* i *will prefer endogamy.*

Suppose that after observing its beliefs over the distribution of groups for which $m_{jt} = 1$, $G_{it}(\theta)$, group i chooses a cutoff type $\bar{\theta}_{it}$. To prove the Lemma, I need to show there exists a cutoff type such that after θ_{it} is realized, group i strictly prefers exogamy if $\theta_{it} < \bar{\theta}_{it}$ and strictly prefers endogamy if $\theta_{it} > \bar{\theta}_{it}$.

Proof. Let i's belief about the distribution of groups agreeing to marry exogamously in period t be $G_{it}(\theta)$. Then we can integrate over group i's beliefs in the types willing to marry exogamously, $G_{it}(\theta_{jt})$ to get the expected utility of exogamy. Group i prefers exogamy iff:

Therefore, exogamy results in strictly faster growth unless the population has zero fitness depression. The only way fitness depression could be zero on average is if all groups married exogamously in the previous period.

$$\theta_{it} + \int_0^{\infty} \theta_{jt} dG_{it}\left(\theta_{jt}\right) > 2\left(1-\lambda\right)\theta_{it}$$

Collecting terms, I get:

$$\int\limits_0^{\infty} \theta dG_{it}\left(\theta\right) > \left(1-2\lambda\right)\theta_{it} \tag{1.7}$$

If $\lambda > 1/2$, equation (1.7) holds regardless of beliefs and type (if $\lambda = 1/2$, this holds regardless of type provided $\int_0^{\infty} \theta dG_{it}\left(\theta\right) > 0$). That is, exogamy is always preferred, which can be represented by a cutoff of ∞.

If $0 < \lambda \leq 1/2$, the left-hand side is constant in θ_{it} whereas the right-hand side is increasing on the domain from 0 to ∞, and so a cutoff type must exist (although it may be 0). The sole exception is if $\lambda = 1/2$ and agent i believes only types with quality 0 will participate in the market.[18] $\square$

This result establishes that, if inbreeding aversion is sufficiently high, everyone will choose exogamy. For moderate to low levels of inbreeding aversion, high types will choose endogamy and low types will choose exogamy. This is somewhat counterintuitive to the modern experience, where people who are successful academically or in business tend to travel farther from their hometowns for school and work and thereby encounter and marry spouses from further away.

The next result establishes that the cutoff will be at the same quality level for all groups, and characterizes equilibrium threshold values for each value of inbreeding aversion. It does not, however, rule out multiple equilibria. It also does not guarantee the existence of any equilibrium with positive levels of exogamy. The questions of existence and uniqueness will be addressed later by Proposition 1.3.

Lemma 1.2. *For any $\lambda \in [0,1]$, a PBNE requires groups to follow a common cutoff strategy, $\bar{\theta}_t$, in each period. Types below the cutoff prefer exogamy while types above prefer endogamy.*

---

[18] In this case, the agent would be indifferent between exogamy and endogamy regardless of type, so any strategy is supported. However, even this special case is not a PBNE, and restricting the agent to cutoff strategies would not reduce their expected utility.

Such a cutoff satisfies:

$$\bar{\theta}_t \in \begin{cases} \left\{0, x : (1-2\lambda)^{-1} \frac{1}{F(x)} \int_0^x \theta dF(\theta) = x\right\} & \text{if } \lambda < \frac{1}{2} \\ \{0, \infty\} & \text{if } \lambda = \frac{1}{2} \quad \{0,\infty\} \\ \{\infty\} & \text{otherwise} \end{cases} \tag{1.8}$$

Proof. Lemma 1.1 shows that restricting the set of strategies to cutoff strategies does not restrict the set of BNE, and therefore also does not restrict the set of PBNE. Similarly to Lemma (1.1), group $i's$ problem can be expressed as “choose exogamy if and only if”:

$$\theta_{it} + \int_0^\infty \theta_{jt} dG_{it}(\theta_{jt}) > 2(1-\lambda)\theta_{it} \tag{1.9}$$

The left-hand side is the utility from exogamy, integrated over the distribution of types who choose to enter the marriage market, while the right-hand side is the utility from endogamy, which does not depend on the type distribution. Next, restricting the solution set to PBNEs, all agents must hold the same, correct beliefs about the distribution willing to marry exogamously. Therefore, we can replace $G_{it}(\theta)$ in equation (1.7) with $G_t(\theta)$, all groups' common belief of the distribution of quality for groups willing to marry exogamously. Agents prefer exogamy if and only if:

$$\int_0^\infty \theta_{jt} dG_t(\theta_{jt}) > \theta_{it}(1-2\lambda) \tag{1.10}$$

If $\lambda \geq 1/2$, there is an equilibrium with $\bar{\theta}_{it}$ on the boundary at ∞. In this case $\int_0^\infty \theta dG_t(\theta) = E[\theta]$, and the inequality in equation (1.9) holds strictly. From the perspective of group i when $\lambda > 1/2$, even if the expected value of θ_{jt} is zero, the group still receives a payoff of θ_{it} on the exogamous market, which is greater than the endogamous payoff of $2(1-\lambda)\theta_{it}$, which is strictly less than θ_{it}. Thus, even if one's exogamous option is completely unproductive, if inbreeding aversion is sufficiently high, it is still preferable to marry exogamously. [19]

If $\lambda \leq 1/2$, there is a boundary equilibrium where $\bar{\theta}_{it} = 0$ and thus the left-hand side of equation 1.10 is zero. Since $\lambda < 1/2$, the right-hand side is strictly positive, so endogamy is strictly preferred. If $\lambda = 1/2$, then the right-hand side is zero also. If groups believe that no groups will offer to marry exogamously or only those that have zero quality, all groups

---

[19] It may appear that $\bar{\theta}_{it} = 0$ is an equilibrium even if $\lambda > 1/2$, because if no other group would be willing to match, choosing a threshold of 0 does not make a group worse off. This would be the case if groups chose their thresholds before they drew types. However, groups only behave “as if” they prespecify thresholds; by assumption, they cannot do this and so cannot commit to strategies that are not subgame perfect.

will be indifferent to exogamous vs endogamous marriage. However, in both of these perfect Bayesian Nash equilibria, all groups play the same cutoff.[20]

This equilibrium represents a complete collapse of the exogamous marriage market. If this were to occur it would be disastrous, especially if beliefs about the quality of individuals willing to marry exogamously were persistent across periods.

Finally, if $\lambda < 1/2$, there may be interior equilibria where $\bar{\theta}$ satisfies equation (1.10) at equality, i.e., $x = (1-2\lambda)^{-1} \frac{1}{F(x)} \int_0^x \theta dF(\theta)$. In these equilibria, agents with quality $\bar{\theta}$ are indifferent between endogamy and exogamy, but these groups are a measure 0 share of the population.

□

Population Collapse:

The next preliminary result establishes that if there is no inbreeding aversion or if it is too low, endogamy will happen with certainty every period. Asymmetric information causes the market to unravel completely. Later, Proposition 1.3 will show that with additional assumptions on the functional form, $f(\theta)$, the market unravels and there is no exogamy for all $\lambda < \frac{1}{4}$, and therefore that inbreeding aversion must be substantial to avoid this population collapse.

Proposition 1.2. *Without endogamy aversion (i.e., $\lambda = 0$), the unique PBNE is for all individuals to marry endogamously. The population then dies out within D+1 periods.*

Proof. From equation (1.10), plugging $\lambda = 0$ into the right side, and setting $G_t(\theta) = \frac{\int_0^{\bar{\theta}} \theta dF(\theta)}{F(\bar{\theta})}$, agents marry exogamously iff:

$$\frac{\int_0^{\bar{\theta}} \theta dF(\theta)}{F\left(\bar{\theta}\right)} > \bar{\theta}$$

But this cannot be true for any $\bar{\theta}_t > 0$ because all absolutely continuous random variables with finite expectation have a mean that is strictly less than their maximum.

The second part of the proposition follows directly from equation (1.5). Complete endogamy implies $m_{it} = 0 \ \forall \ i, t$. This, and equation (1.5) imply $\delta_{it} = \min\{\delta_{it-1} + \epsilon, 1\}$. Working this backwards, I get:

---

[20] In these equilibria, it is possible for a measure zero set of agents to play some other cutoffs, subject to some conditions on the distribution of groups who do so. In what follows I ignore these equilibria because the population growth rate is unaffected.

$$\delta_{iD} = \min\left\{\delta_{i0} + D\epsilon, 1\right\} = \min\left\{\delta_{i0} + 1, 1\right\} = 1$$

After D periods, all surviving groups (those with $\delta_{i0} = 0$) will be completely inbred. Any groups still existing at time D will again marry endogamously. Because a completely inbred, endogamously-marrying, group dies with probability 1, these remaining groups also die. □

Corollary 1.1. *If $\lambda < 1/2$ and all groups, every period, believe that only types of quality zero will choose exogamy, all groups choose endogamy and the population dies out within D+1 periods.*

Proof. Follows directly from equations (1.8) and (1.5). □

To rule out the disastrous equilibria and focus attention on a viable one whenever it exists, I make the following definition and assumption:

Definition 1.1. Let $\kappa(\lambda) := max\left\{\bar{\theta}(\lambda)\right\}$ where $\bar{\theta}(\lambda)$ is a PBNE corresponding to a given λ.

Assumption 2. For a given λ, a population will always play the equilibrium cutoff that maximizes the population growth rate. By Proposition 1.1 this is the equilibrium with the highest supportable cutoff, $\kappa(\lambda)$.

For some distributions and values of endogamy aversion, multiple interior equilibria exist. From now on, I focus on the highest possible cutoff type, $\kappa(\lambda)$, that can be sustained for a given level of endogamy aversion, λ. This focus is motivated by an appeal to a weak form of natural selection at the population level: if there were multiple distinct populations endowed with endogamy aversion λ, the population that coordinated on the $\kappa(\lambda)$ cutoff would be less inbred in the long-run, and would therefore grow faster over time, replacing other populations.

In order to rule out multiple interior equilibria, it is necessary to make an assumption on the quality distribution:

Assumption 3. Let $R = \frac{1}{\bar{\theta}F(\bar{\theta})}\int_0^{\bar{\theta}} \theta dF(\theta)$. The quality distribution $F(\theta)$ satisfies the following property:

$$\frac{dR}{d\bar{\theta}} < 0 \tag{1.11}$$

i.e.

$$f\left(\bar{\theta}\right)F\left(\bar{\theta}\right)\bar{\theta}^2-\left[F\left(\bar{\theta}\right)+\bar{\theta}f\left(\bar{\theta}\right)\right]\int\limits_0^{\bar{\theta}}\theta dF\left(\theta\right)<0$$

As I will show in Proposition 1.3, this assumption insures that higher levels of inbreeding aversion do in fact lead to more exogamy and less fitness depression. i.e., that $\frac{d\kappa(\lambda)}{d\lambda}>0$. Without it the problem is less tractable.

One final assumption:

Assumption 4. $F\left(\theta\right)$ is concave. i.e. $f'\left(\theta\right)\leq 0$.

Lemma 1.3. *Any i. Uniform distribution, ii. Pareto Distribution, or iii. Exponential distribution satisfies assumptions 4 and 3, i.e.* $\frac{dR}{d\bar{\theta}}<0$, *where* $R=\frac{1}{\bar{\theta}F\left(\bar{\theta}\right)}\int_0^{\bar{\theta}}\theta dF\left(\theta\right)$

substituting x *for* θ*,*

Proof. For the three cases:

i. Consider a uniform quality distribution, $F\left(x\right)=\frac{x-a}{b-a}$ on $[a,b]$. $f\left(x\right)=\frac{1}{b-a}$. In this case, $R=\frac{1}{\bar{\theta}F\left(\bar{\theta}\right)}\int_0^{\bar{\theta}}\theta dF\left(\theta\right)=\frac{1}{2}+\frac{a}{2x}$, which is decreasing in x on $[a,b]$.

ii. Consider the Pareto Distribution $F\left(x\right)=1-x^{-\beta}$ for $x\geq 1$ and $\beta>1$. Hence $f\left(x\right)=\beta x^{-\beta-1}$. $\frac{1}{\bar{\theta}F\left(\bar{\theta}\right)}\int_0^{\bar{\theta}}\theta dF\left(\theta\right)$ is then $\frac{1}{x\left(1-x^{-\beta}\right)}\int_1^x x\beta x^{-\beta-1}dx$. Integrating, it becomes $\frac{\beta}{x\left(1-x^{-\beta}\right)}\left(\frac{1-x^{1-\beta}}{\beta-1}\right)$. This function is decreasing in x (for all $x>1$) if $\left(\beta-1\right)x^{1-\beta}-1+x^{1-\beta}-\left(\beta-1\right)x^{-\beta}=\varphi\left(x\right)<0$. $\varphi\left(1\right)=0$, and $\varphi'\left(x\right)=-\beta\left(\beta-1\right)x^{-\beta-1}\left(x-1\right)<0$, so $\varphi\left(x\right)<0$ for all $x>1$.

iii. Finally, consider the Exponential Distribution, $F\left(x\right)=e^{\lambda x}$ for all $x>0,\lambda>0$. $R=-\frac{1}{e^{\lambda x}-1}+\frac{1}{\lambda x}$. $\frac{dR}{dx}=\frac{\lambda e^{\lambda x}}{\left(e^{\lambda x}-1\right)^2}-\frac{1}{\lambda x^2}=\frac{\left[\lambda xe^{\frac{\lambda x}{2}}-\left(e^{\lambda x}-1\right)\right]\left[\lambda xe^{\frac{\lambda x}{2}}+e^{\lambda x}-1\right]}{\left(e^{\lambda x}-1\right)^2\lambda x^2}$. Both the denominator and the second multiplier in the numerator are positive. Consider the sign of the first multiplier in the numerator, letting $t=\lambda x$, $t>0$. $\frac{d\left(te^{t/2}-e^t+1\right)}{dt}=e^{t/2}+\frac{t}{2}e^{t/2}-e^t=e^{t/2}\left(1+\frac{t}{2}-e^{t/2}\right)$. $1+\frac{t}{2}-e^{t/2}$ is decreasing and always negative on $t>0$. Thus, the first multiplier, $\lambda xe^{\frac{\lambda x}{2}}-\left(e^{\lambda x}-1\right)$, is always decreasing on $t>0$. Since it equals to zero when $t=0$, we have shown that it is negative on the domain. Thus $\frac{dR}{dx}<0$. □

Now we are ready to show that the cutoff, and thus the proportion of the population marrying exogamously, is weakly increasing in the level of inbreeding aversion.

Proposition 1.3. *Under assumptions 1, 2, and 3,* $\kappa\left(\lambda\right)$ *is 0 below* $1/4$*, strictly increasing on the domain* $\lambda\in\left[1/4,1/2\right]$*, and infinite above* $1/2$*.*

Proof. From equation (1.8) $\kappa(\lambda) = \infty$ if $\lambda > 1/2$.

Also from equation (1.8), any finite equilibria satisfies $\frac{1}{F(\kappa(\lambda))}\int_0^{\kappa(\lambda)} \theta dF(\theta) = (1-2\lambda)\kappa(\lambda)$. Towards a contradiction, suppose $\kappa(\lambda) > 0$ and $\lambda < 1/4$. Then $(1-2\lambda) > \frac{1}{2}$, so $2 > \frac{1}{1-2\lambda}$, so $\frac{2}{F(\kappa(\lambda))}\int_0^{\kappa(\lambda)} \theta dF(\theta) > \kappa(\lambda)$. But given Assumption (4), the p.d.f. of the distribution does not slope upward anywhere, and so its expected value cannot be greater than half its maximum. Therefore, if $\lambda < 1/4$, $\kappa(\lambda) = 0$.

Now suppose $0 < \bar{\theta}_{it} < \infty$. Then, from equation (1.8) $\lambda < 1/2$ and, plugging in the true value of $G_t(\theta)$ (the common belief on the distribution of quality among exogamously marrying groups), I get:

$$\bar{\theta}_{it} = (1-2\lambda)^{-1}\frac{1}{F\left(\bar{\theta}_t\right)}\int_0^{\bar{\theta}_t} \theta dF(\theta)$$

Rearranging:

$$\frac{1}{F\left(\bar{\theta}_t\right)}\int_0^{\bar{\theta}_t} \theta dF(\theta) - \bar{\theta}_{it}(1-2\lambda) = 0 \tag{1.12}$$

This is the first order condition of group i's utility maximizing problem. At the threshold $\bar{\theta}_{it}$, increasing $\bar{\theta}_{it}$ potentially costs the value of the foregone endogamous match, $\bar{\theta}_{it}$. This value is deflated by $(1-2\lambda)$ because both parties in the endogamous match will be subject to a fitness depression penalty. The second order condition of equation (1.12) w.r.t. $\bar{\theta}_{it}$ is $-(1-2\lambda)$, verifying that if the F.O.C. is satisfied, the optimum is global.

Taking equation (1.12) and setting $\bar{\theta}_{it} = \bar{\theta}_t$, rearranging, and dropping the t subscript:

$$(1-2\lambda) = \frac{1}{\bar{\theta}F\left(\bar{\theta}\right)}\int_0^{\bar{\theta}} \theta dF(\theta) \tag{1.13}$$

Taking the derivative of equation (1.13) with respect to $\bar{\theta}$:

$$-2\frac{\partial\lambda}{\partial\bar{\theta}} = \frac{f\left(\bar{\theta}\right)F\left(\bar{\theta}\right)\bar{\theta}^2 - \left[F\left(\bar{\theta}\right) + \bar{\theta}f\left(\bar{\theta}\right)\right]\int_0^{\bar{\theta}} \theta dF(\theta)}{F\left(\bar{\theta}\right)^2\bar{\theta}^2}$$

Rearranging:

$$\frac{\partial\lambda}{\partial\bar{\theta}} = \frac{\left[F\left(\bar{\theta}\right) + \bar{\theta}f\left(\bar{\theta}\right)\right]\int_0^{\bar{\theta}} \theta dF(\theta) - f\left(\bar{\theta}\right)F\left(\bar{\theta}\right)\bar{\theta}^2}{2F\left(\bar{\theta}\right)^2\bar{\theta}^2}$$

Because $F\left(\bar{\theta}\right)$, $\bar{\theta}$, $f\left(\bar{\theta}\right)$, and $\int_0^{\bar{\theta}} \theta dF(\theta)$ are all continuous, and $F\left(\bar{\theta}\right)$ and $\bar{\theta}$ are both strictly positive $\forall\bar{\theta} > 0$, $\frac{\partial\lambda}{\partial\bar{\theta}}$ is continuous $\forall\bar{\theta} \geq 0$. In accordance with assumption (3), $f\left(\bar{\theta}\right)F\left(\bar{\theta}\right)\bar{\theta}^2 - \left[F\left(\bar{\theta}\right) + \bar{\theta}f\left(\bar{\theta}\right)\right]\int_0^{\bar{\theta}} \theta dF(\theta) < 0$, and therefore $\frac{\partial\lambda}{\partial\bar{\theta}} > 0$.

What is the range of $\lambda\left(\bar{\theta}\right)$? From equation (1.12),

$$\lambda = \frac{1}{2} - \frac{\frac{1}{F(\bar{\theta}_t)}\int_0^{\bar{\theta}_t} \theta dF(\theta)}{2\bar{\theta}}$$

If $\bar{\theta} \to \infty$, $\lambda \to 1/2$ (if $\bar{\theta} \to \infty$, the R.H.S. must go to zero by the finite expectation of $F(\cdot)$). If $\bar{\theta} \to 0$, $\lambda \to 1/4$. This result follows from the fact that as $\bar{\theta}$ goes to zero, the ratio of the expectation of the distribution to its maximum must converge to $1/2$.

In sum, $\lambda\left(\bar{\theta}\right)$ is a monotonically increasing, continuous function of $\bar{\theta}$ on the domain $\overline{\mathbb{R}}_+$, with range $[1/4, 1/2]$. But this implies the inverse function, $\kappa(\lambda)$, is monotonically increasing continuous function on the domain $[1/4, 1/2]$ with range $\overline{\mathbb{R}}_+$. $\square$

I have established that the exogamy threshold and rate are increasing in the level of inbreeding aversion. In subsection (1.4.4) I will endogenize the inbreeding aversion coefficient, λ. The following corollary will be useful:

Corollary 1.2. *By choosing* $\lambda \in [1/4, 1/2]$*, nature can induce any cutoff* $\bar{\theta} \in \overline{\mathbb{R}}_+$.

Next we establish a system of two equations for λ and δ. In subsection (1.4.3) we take λ to be exogenous and solve for the corresponding level of fitness depression. In subsection (1.4.4) we solve for the level of inbreeding aversion that is evolutionarily stable against invaders, given δ. Then in Theorem 1.1 I establish existence and uniqueness of an equilibrium. In what follows, $\bar{\theta}$ is used as shorthand for $\kappa(\lambda)$. The long-run mean of the distribution of δ_{it} arising in a population playing threshold $\bar{\theta}$ (or with endogamy aversion λ), $\lim_{t\to\infty}\sum_{j=0}^{D} \epsilon j h(j\epsilon|\lambda)$ is often written simply as $\delta\left(\bar{\theta}\right)$, (or $\delta(\lambda)$).

1.4.3 Endogenizing Fitness Depression

In this subsection I define a long-run equilibrium with respect to fitness depression, and show that the equilibrium level of fitness depression is decreasing in the exogamy threshold $\bar{\theta}$.

Definition 1.2. A population is in a long run equilibrium with respect to fitness depression if the mean level of fitness depression is constant over time.

Consider a population that is all playing some exogenous threshold strategy, with $\lambda \in [0, 1]$. There are three possible cases, according to Proposition 1.3. The edge cases are straightforward. First, if $\lambda \leq \frac{1}{4}$, then $\bar{\theta} = 0$ by Proposition (1.3), all marriages are endogamous, over time δ goes to zero and the population becomes extinct by Proposition 1.2. Second, if $\lambda \geq \frac{1}{2}$, by Proposition 1.3 $\bar{\theta} = \infty$ and all unions are exogamous and δ goes to 1 over time. I will show in subsection 1.4.4 that this case is not proof against invaders with a different λ_i.

Finally, and for the rest of this subsection, consider the case where $\lambda \in \left(\frac{1}{4}, \frac{1}{2}\right)$. In this case, there must be an interior equilibrium, where the mass of daughter groups resulting from exogamous marriages is equal to the mass resulting from endogamous marriages, after accounting for the effect of fitness depression on the fecundity of the endogamously marrying groups.

In a long run equilibrium, the average value of $\delta\left(\bar{\theta}\left(\lambda\right)\right)$ does not change from one period to the next. By the definition of the law of motion for fitness depression in equation (1.5), an offspring group has fitness depression that is ϵ higher than its parental group if the parent group married endogamously and ϵ lower if the parental group married exogamously. Therefore, in equilibrium, the measure of groups whose fitness depression increases by ϵ must exactly be cancelled out by an equivalent measure of groups whose fitness depression decreases by ϵ, provided that the measures of the groups account for the creation of daughter groups and fitness depression. This equality allows the average level of fitness depression in the population, $\delta\left(\bar{\theta}\left(\lambda\right)\right)$, to be pinned down as a function of the degree of exogamy.

In an interior long-run equilibrium, the mass of children born from exogamous unions must be equal to that born from endogamous unions. For convenience, normalize the population entering the period to have size one. From the conditional production function, equation (1.1), each exogamous group on expectation will have $\pi_{it} = \theta_{it} + \theta_{jt}$ offspring groups. Integrating over the segment of the population that marries exogamously (and removing the groups that have $\delta_{it} = 0$; according to (1.5) groups cannot become less inbred if they already have zero fitness depression) gives the mass of groups becoming less inbred:

$$\left(1 - P\left[\delta_{it} = 0|\bar{\theta}\right]\right) \int_{0}^{\bar{\theta}} \int_{0}^{\bar{\theta}} \left(\theta_{it} + \theta_{jt}\right) dF\left(\theta_{it}\right) dF\left(\theta_{it}\right) \tag{1.14}$$

While it is possible to solve for the long-run value of $P\left[\delta_{it} = 0|\bar{\theta}\right]$ as a function of $E\left[\theta\right]$, $F\left(\bar{\theta}\right) \int_0^{\bar{\theta}} \theta dF\left(\theta\right)$, and ϵ, it is not very tractable. Instead, I take the limit of the above expression as $\epsilon \to 0$ in the long-run steady state. I show in Appendix A.1 that for all λ such that $F\left(\bar{\theta}\left(\lambda\right)\right) \in [0, 1)$, $\lim_{\epsilon \to 0} \lim_{t \to \infty} P\left[\delta_{it} = 0|\bar{\theta}\right] = 0 \; \forall \bar{\theta} < \infty$. Expression (1.14) simplifies to:

$$2F\left(\bar{\theta}\right)\int\limits_{0}^{\bar{\theta}}\theta dF\left(\theta\right) \tag{1.15}$$

after plugging in 0 for $P\left[\delta_{it}=0|\bar{\theta}\right]$. After suppressing the time subscript (due to the fact that the population is in a long-run steady state where $t\rightarrow\infty$), and writing $\delta\left(\bar{\theta}\left(\lambda\right)\right)$ with $\delta\left(\bar{\theta}\right)$ to simplify the notation, the mass of groups becoming more inbred will be:

$$\left(\left(1-F\left(\bar{\theta}\right)\right)\int\limits_{0}^{\bar{\theta}}2\theta_i dF\left(\theta_i\right)+\int\limits_{\bar{\theta}}^{\infty}2\theta_i dF\left(\theta_i\right)\right)\left(1-\delta\left(\bar{\theta}\right)\right) \tag{1.16}$$

The above expression is the sum of the mass of groups who tried to marry exogamously but were rejected, plus the mass of groups who drew a θ_i above the threshold $\bar{\theta}$ and therefore chose endogamy. This mass is deflated by the fitness depression factor, because in endogamous unions this percentage of the population is removed due to fitness depression and therefore does not contribute to the mass that increases its fitness depression factor. Any group whose $\delta_{it}=1$ will die with certainty if they marry endogamously again, so there is no need to remove these groups. Expression (1.16) can be rewritten as:

$$2\left(1-\delta\left(\bar{\theta}\right)\right)\left(E\left[\theta\right]-F\left(\bar{\theta}\right)\int\limits_{0}^{\bar{\theta}}\theta dF\left(\theta\right)\right) \tag{1.17}$$

Setting the mass from endogamous union equal to the mass from exogamous unions yields:

$$\delta\left(\bar{\theta}\right)=\frac{E\left[\theta\right]-2F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF\left(\theta\right)}{E\left[\theta\right]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF\left(\theta\right)} \tag{1.18}$$

One important implication of this equation is that although individual groups are restricted to a discrete set of fitness depression levels, the long-run population average $\delta\left(\bar{\theta}\right)$ can be any element of $[0,1]$ (this fact also holds for the population average δ_t in any period t).

Lemma 1.4. *The average level of fitness depression is decreasing in the exogamy rate and in endogamy aversion.*

Proof. From (1.18), $\delta\left(\bar{\theta}\right)=\frac{E[\theta]-2F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)}{E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)}$, so

$$\frac{d\delta\left(\bar{\theta}\right)}{d\bar{\theta}}=\frac{-E\left(\theta\right)}{\left[E\left(\theta\right)-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF\left(\theta\right)\right]^2}\times\frac{d\left[F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF\left(\theta\right)\right]}{d\bar{\theta}}$$

$-E\left(\theta\right)<0$, $\left[E\left(\theta\right)-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF\left(\theta\right)\right]^2>0$, and $\frac{d\left[F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\right]}{d\bar{\theta}}>0$, and so $\frac{d\delta\left(\bar{\theta}\right)}{d\bar{\theta}}<0$.

From subsection (1.4.3), those below the threshold choose exogamy, so a higher threshold means a higher fraction of the population is both below the threshold and is also paired with another group below the threshold; a higher threshold therefore results in a higher exogamy rate. There is a positive, monotonic function between $\bar{\theta}$ and λ, so δ is decreasing in λ as well. □

Thus, the higher the threshold for exogamy (and therefore the higher the fraction of the population that is exogamous, or the higher the endogamy aversion), the lower the level of fitness depression in the population, as one would expect.

1.4.4 Endogenizing Inbreeding Aversion:

Next, I turn to in endogenizing inbreeding aversion, essentially creating a second equation which, together with equation (1.18), describe a general equilibrium between the two variables δ and λ. Given $\delta\left(\bar{\theta}\left(\lambda\right)\right)$ as characterized by equation (1.18), what level of λ would emerge via evolution by natural selection? To characterize this level, I need to define an equilibrium concept.

Consider a population playing a particular threshold, $\bar{\theta}=\kappa\left(\lambda\right)$. Recall that the population has a corresponding long-run distribution of fitness depression parameters, $\lim_{t\to\infty}H_t\left(\cdot\right)$, which we denote $H^{\lambda}\left(\cdot\right)$ (the existence and uniqueness of this long-run distribution is proved in Appendix A.1).

Definition 1.3. The population is in an evolutionary stable equilibrium, or proof against invaders, if it is not possible for any group to play some $\lambda_{it}\neq\lambda$, and have a strictly higher growth rate.[21]

---

[21] It would probably be more realistic to model the invading λ_i as a hereditary trait, but this would introduce significant complication, both algebraic and substantive. My intuition is that if inbreeding aversion was hereditary, the λ_i that maximized the long run growth of group i would be higher than the one resulting from the current assumption. The group would internalize the negative externality that endogamy has on

To operationalize this definition I introduce a mutation that causes group i to have a different preference for endogamy aversion λ_{it}, and therefore choose a different threshold for one period. Recall that if nature could choose a homogenous λ for the entire population, she would chose $\lambda \geq 1/2$ and total exogamy. In contrast, we would intuitively expect nature to choose a lower λ_{it}, and therefore less exogamy, for the invader. The reason for this difference is that while a lower λ_{it} and therefore less exogamy, will result in greater fitness depression in future periods, from the perspective of the invader there is a chance that this fitness depression will be borne by a group that no longer has the invading λ_{it}. Effectively, the invading mutation can free-ride off the incumbent preference.

This depends on the very reasonably requirement that endogamy aversion preferences are not perfectly persistent at the group level; the preferences are genetic, after all, and the entire purpose of sex is to allow for the exchange of genetic material between individuals (and in this case, groups). In principle, one could operationalize this by modelling a mutation that survives with probability 0.5 if the invader marries exogamously, and 1 otherwise.

Instead, I make the starker assumption that mutations are completely transitory in nature, i.e., $E\left[\lambda_{it+1} | \lambda_{it} \neq \lambda\right] = \lambda$. To motivate this assumption, consider a mutation that manifests only under specific conditions that occur with probability q, and take the limit as $q \to 0$. This assumption captures the key feature of the free-riding mechanism. Conveniently, it ensures that the distribution of fitness depression of the invader remains the same as the incumbent.

Denote the invader threshold by $\bar{\theta}_i = \kappa_i\left(\lambda_i, \lambda\right)$. I assume that the invaders are a measure zero share of the total population. Therefore, one invader encountering another on the marriage market is a zero probability event.

Subsection (1.4.4) considers the principal-agent problem regarding group i's preferences and decision. The population level λ, corresponding $\bar{\theta}\left(\lambda\right)$ and mean inbreeding aversion level $\delta\left(\lambda\right)$ (as specified in (1.18)) are exogenous to this problem. I define nature's objective function for the invading group to be g_i. g_i is i's long-run growth rate when facing a population with inbreeding aversion level $\lambda \in [0, 1]$ and equilibrium fitness depression distribution $H^{\lambda}\left(\cdot\right)$ according to equations (1.2) and (1.18). Neither the principal nor the agent observes δ_i, so nature can not condition the invader's preferences on the invader's fitness depression coefficient. However, because the population is in a long-run equilibrium with respect to

---

proceeding generations, thereby lowering the endogamy rate and fitness depression, and resulting in a higher λ^* and lower δ^*. This assumption complicates the model though because an individual from other groups can marry into group i and λ_j would presumably have an effect on the inbreeding aversion level of generation $i, t+1$.

fitness depression, nature has learned the equilibrium distribution of δ_{it}, which is the same as the rest of the population, $H^{\lambda}(\cdot)$, and therefore can condition on this distribution.

First, I consider the fully exogamous case. By equation 1.3 $\lambda \geq \frac{1}{2}$, $\bar{\theta} = \infty$, and by subsection 1.4.3, $\delta = 0$. As all groups are choosing exogamy, the expected payoff for choosing exogamy is $\theta_i + E[\theta]$. This population is not proof against an invader with $\lambda_i = 0$ and $\bar{\theta} = E(\theta)$ because there are information benefits and no fitness penalty for choosing endogamy. If group i drew $\theta_i > E(\theta)$, they would have a strictly higher growth rate by choosing endogamy. By equation (1.2) if they choose endogamy they will produce on expectation $2\theta_i$ daughter groups, which is strictly larger on expectation than the $\theta_i + E[\theta]$ daughter groups they produce if they choose exogamy.

Next, in subsection (1.4.4), I consider the interior case described by $\lambda \in \left(\frac{1}{4}, \frac{1}{2}\right)$ and $\delta(\lambda)$ as in equation (1.18). I show that nature can induce group i to maximize this growth rate by choosing a specific λ_i. I show that group $i's$ utility function U_i from equation (1.3) is the same as nature's objective function when $\lambda_i = \delta(\lambda)$, where $\delta(\lambda)$ is the long-run average level of fitness depression in the population.

The population is only proof against invaders when all individuals have these preferences, i.e., when $\lambda_i = \lambda = \delta(\lambda)$. Otherwise, some mutation could result in a group having a higher expected growth rate. Therefore, in equilibrium $\lambda = \delta$.

For the population to be in a long-run equilibrium it must be true that $\lambda \in (1/4, 1/2)$ and therefore that δ is also in this range. Thus, there is a significant level of inbreeding aversion and fitness depression in the population.

Nature's objective function for Group i

Integrating the group's production function (equation (1.2)) over the type distribution and the possible levels of fitness depression gives the period t expected growth rate of group i:

$$g(\lambda_i, \lambda, \delta) = \sum_{k=0}^{N} \int_0^{\infty} \int_0^{\infty} m_{it}(\theta_{it} + \theta_{jt}) + 2(1 - m_{it})(1 - \delta_{it})\theta_{it} dF(\theta_{it}) dF(\theta_{jt}) h(\delta_{it} = k\epsilon \mid \lambda) \tag{1.19}$$

This growth rate integrates over the realizations of the quality draws of both groups, θ_{it} and θ_{jt}. Because the level of fitness depression takes discrete values, $\delta_{it} \in \{0, \epsilon\}$, the outer term sums (instead of integrates) over the different possible levels of fitness depression.

Next we manipulate and simplify this expression in order to show that for a certain value of λ_i, the growth function is equivalent to the utility function, and thus that nature can induce the group to play the growth-maximizing strategy by assigning them this level of aversion to endogamy.

Nature cannot assign preferences that are contingent on groups realizations of fitness depression, δ_{it} to either incumbents or invaders. However, Nature is assumed to understand that her choice of λ (but not λ_i) will affect the long-run distribution of fitness depression coefficients, $H^{\lambda}(\cdot)$. The invader's δ will be drawn from the same distribution as the rest of the population. Because θ_{it} and δ_{it} are independent, we can sum equation (1.19) over the long-run distribution of δ to get:

$$g(\lambda_i, \lambda) = \int_0^{\infty}\int_0^{\infty} m_{it}(\theta_{it} + \theta_{jt}) + 2(1 - m_{it})(1 - \delta(\lambda))\theta_{it} dF(\theta_{it}) dF(\theta_{jt})$$

Substituting in from equation (1.1) and integrating over the quality ranges that would result in exogamy and endogamy:

$$g(\lambda_i, \lambda) = \int_0^{\kappa(\lambda)}\int_0^{\kappa(\lambda_i,\lambda)} (\theta_{it} + \theta_{jt}) dF(\theta_{it}) dF(\theta_{jt}) \tag{1.20}$$

$$+2(1 - \delta(\lambda))\left(\int_0^{\kappa(\lambda_i,\lambda)}\int_{\kappa(\lambda)}^{\infty} \theta_{it} dF(\theta_{jt}) dF(\theta_{it}) + \int_{\kappa(\lambda_i,\lambda)}^{\infty} \theta_{it} dF(\theta_{it})\right)$$

The first double integral is when both groups are below the threshold, the second is when j is above and i is below. The last integral is when i is above (and j is below or above).

Specifying the Optimal Invader's preferences

Now that I have specified the evolutionary objective function, I investigate how nature can specify group preferences in order to maximize its growth rate. Effectively, we are looking for the λ_i that will cause the invader to choose a growth maximizing threshold, $\bar{\theta}_i$, in response to a population endowed with λ and therefore playing $\kappa(\lambda)$ according to Proposition 1.3.

Proposition 1.4. *Nature's best response to a population playing a cutoff* $\bar{\theta}$ *is to set* $\lambda_i = \delta(\lambda)$*, where* $\delta(\lambda) = \frac{E[\theta] - 2F(\kappa(\lambda))\int_0^{\kappa(\lambda)} \theta dF(\theta)}{E[\theta] - F(\kappa(\lambda))\int_0^{\kappa(\lambda)} \theta dF(\theta)}$ *in accordance with equation* (1.18).

Proof. The group's problem will be:

$$\max_{\bar{\theta}_{it}} U_{it}\left(\bar{\theta}_t, \bar{\theta}_{it}\right)$$

Substituting in from the payoffs defined in equation (1.4), $U_{it} = m_{it}(\theta_{it} + \theta_{jt}) + 2(1 - m_{it})(1 - \lambda)\theta_{it}$, and integrating over the appropriate ranges of quality:

$$\max_{\bar{\theta}_{it}} \int_0^{\bar{\theta}_{it}} \int_0^{\bar{\theta}_t} (\theta_{it} + \theta_{jt})\, dF(\theta_{jt})\, dF(\theta_{it}) + 2(1 - \lambda)\left[\int_{\bar{\theta}_t}^{\infty} \int_0^{\bar{\theta}_{it}} \theta_{it} dF(\theta_{it})\, dF(\theta_{jt}) + \int_{\bar{\theta}_{it}}^{\infty} \theta_{it} dF(\theta_{it})\right]$$

The first expression is the probability weighted payoff if both candidates/groups are below the cutoff, and thus marry exogamously. The second term is the payoff if i marries endogamously, integrated over the ranges of quality that would result in this outcome (either below the cutoff and rejected by group j, or above the cutoff).

But this problem is very similar to nature's problem in (1.20); the only difference is that λ replaces $\delta\left(\bar{\theta}\right)$.

Suppose that $\lambda \in (1/4, 1/2)$. If nature chooses $\lambda_i = \delta\left(\bar{\theta}\right)$, the two problems are identical. Therefore, nature can induce individuals to choose the growth maximizing $\bar{\theta}_i$ by assigning the preference $\lambda_i = \delta\left(\bar{\theta}\right)$ (the uniqueness of this assignment will be discussed in Theorem 1.1 below).

Therefore, the λ_i that maximizes the growth rate must result in a threshold of $\kappa(\delta(\lambda), \lambda)$. The problem is symmetric for all groups and so $\lambda_i = \lambda$ for all i.

For $\lambda \leq \frac{1}{4}$, the population dies within a finite number of periods and so cannot be said to be in a long-run equilibrium. For $\lambda > 1/2$, a similar argument shows that the solutions to the two problems coincide. However, if $\lambda > 1/2$, the population level of fitness depression will reach zero in a finite number of periods. After this point, any group with a draw of $\theta_i > E[\theta]$ can do better by deviating to endogamy, so no λ in this range can be robust to invaders. □

Essentially, Proposition 1.4 states that agents will discount endogamous marriages at the expected rate they are destructive of productivity, which is intuitively appealing.

General Equilibrium

Theorem 1.1. *Under assumptions 1, 2, 3, and 4 there exists a unique* $\lambda^* \in (1/4, 1/2)$ *that is evolutionarily stable. By Corollary 1.2, this implies the existence of a unique evolutionarily stable* $\bar{\theta}^* \in (0, \infty)$*, and by Proposition 1.4, a unique* $\delta^* \in \left(\frac{1}{4}, \frac{1}{2}\right)$.

Proof. Collect results from equation (1.12), equation (1.18) and Proposition 1.4 above. This gives us three equations (in order):

$$
\begin{gathered}
\left(1-2\lambda_i\right)\bar{\theta}_i = \frac{1}{F\left(\bar{\theta}\right)}\int_0^{\bar{\theta}} \theta dF\left(\theta\right) \\
\delta\left(\bar{\theta}\right) = \frac{E[\theta]-2F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)}{E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)} \\
\lambda_i = \delta\left(\bar{\theta}\right)
\end{gathered}
$$

If the equilibrium is robust to invaders, it must be the case that $\lambda_i = \lambda$ (and therefore $\bar{\theta}_i = \bar{\theta}$). That is, the best an invader can do is to play the same strategy as the other groups in the population. Consequently, I can collapse the system of three equations into two:

$$
\begin{gathered}
R_1\left(\bar{\theta}\right) = \frac{1}{2}\left(1-\frac{1}{\bar{\theta}F\left(\bar{\theta}\right)}\int_0^{\bar{\theta}}\theta dF\left(\theta\right)\right) \\
R_2\left(\bar{\theta}\right) = \frac{E[\theta]-2F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)}{E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)}
\end{gathered}
$$

To show existence and uniqueness, I will show that the first expression is increasing in $\bar{\theta}$ and the second is decreasing. These properties follow because the first expression starts below the second and after some point is above it, so they must cross exactly once. To elaborate, consider:

$$
\frac{dR_1\left(\bar{\theta}\right)}{d\bar{\theta}} = \frac{d\left(-\frac{1}{2\bar{\theta}F\left(\bar{\theta}\right)}\int_0^{\bar{\theta}}\theta dF\left(\theta\right)\right)}{d\bar{\theta}} = -\frac{1}{2}\times\frac{d\left(\frac{1}{\bar{\theta}F\left(\bar{\theta}\right)}\int_0^{\bar{\theta}}\theta dF\left(\theta\right)\right)}{d\bar{\theta}}
$$

From Assumption 3 $\frac{d\left(\frac{1}{\bar{\theta}F\left(\bar{\theta}\right)}\int_0^{\bar{\theta}}\theta dF(\theta)\right)}{d\bar{\theta}} < 0$, so $\frac{dR_1\left(\bar{\theta}\right)}{d\bar{\theta}} > 0$

From Lemma (1.4), $\frac{dR_2\left(\bar{\theta}\right)}{d\bar{\theta}} < 0$.

Consider $R_1\left(\bar{\theta}\right)$. Using l'Hopital's Rule twice, assuming that $f(0) > 0$; and that $f'(0)$ exists, it follows that $R_1(0) = \frac{1}{4}$. $R_2(0) = 1 > R_1(0)$. R_2 is decreasing in $\bar{\theta}$ whereas R_1 is increasing in $\bar{\theta}$. Therefore, if they cross they must cross only once.

Since $R_1(0) = 1/4$ and is an increasing function, it must be strictly positive everywhere. If $F\left(\bar{\theta}\right) = \sqrt{\frac{1}{2}}$ then $R_2\left(\bar{\theta}\right) = 0$, and thus at $\bar{\theta} : F\left(\bar{\theta}\right) = \sqrt{\frac{1}{2}}$, $R_1 > R_2$. The functions are both continuous. Therefore, they must cross exactly once.

Denote this value of $\bar{\theta}$ as $\bar{\theta}^*$. The corresponding value of $R_1\left(\bar{\theta}^*\right) = R_2\left(\bar{\theta}^*\right) \equiv \lambda^*$. Because $R_1(0) = \frac{1}{4}$ and R_1 is an increasing function, $\lambda^* > \frac{1}{4}$. From Proposition 1.4 $\delta^* = \lambda^*$.

In sum, the model predicts not only positive levels of endogamy, fitness depression, and lower growth than is socially optimal, but also that the degree of fitness depression and decreased growth would have been quantitatively different from zero. The primary driver of this result is adverse selection in the marriage market. The presence of adverse selection may in itself be unsurprising. What is more revealing is that this adverse selection is quite difficult to counteract. Even endowing nature with the ability to modify group preferences so they dislike endogamy cannot overcome the problem, as free-riding at the level of the gene results in continued fitness depression and lower population growth. □

1.5 Conclusion

For the vast majority of time that Homo sapiens have been on the planet, low population density meant that there would have been relatively few opportunities for observing the quality of potential exogamous marriage partners. Adverse selection - resulting from the unwillingness of high types to accept a partner of unknown quality - likely afflicted the exogamous marriage market, causing high types to choose endogamy. The resulting fitness depression would have been significant and potentially result in the complete collapse of some populations. An innate preference for aversion to endogamy would likely have evolved. This would have enabled significant amounts of exogamy, reducing the population wide level of fitness depression.

An interesting avenue for further research is how humans overcame this problem. Various additional preferences, strategies and physical traits would likely have been helpful in reducing fitness depression, especially those that provided opportunities to observe the quality of potential mates. Social institutions for congregation would likely have been beneficial, especially if these included costly signaling activities such as dancing, athletic competition and displays of wealth. Other types of traits that probably would have been helpful include physical traits that signal health, such as long head hair, beards and excess adipose tissue.

Chapter 2

Cooperation, Competition, and Linguistic Diversity

2.1 Introduction

Over 7,000 human languages are currently spoken on the planet. (Lewis et al., 2013). Prior to colonialization, approximately 850 of these were spoken in the area that is now Papua New Guinea.[1] Nearby Australia, on the other hand, was home to only about 392 languages, despite being over 16 times as large.[2]

Why are there so many languages in some places and so few in others? Is it possible that economic incentives play a role in effecting this diversity? In this paper we show that repeated conflict and cooperative interactions can plausibly generate linguistic change and explain spatial and temporal variations in the distribution of languages.

From the empirical literature on linguistic change and linguistic diversity, we extract three stylized facts. First, linguistic diversity varies greatly across the earth's surface, correlated with several geographic factors. Second, a pair of languages can merge, diverge, or maintain a stable relationship. Sometimes a language goes extinct, or is revived from near-extinction. Third, there is large variation in the rates that languages diverge, and they tend to explode apart very quickly when they first separate. We will describe and discuss these patterns in more detail in section 2.5.

We explain the phenomena described above by explicitly modeling a process of linguistic development in small-scale, sedentary societies. We develop a formal economic model to explain how strategic incentives can induce language change and thereby explain linguistic diversity. This model is dynamic, micro-founded, incentive based and it produces a rich set of predictions.

We propose that conflict may be an important causal variable for linguistic change, particularly the divergence of languages over time and their diversity over space. Conflict here is defined as any competitive interaction between two groups, each containing multiple members. In addition to conflict as a diverging influence on languages, our model also incorporates the traditional, more intuitive idea that cooperation fosters linguistic homogeneity. As a result, our theory generalizes to a broad set of situations. This paper complements the existing literature on trade and conflict by suggesting how language may affect these interactions, and how it is in turn influenced by them.

We begin by choosing modeling assumptions appropriate to the environment and strategic interactions that small-scale, sedentary societies faced.

---

[1] Based on Ethnologue, Papua New Guinea. 851 languages minus one that is not indigenous

[2] Based on Ethnologue, Australia. 403 languages minus 11 that are non indigenous

We develop a formal model of language and linguistic change, including a set of elements that forms a vocabulary, definitions for 'dialect' vs 'language' and a mechanism for change. We define a game that is played each period.

We show how the relationship between a pair of languages can affect basic economic interactions between two groups. Common language facilitates co-operation, whereas language unique to one group gives that group an advantage in competitive interactions. We argue that the relationship between a pair of languages, and linguistic diversity itself, is important in that it impacts the outcomes of these common strategic games.

Next, we show how these interactions create incentives for groups to modify the language they use, and characterize those changes for a given period. Repeated changes over multiple periods results in convergence or divergence for a pair of dialects.

Finally, we show that if the underlying probability of conflict vs cooperative games remains constant, these trajectories result in a stable pattern of linguistic diversity in a region. This pattern should be of interest to economists because it feeds back into the probable outcomes of cooperative and competitive games between groups.

A high frequency of cooperative games, such as trade, has a homogenizing effect on language, thus leading to low linguistic diversity. High frequency of competition such as warfare, on the other hand, raises the value of private information. This leads to the generation of new linguistic expressions and, over time, high linguistic diversity.

2.1.1 Preview of the Model

In our model, groups interact periodically in pairs, in one of three types of activities: cooperation, competition, and non-interaction (hereafter referred to as a "null game"). Each group speaks a *dialect*, and two dialects are considered the same *language* if they are sufficiently similar.[3] In section B.1, we show that the groups' expected payoffs in a game are a function of the relationship between their dialect and their opponent's dialect. In cooperation, modeled as a coordination game, group i's payoff is increasing in the commonality between i and j's dialects. This captures the intuition that coordination is more likely to be successful when the two groups can communicate in an accurate and inexpensive manner.

During conflict, modeled as a zero-sum game, i's expected payoff increases in how much it understands the opponent j's dialect, while decreasing in the level of j's comprehension of i's dialect. Intuitively, when i and j are in an antagonistic relationship, the more information i has about j, the more advantaged i would be in that relationship. Since language is the

---

[3] The definitions of *dialect* and *language* will be made more precise in section 2.3.

most common medium of intra-group communication, a better understanding of j's dialect leads to a higher chance of obtaining j's private information. By the same token, the less j understands i's dialect, the better i would be able to protect its own information.[4]

Lastly, if two groups do not interact in a period, which we model by a "null game", payoffs are constant in all outcomes.

Each group interacts with every other group exactly once per period, and the nature of each interaction (i.e. cooperative, competitive, or null) is determined by a random draw for every pair from among the three games. For a pair i, j, the probability of drawing a cooperative game is p_{ij}, the probability of drawing a competitive game is q_{ij}, and the probability of drawing a null game is $1 - p_{ij} - q_{ij}$. The null game is more likely to occur when groups are farther apart geographically. To maximize expected payoff, group i can make a costly effort to change its dialect at the beginning of a period by (i) learning parts of the dialect(s) spoken by some other group(s), and/or (ii) inventing novel linguistic expressions that are only understood by members of group i. The benefit of learning from an existing dialect is twofold: first, it boosts i's chance to succeed in a cooperative interaction with speakers of that dialect; second, it increases i's competitive edge in conflict against speakers of that dialect, because i would be more likely to decipher j's intra-group communication. Inventing new expressions for intra-group communication, on the other hand, improves i's odds of winning a conflict against all other groups, as the new expressions make it harder for all other groups to acquire information on i.

For simplicity we assume that agents are myopic, i.e. they do not consider the impact of the changes they make on future generations.[5]

The main results we derive in this paper fall into three categories: (i) optimal linguistic change in a period game; (ii) the existence and characterization of a steady state in the infinitely repeated game, and convergence towards this state; and (iii) the steady state number of languages in a region and its comparative statics.

In any given period, it will always be optimal for a group to invent some new linguistic expressions, assuming that competitive games happen with positive probability. With regard to learning existing dialects, we show that if the probability of a non-trivial interaction

---

[4] There is an additional, complementary benefit to a group of speaking a differentiated dialect, which we do not discuss at length or model explicitly. Differentiated language—including both vocabulary and accent—can help group members identify each other, which in turn facilitates the in-group distribution of sensitive information.

[5] Considering future generations would likely have a symmetric effect, enhancing both incentives to invent new expressions and to learn other groups' languages.

(i.e. cooperation or competition) with a neighbor drops sufficiently fast with an increase in geographical distance, then it is optimal for a group to learn from the dialect spoken by its closest neighbor(s) first.

In the dynamic setting, where the period game is played repeatedly, we establish the existence and uniqueness of a steady state in which the set of groups speaking each language remains unchanged over time. We emphasize that while languages themselves continue to evolve in the steady state, *the set of groups* speaking any particular language does not. This in turn generates a unique steady state number of languages. If the initial linguistic composition is of a particular symmetric structure, we show that the number of languages in the region converges to that in the steady state. Lastly, we show that the steady state number of languages is weakly decreasing in the relative probability of cooperative vs. competitive interactions.

The rest of the paper is organized as follows. Section section 2.2 describes the environmental and institutional context of the societies in question and reviews the economic literature relevant to the paper. We present our model in section 2.3, and the theoretical results in section 2.4. In subsection 2.4.1, we derive optimal language change in a typical period. subsection 2.4.2 introduces the dynamic setting, establishes the existence, uniqueness and characterization of the steady state, and convergence thereto. The comparative statics of the steady state are given in subsection 2.4.3.

section 2.5 summarizes the multidisciplinary literature on linguistic diversity and identifies stylized facts pertaining to our model. Existing explanations for linguistic diversity are summarized and discussed. section 2.6 concludes.

2.2 Context

This paper draws on several different theoretical literatures and is consistent with empirical results from a variety of fields of inquiry. In this section we outline the economic conditions of the societies we are modeling and the relevant economic literature. In section 2.5 we will discuss linguistic research from other academic disciplines that has bearing upon the question.

Modern linguistic diversity is the result of many thousands of years of change, much of which occurred in small-scale societies. More recent conditions should be weighted more heavily, and sedentism has been widespread globally for the past few thousand years. ’Small-scale’ is defined here to include societies with the highest level of political organization being the family, the village, the clan, the big man collectivity or the simple chiefdom, as described in Johnson and Earle (2000). Complex chiefdoms and states are beyond the scope of the model. Sedentary peoples live largely in one place, and in small-scale societies their descendants are likely to live there as well.

The societies we are modeling may be foragers or farmers, they may be stratified or egalitarian, but literacy is virtually nonexistent, and there are no standing armies or specialized bureaucrats or merchants. Population of a politically autonomous group is typically less than 2000 (Johnson and Earle, 2000, p.246) .

The results of our model, regarding variation in linguistic diversity over time, space and trajectory, should be interpreted under these specific circumstances. Societies organized into states are beyond the scope of this model primarily because local language is not free to evolve independently of a region wide political power. State bureaucracies use a standardized language to administer government programs and often provide incentives and pressure for local groups to adopt or at least learn this common language.

Small scale societies often engage in strategic interactions of cooperation and conflict. Trade, including long distance trade, exogamous marriage, political alliance, warfare, feasting, and raiding are features of many small scale societies. Their frequency, however, varies greatly across groups, from non-existent to common. For descriptions and explanations of the variation across societies of these institutions, see Dow et al. (2016) for exogamy, Johnson and Earle (2000) for political alliance, Dow et al. (2017) for warfare, Hayden (1996) for feasting.

Small scale societies have much less individual specialization than state-organized societies (Johnson and Earle, 2000); and so most adults would be involved in most activities when the opportunity arose. Few or no people were uniquely merchants, translators, or warriors, and so internal or external coordination would involve the participation of large segments of society, and thus affect their common language. Rapid large scale invasions and large scale political institutions are not considered, as the technology and institutions largely did not exist for such activities. The set of conditions we are analyzing does include, however, unorganized large scale movements of people that take place over many generations such as the Bantu expansion. This movement of people, languages and technology occurred in Sub-Saharan Africa from about 3000-2000 years ago. There was probably no central political organization nor large scale effort to homogenize language. We therefore would predict that dialects, and then languages, would arise over time in the newly occupied territories, in accordance with the frequency of competitive and cooperative interactions. Caution should therefore be used when applying the intuition resulting from this model to more modern circumstances.

In situations involving potential benefit from cooperation between groups, the parties' ability to understand each other is undoubtedly crucial. Trade, exogamous marriage and political alliances rely on clear communication. Perhaps less conventionally, communication is also critical in situations of conflict. Theoretical work on conflict suggests that information plays a vital role in determining the outcome of a hostile contention. In particular, there are "incentives to misrepresent information [...] specifically, each party would like to appear tougher than they really are" (Garfinkel and Skaperdas, 2007). Information regarding the strength of a group and any plans they have regarding attack or defense is particularly important. Success is more likely in these situations when a party can communicate such information within the group while keeping it indecipherable from their opponents. The invention of new terms that are understood only by the "insiders" of the group facilitates the control of information.[6] In addition, the empirical work by Arbatlı et al. (2020) records a positive effect of population (genetic) diversity on the risk and intensity of historical and contemporary civil conflicts. One of the channels of such an effect is through the increased linguistic fractionalization and polarization due to population diversity, which in turn bred mistrust among different linguistic groups, thereby contributing conflicts.

The highly creative, social and linguistic human species has evolved to use language as a tool in strategic interactions. It is easy to imagine several possible micro mechanisms for linguistic change. For example, there could be a leader that decides on the changes, or a more organic process in which everyone takes a turn changing a few things, or even an evolved instinct to learn and innovate, especially under certain conditions. We don't have strong beliefs about the precise mechanism, or the level of intention involved.

Perhaps the literature we draw most heavily on is theory of language and communication. Economic theorists have a long, albeit sporadic, interest in the subject. It is common to regard language as a facilitator of cooperation, especially in the form of trade (Lazear, 1999). Blume et al. (1993), Wärneryd (1993), and Robson (1990), among others, studied how messages in a language become associated with meaning in cooperative interactions with communication. Rubinstein (2000) characterized an "optimal" language based on criteria

---

[6] Group-specific vocabulary and other common elements of language (e.g. accents) may also serve the purpose of identifying insiders from "outsiders" of the group. Accurate identification of fellow members enables groups to communicate strategically sensitive information to its membership without accidentally revealing it to outsiders. Accents would be especially useful for the purpose of identifying members of groups that are closely related linguistically. We do not formally model differentiated language as enabling identification in this way, but this use of language is definitely consistent with our hypothesis that linguistic groups will choose to invent more new terms in high conflict settings. Language as an identifier is especially useful in groups large enough that not all members know each other personally.

such as the ability of a language to identify objects and the ease by which the language is learned.[7] More recently, Blume and Board (2013a) explored the implications of language competence, and the knowledge thereof, on the efficiency of communication in common interest games.

We incorporate versions of these concepts into our model. Specifically, there is a cost associated with learning linguistic elements, as well as a benefit in that learning elements increases the likelihood of understanding future communications. We also have an 'optimal' language, although we innovate in that groups are able to choose which languages to learn and how much to learn of each.

Anderson and Van Wincoop (2004) review the literature showing that communication barriers increase the cost of trade between countries. Blume and Board (2013b) outlines the theoretical case where the misalignment of interest between a sender and a receiver would lead the former to intentionally communicate with messages that are vague, even though more precise alternatives are available.

We bring together cooperative and competitive incentives into a single dynamic model, exploring how they generate together patterns of change and the spatial distribution of languages.

There are large empirical literatures that use geography to predict and/or explain linguistic diversity, linguistic diversity to predict/explain conflict and cooperation, and geography to predict/explain conflict and cooperation. We link these variables together in a logical explanatory pathway from geography to conflict/cooperation to linguistic diversity. Geography is important as it is the most exogenous and has the best available data, resulting in many studies of the relationship between geographic variables and linguistic diversity. Levels of conflict and cooperation are almost certainly affected by technology, culture, cultural differences and linguistic differences, but we mostly leave the exploration of these relationships to other researchers. We focus on the effect of conflict on linguistic diversity as we think this line of thinking is both novel and compelling.

Almost by definition, the societies we are interested in did not leave written records of their languages. Anthropologists and archaeologists have studied many small-scale societies in detail, but this data tends to be focused on describing each society at a single point in time. Panel data is nearly impossible to come by. As a result we rely mostly on qualitative evidence

---

[7] See Lipman (2003) for a good review.

and the results of other researchers on related questions to establish empirical support for our theory.[8]

Having said this, there is a large literature in economic development establishing a strong positive correlation between ethnolinguistic diversity and conflict (Esteban et al., 2012; Desmet et al., 2012). Some authors, however, do not find this relationship (Fearon and Laitin, 2003). Most of the data used is from the modern era during which societies were organized at the state or colonial level, and with most a single cross-section of languages' vocabularies. It is therefore not well suited to studying our question.

Measures of polarization, which rely on measures of ethnolinguistic distance between groups, are found to be significant correlates of conflict (Esteban et al., 2012). Measures of distance between languages are often used as proxies for differences in ethnicity or differences in preferences over public goods. The authors use linguistic data for the explanatory variable of interest because it is more available than data on genetic distance or on differences in preferences, and more exogenous than the latter.

If conflict causes significant linguistic change, however, there is an endogeneity problem in any attempt to measure the causal impact of ethnolinguistic fractionalization or polarization on conflict. This would be particularly problematic in approaches where linguistic similarity alone is used to estimate distance between a pair of groups or time since a linguistic split, or where the number of groups used to measure fractionalization is based solely on linguistic difference.

We do not deny that linguistic distance or its correlates may lead to higher levels of conflict, but we do believe that the reverse is also likely to be true. It would be unfortunate to over-attribute conflict to ethnic division and overlook other possible causes.

We argue that changes in linguistic diversity are, at least partially, influenced by strategic incentives. Over many generations, strategically induced changes may accumulate and generate new languages. We propose a causal channel for such linguistic change: a region's geographic make-up (e.g. climate, terrain, soil, vegetation, natural resources, bodies of water, etc.) and the technology of the society affects the relative probability of cooperative vs. competitive interactions. This relative probability in turn determines the strategic incentives that drive linguistic change.

---

[8] The data available on small scale societies in the cross-cultural standard sample has only a few hundred observations at most, and many of the data points are problematic for various reasons. We have not been able to find or construct any panel data that documents changes in linguistic diversity and conflict levels over time for many small scale societies. This makes it difficult to establish any relationship with the degree of certainty required to report empirical results. For this reason, this paper is mainly theoretical. We have done our best to document and discuss empirical patterns whenever possible

It is therefore helpful to review the literature that studies the impact of geography on linguistic diversity. In this section we focus on studies by economists; for a more detailed review of studies of the relationship between language and geography, see section 2.5.

Michalopoulos (2012) finds that variance in land quality and in altitude are positively and significantly correlated with linguistic diversity. He theorizes that geographically heterogeneous regions foster location-specific human capital that is not easily transferable to a different environment. As a result, possession of location-specific human capital contributes to the immobility of language groups. Immobility thus leads to a higher linguistic diversity within the region. Michalopoulos (2012) also notes that, as geographically uniform territories are easier to conquer, and invasion has a homogenizing effect on language and culture, so one would expect to see a positive correlation between linguistic diversity and geographic heterogeneity within a region.

Michalopoulos does not cite any evidence that location-specific human capital or any related variables are correlated with linguistic diversity.

According to a line of argument mostly mapped out in Dow and Reed (2015), Dow et al. (2016) and Dow et al. (2017), environmental productivity and climate shocks are important conditions for prehistoric warfare. High resource productivity leads to sedentism and larger communities. Larger communities results in more endogamy, which along with sedentism makes it harder for individuals to relocate from one site to another in response to disproportional productivity shocks.

It seems likely that geographic heterogeneity makes the effect of climate shocks more likely to be disproportional across sites. The site that benefited most (or was harmed least) by the shock becomes an attractive attack target for a group living at a less favoured site. Thus, we posit the following chain of causation: high regional productivity and high geographic variation causes increased conflict, which over time causes high linguistic diversity.

Arbatlı et al. (2020) identify population (genetic) diversity as a pivotal factor that contributes to both contemporary and historical conflicts within and across groups. The authors provide empirical support for a number of channels through which population diversity may lead to conflicts. Most relevant for this paper are the channels involving (ethno)linguistic diversity. Two measures of ethnolinguistic diversity are used: ethnolinguistic fractionalization and ethnolinguistic polarization, both of which capture essentially the proportional representation of ethnic groups within a society. The authors find that higher ethnolinguistic diversity not only has a direct, positive association with more conflicts, but also contributes indirectly to conflicts through the erosion interpersonal trust, which in turn breeds animosity among members of a society.

Last but not least, Chen (2017) reports findings from a lab experiment in which linguistic diversification is observed. The design of the experiment follows closely of the model in this paper. In the experiment, pairs of subjects endowed with a set of words, or vocabulary, interact repeatedly in a series of underlying games, in which they use the words to signal their intended action. The underlying games are either coordination or zero-sum, and the subjects are allowed to modify their vocabularies by learning words from their counterpart and creating new words. The author observes that, over time, the pairs of vocabularies in coordination games tend to converge, while in zero-sum games, the vocabularies experience constant pressure to diverge. These findings provide strong support for the theoretical predictions made in this paper.

2.3 Model

Groups residing in a region interact myopically in every period. There are three types of interactions—cooperative, competitive, and null—which we use to model cooperation, competition, and no interactions, respectively. The probability of each type of interaction occurring is determined by the geographic environment, which is captured by a fixed, exogenous, region-wide parameter, and a measure of distance between the interacting groups. The payoffs of the interactions depend on the groups' understanding of each other's dialects. Therefore, at the beginning of a period, the groups, anticipating the type of interactions they may be involved in, choose to update their existing dialect by either learning words from other groups or inventing new linguistic expressions. These linguistic changes, as well as the region's linguistic composition, in turn determine the groups' payoffs in the periodic interactions.[9] The geographic parameters are constant over time, while the linguistic variables, to be introduced in subsection 2.3.2, are in general time varying. We use a superscript t on the linguistic variables to index time. However, in this and the next section, where it is clear from the context that the variable refers to one in a particular period t, we may suppress this time index.

2.3.1 Geography

In this subsection, we impose a geographic structure on a region. This structure will allow us to discretely count the dialects and languages and to determine whether two groups speak

---

[9] A region's linguistic composition is a description of which group speaks which language, and how much each group understands of the other groups' dialects. The term is formally defined in 2.5.

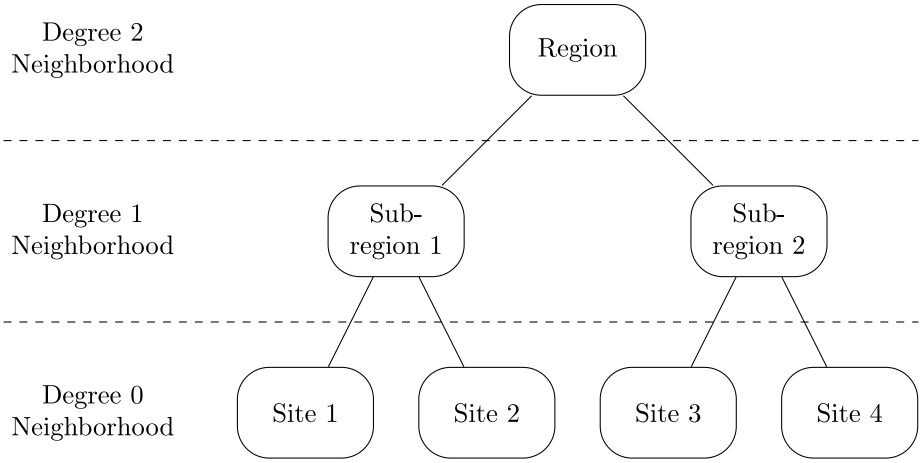


Figure 2.1: Neighborhood Illustration

the same language. A region of size normalized to 1 is populated by groups from the set $\mathscr{G} = \{1, \ldots, G\}$, where $G = 2^n$, $n \in \mathbb{N} \cup \{0\}$. Each group $i \in \mathscr{G}$ controls a site in the region. To model distances between sites, we introduce a neighborhood structure analogous to a binary tree, as depicted in Figure 2.1. A degree 1 neighborhood is inhabited by two groups; they are each other's closest neighbors. A degree 2 neighborhood consists of two degree 1 neighborhoods, or four groups; and a degree k neighborhood consists of two degree $k-1$ neighborhoods. Hence, there are 2^k groups living in a degree k neighborhood.

The degree of neighborhood between two groups can be loosely interpreted as a measure of the travel time between them, including both distance and geographic barriers. For example, we can think of Sub-region 1 in Figure 2.1 as including a common meeting place,and the degree 1 as a measure of the difficulty with which groups 1 and 2 (living in Sites 1 and 2) reach the meeting place. In this example the Region (or degree 2 neighbourhood) includes places for groups 1 or 2 to meet groups 3 or 4. Due to geographic barriers/distance this degree 2 meeting place is assumed to be less accessible than a degree 1 meeting place.

For any two groups i and j, let $d_{ij} \in \{0, 1, \ldots, D\}$ denote the smallest degree neighborhood i and j share.[10] Then d_{ij} is a proxy for the traveling time for the members of group i to meet with group j. A higher d_{ij} indicates a farther distance between groups i and j.[11]

---

[10] Trivially, $d_{ij} = d_{ji}$. Since any group is a degree 0 neighbor with itself, $d_{ij} = 0 \Leftrightarrow i = j$.

[11] The binary tree neighborhood structure is a highly stylized way to model distance between groups. We choose this approach mainly for its tractability. This binary tree neighborhood structure allows us to use a single parameter, d, to keep track of both the distance between any two groups and the number of groups within a particular degree of neighborhood (2^d). The latter feature is useful in proving 2.1, and subsequently establishing the existence of a steady state.

As will be clear in subsection 2.3.3, the higher a pair's neighborhood degree, the less likely they will interact in a given period. D is the highest degree of neighborhood possible in the region (in Figure 2.1, for example, $D = 2$). Thus, $D = \log_2 G$, or equivalently, $G = 2^D$.

2.3.2 Language

For simplicity, we model a language as a list of sound-meaning pairs, which we call *linguistic elements*. Let $\mathscr{L}$ be the set of all possible linguistic elements.[12] For technical convenience, we assume that $\mathscr{L} = \mathbb{R}$.[13] Let $\mathcal{B}$ be the Borel σ-algebra on $\mathbb{R}$, and $|\cdot|$ be the Lebesgue measure on $(\mathscr{L}, \mathcal{B})$. Therefore, $(\mathscr{L}, \mathcal{B}, |\cdot|)$ forms a measure space.

Definition 2.1. A *dialect* spoken by group i at the beginning of period t is a non-empty, measurable (with respect to $|\cdot|$) subset $L_i^t \subset \mathscr{L}$ with a finite measure $|L_i^t| < \infty$. Let $L_0^t = \mathscr{L} \setminus \bigcup_{k \in \mathscr{G}} L_k^t$ denote the set of linguistic elements that are not used in any dialect.

Group i can choose to enrich its dialect in two ways: (i) adopting linguistic elements from other groups' dialects, i.e. from $\{L_j^t \setminus L_i^t : j \neq i\}$, or (ii) acquiring elements from L_0^t, which are not in use by any group.

Denote by E_i^t the set of linguistic elements i learns from other groups' dialects. We write

$$E_i^t = \bigcup_{k \in \mathscr{G} \setminus \{i\}} E_{ik}^t$$

where $E_{ik}^t \subseteq L_k^t \setminus L_i^t$ is the (possibly empty) set of linguistic elements adopted from L_k^t by i. Assume that each E_{ik}^t is measurable. A non-empty E_i^t indicates that group i is adopting/learning part of other groups' vocabulary. This captures, for example, the adoption of English words by Spanish as discussed in section 2.5.

Let $N_i^t \subseteq L_0^t$ denote the set of linguistic elements i acquires from L_0^t. Again, each N_i^t is measurable.[14] We further assume that groups acquire elements from L_0^t in an uncoordinated fashion, and therefore that N_i^t and N_j^t are disjoint (up to a subset of measure zero) for any

---

[12] That is, $\mathscr{L}$ contains all possible sound-meaning pairs. For example, "perro"-dog, "haha"-dog, "cat"-dog and "dog"-dog would be four of the elements of $\mathscr{L}$. Note also that $\mathscr{L}$ is not time dependent.

[13] We do not require a notion of closeness of any two linguistic elements, however. Using an alternative assumption—letting $\mathscr{L}$ be countable—will not change our results qualitatively. It is just awkward to work with a discrete set.

[14] In the remainder of the paper, whenever we refer to a subset of $\mathscr{L}$, we assume that it is measurable, unless otherwise noted.

i and j.[15] The latter assumption is plausible, in that L_0^t is of an infinite measure while each N_i^t has a finite measure, and so it is unlikely that any two groups would adopt the same subset of elements from L_0^t when their actions are uncoordinated. This type of learning can be thought of as people in group i inventing new expressions for their own dialect.

We consider several dialects to be variants of the same language if they are sufficiently similar, as follows:

Definition 2.2. A subset $\widetilde{\mathscr{G}}_{\lambda^t} \subseteq \mathscr{G}$ of groups speak the *same language* λ^t at the end of period t if $L_i^t \cup E_i^t = L_j^t \cup E_j^t$, up to a subset of measure zero for all $i, j \in \widetilde{\mathscr{G}}^t$. The language λ^t is defined as the set $\lambda^t = \bigcup_{k \in \widetilde{\mathscr{G}}_{\lambda^t}} (L_k^t \cup E_k^t \cup N_k^t)$.

Thus, two dialects are deemed the same language if they consist of the same set of linguistic elements acquired either through inheritance from a previous generation (i.e. elements in the set L_k^t) or learning (i.e. elements in the set E_k^t).[16]

Elements in N are linguistic innovations, as discussed in subsection 2.3.3, and often considered to be new “slang”. By counting English as a single language we declare Canadian English and American English, for example, to be the same language. These two variants of English, however, are by no means identical. What distinguishes them, differences in spelling, accent, language use, etc., exist partially due to the diverging force of competition.

---

[15] Note also that L_i^t, E_i^t and N_i^t are disjoint pairwise. The names of the sets are chosen to signify their properties: E stands for “existing”, so elements in E_i^t are chosen from existing dialect (other than L_i^t); and N stands for “non-existing”, so that it contains elements that are not from an extant dialect.

[16] Some readers may wonder why we choose $L_i^t \cup E_i^t = L_j^t \cup E_j^t$, as opposed to $L_i^t = L_j^t$ (or $L_i^t \cup E_i^t \cup N_i^t = L_j^t \cup E_j^t \cup N_j^t$), as the criterion of similarity between i and j's dialects. Our choice is based on two reasons. First, languages evolve continually, and so it is unlikely for two dialects—in our model as well as in reality—to be identical at any point in time. Therefore, $L_i^t = L_j^t$ as a criterion of similarity, which requires i and j's dialects to be identical up to a subset of measure zero, would be too stringent. Second, in our model, especially the dynamic part in subsection 2.4.2, groups learn and invent linguistic elements in every period. In this setting, $L_i^t \cup E_i^t = L_j^t \cup E_j^t$ is the most similar two dialects can get *at the end of a given period*. Hence 2.2 is already using a very strict criterion.

2.2 is also consistent with the two criteria of language classification in the *Ethnologue* (see footnote 35): common ethnolinguistic source and mutual intelligibility. Every language has many variants, and these variants contain features specific to a particular linguistic group. The degree of similarity of pairs of variants is of a continuous nature. Above some arbitrary degree of dissimilarity, variants are referred to as distinct dialects, and above a second but higher arbitrary degree of dissimilarity they are counted as distinct languages. In defining language and dialect, the *Ethnologue* makes the following comment: “Every language is characterized by variation within the speech community that uses it. Those varieties, in turn, are more or less divergent from one another. These divergent varieties are often referred to as dialects. They may be distinct enough to be considered separate languages or sufficiently similar to be considered merely characteristic of a particular geographic region or social grouping within the speech community.” (Lewis et al., 2013) The elements in N_i^t and N_j^t represent exactly this aspect of dialects; their presence does not disqualify dialects from being counted as a single language.

Hence, we consider it appropriate to ignore elements in N_i^t and N_j^t when judging whether groups i and j speak the same language.

Defining a language in this way allows us to determine the speakership of each language, and then count the number of languages in the region. This can be done for any period, and becomes particularly important in subsection 2.4.2, when the model is extended to a dynamic setting where speakerships may change from period to period.

The cost of acquiring new linguistic elements within a period depends on the sizes of the sets acquired, E_i^t and N_i^t,[17] and is assumed to have the following (time independent) functional form:

$$C\left(|E_i^t|, |N_i^t|\right) = c(|E_i^t|) + c(|N_i^t|)$$

where $c : [0, \infty] \to \mathbb{R} \cup \{\infty\}$ is a strictly increasing and strictly convex function with $c(0) = 0$.[18] Assume that c is the following quadratic:

$$c(|\cdot|) = \frac{1}{2}|\cdot|^2, \tag{2.1}$$

where $|\cdot|$ is the measure of a set. Convexity represents the increasing marginal costs that are likely to occur with such activities.

2.3.3 Strategic Interactions

Each pair of groups in the region interacts periodically in one of three activities: cooperation, competition, or non-interaction.[19]

---

[17] In this paper, we use the words "size" and "measure" interchangeably.

[18] This functional form assumes that there is no complementarity between the costs of the two modes of acquisition. Learning another language and inventing new expressions occur in very different contexts. Learning happens in an inter-group environment, in which multiple groups must spend time together; whereas inventing occurs in an intra-group context, where members of a single group spend time together inventing novel linguistic expressions. The social and mental resources required for engaging in these activities seem different enough to justify a zero cross-partial derivative.We do not believe that relaxing the additive separability assumption would change our results in a qualitative way. At the very least, we conjecture that our conclusions would still hold if we allow for a sufficiently small cross partial between the costs in the two modes of acquiring new linguistic elements.

[19] This subsection deals with what happens in a typical period t, so we suppress the superscript t on the linguistic variables.

Formally, we model a cooperative interaction as a pair of groups trying to coordinate on accomplishing some task using language. The expected payoff from a cooperation is increasing in the measure of the intersection $(L_i \cup E_{ij}) \cap (L_j \cup E_{ji})$.[20]

The intuition is that, since i and j use language to coordinate their actions, the more linguistic elements they have in common, the more accurate their communication will be, and consequently the better the two groups will perform in the cooperative interaction. As we show in section B.1, i's expected payoff from the cooperative interaction has the following reduced form:

$$u_i(E_i, E_j) = |(L_i \cup E_{ij}) \cap (L_j \cup E_{ji})|. \tag{2.2}$$

Observe that N_i and N_j do not affect u_i. It would thus be equivalent to define u_i as $|(L_i \cup E_{ij} \cup N_i) \cap (L_j \cup E_{ji} \cup N_j)|$, since $N_i \cap (L_j \cup E_{ji} \cup N_j) = \varnothing$.

A competitive game represents a situation where the two groups are antagonistic towards each other, for example when they are competing over the use of resources.[21]

We formally model the competitive game as a form of zero-sum game. The role of language in this context is to communicate *within* a group itself, for example for group i to organize a show of strength, an attack, or a plan for defense. With positive probability, i's within-group communication may be intercepted/overheard by someone from group j. If members of group j know a large portion of the elements in i's dialect, there is a high probability that an intercepted communication will be understood and used to group i's disadvantage. We show in section B.1 that the expected payoff from the competitive game has the following reduced form:

$$v_i(E_i, N_i, E_j, N_j) = \beta \left[|(L_i \cup E_i \cup N_i) \setminus (L_j \cup E_j)| - |(L_j \cup E_j \cup N_j) \setminus (L_i \cup E_i)| \right]. \tag{2.3}$$

This is the difference between how much of i's dialect is private from j, and how much of j's dialect is private from i. The first term in the square brackets captures i's ability to conceal information from j, and the second term reflects j's ability to conceal information from i. Therefore, i's expected payoff from the competitive game is increasing in the former and

---

[20] See section B.1 for an explanation

[21] See, for example, Dawson (1891). The real world counterparts to this game, however, need not involve actual violence.

decreasing in the latter.[22] The parameter β represents the relative magnitude of competitive payoffs vs. cooperative payoffs. Note also that $v_i(\cdot) = -v_j(\cdot)$.

The null game can be interpreted as a scenario in which i and j do not interact. When a pair of groups play a null game, each gets a payoff of zero with certainty.

For any matched pair i and j, the probability of a cooperative interaction occurring is p_{ij}, a competitive game q_{ij}, and a null game $1 - p_{ij} - q_{ij}$. These probabilities are related to the distance between i and j, d_{ij}, and the geography of the region through a parameter r which summarizes the region's geographic conditions.

Let the relative probability of the cooperative vs. competitive game in a given period be

$$\frac{p_{ij}}{q_{ij}} = r, \tag{2.4}$$

where r is exogenous and constant throughout the region.[23] The fraction p_{ij}/q_{ij} is the defining characteristic of the region and is the same for all pairs $i, j \in \mathscr{G}$.

Let $\pi : \{0, \ldots, D\} \to [0, 1]$ be a function that relates the geographic distance between i and j, d_{ij}, to the probability that these two groups will interact either cooperatively or competitively, i.e. $p_{ij} + q_{ij}$. Then, the probability of i and j playing a null game, i.e. they do not interact in a period, is

$$1 - p_{ij} - q_{ij} = 1 - \pi(d_{ij}). \tag{2.5}$$

We assume that $\pi(\cdot)$ satisfies the following property:

$$\frac{\pi(d)}{\pi(d+1)} \geq 2^d, \qquad \forall d \in \{0, \ldots, D\} \tag{2.6}$$

with the requirement that $\pi(0) = 1$.[24] Property (Equation 2.6) implies that $\pi(\cdot)$ is a decreasing function. Thus, the farther apart two groups are geographically (i.e. the larger d_{ij} is), the less likely they will interact with each other. Furthermore, property (Equation 2.6)

---

[22] Notice that the functionality distinction is only relevant in a cooperative game. Thus, even if both i and j learn the same (measurable) subset of elements from a third group k, knowledge of that subset of elements does not affect the payoffs of a competitive game between i and j.

[23] Our theory does not depend critically on the relationship between r and any specific geographic variables, as long as it is determined jointly by a set of relevant geographic factors, and is roughly constant throughout the region. For example, a high r could mean high complementarity between the resources in the region, so that groups need to cooperate in order to produce final goods or diversify their diet.

[24] An example would be $\pi(d) = \gamma 2^{-d(d-1)/2}$, where $\gamma = \pi(1) \leq 1$.

requires that the probability of interaction between any two groups drops sufficiently fast as the distance between them increases. In other words, we assume that the probability of a group interacting with one degree d neighbor is greater than the sum of the probabilities of interacting with all of its degree $d+1$ neighbors.[25]

Together, (Equation 2.4) and (Equation 2.5) imply that

$$p_{ij} = \frac{r}{1+r}\pi(d_{ij}) \qquad \text{and} \qquad q_{ij} = \frac{1}{1+r}\pi(d_{ij}). \tag{2.7}$$

The parameter r and the set of pairwise distance measures, $\{d_{ij} : i, j \in \mathscr{G}\}$ are exogenous and constant over time. They are also common knowledge. Within a given period, each group plays a total of $G-1$ games, one with every other group in the region. In every period, therefore, a total of $G(G-1)/2$ games are played.

The timing of events within a period, illustrated by Figure 2.2, is as follows:

1. At the beginning of a period, each group observes $\{|L_j \setminus L_i| : i, j \in \mathscr{G}\}$. In other words, each group knows the measure of the set that exists to be learned from every other group's dialect.
2. $\{E_i\}_i$ are chosen simultaneously.
3. $\{N_i\}_i$ are chosen simultaneously.
4. A period game is then drawn for every possible pair of groups, according to the probabilities described in (Equation 2.5) and (Equation 2.7).
5. Lastly, all $G(G-1)/2$ games are played, and payoffs are realized.

In subsection 2.4.1, we derive a group's optimal decisions, N_i^* and E_i^*, for a typical period. In subsection 2.4.2, we examine the long run implications of these short run results.

2.4 Results

2.4.1 Short Run

Our first result establishes the optimal size of N_i for every group i. We show that each N_i^*, the optimal subset of linguistic elements acquired by group i from L_0, has the same measure.

---

[25] This assumption is necessary for the proof of 2.1, which shows that i always prefers to learn, when possible, from a closer neighbor's dialect than from a farther neighbor's. Without this simplifying assumption, we would lose tractability when characterizing the equilibrium.

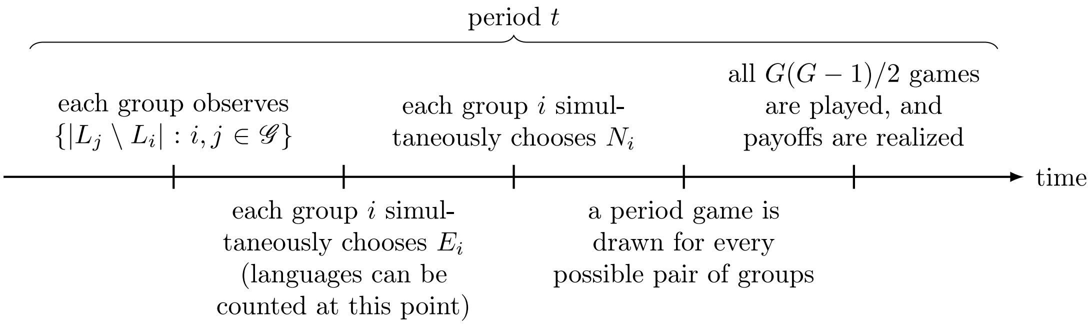


Figure 2.2: Timing of events in a typical period t

Proposition 2.1 (Optimal size of N_i). *For $i \in \mathscr{G}$, we have $|N_i^*| = |N^*|$, where*

$$|N^*| = \frac{\beta}{1+r} \sum_{k=1}^{D} 2^{k-1} \pi(k). \tag{2.8}$$

Proof. From (Equation 2.2) and (Equation 2.3), the *ex ante* expected payoff of group i is

$$\begin{aligned} U_i(\cdot) = &\sum_{j \neq i} p_{ij} \left[|(L_i \cup E_{ij}) \cap (L_j \cup E_{ji})| \right] - c(|E_i|) \\ &+ \sum_{j \neq i} \beta q_{ij} \left[|(L_i \cup E_i \cup N_i) \setminus (L_j \cup E_j)| - |(L_j \cup E_j \cup N_j) \setminus (L_i \cup E_i)| \right] - c(|N_i|). \end{aligned} \tag{2.9}$$

Observe that only the second line involves N_i. Thus group i's maximization problem is

$$\max_{|N_i|} \sum_{j \neq i} \beta q_{ij} \left[|(L_i \cup E_i \cup N_i) \setminus (L_j \cup E_j)| - |(L_j \cup E_j \cup N_j) \setminus (L_i \cup E_i)| \right] - c(|N_i|).$$

Since the marginal benefit is positive everywhere, cost is strictly convex in $|N_i|$, and $c'(0) = 0$, there exists a unique interior maximum. The maximum is given by the first order condition where marginal cost is equal to marginal benefit.[26] Notice that marginal benefit of $|N_i|$ is constant, described by

$$\beta \sum_{j \neq i} q_{ij} = \beta \sum_{j \neq i} \frac{1}{1+r} \pi(d_{ij}) = \frac{\beta}{1+r} \sum_{k=1}^{D} 2^{k-1} \pi(k),$$

where the first equality follows from equation (Equation 2.7) and the last equality follows from the fact that i has 2^{k-1} degree k neighbors. The marginal cost of acquiring elements is

---

[26] Linear benefit and strictly convex cost ensures that the second order condition holds as well.

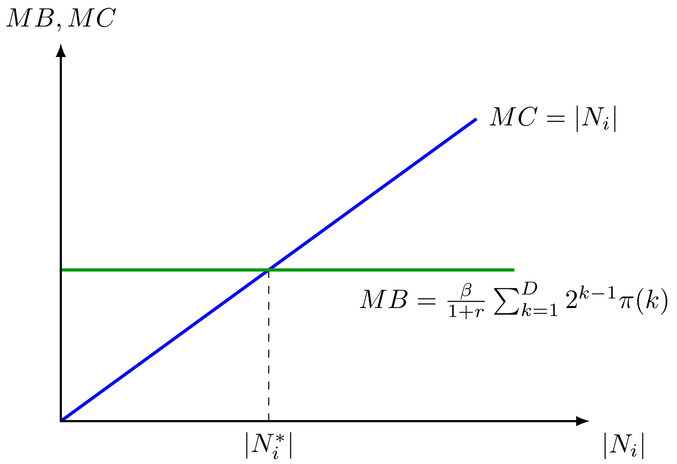


Figure 2.3: Optimal $|N_i^*|$

just $|N^*|$, according to (Equation 2.1). Therefore condition (Equation 2.8) is precisely the first order condition. Since the problem is symmetric for every i, it follows that (Equation 2.8) holds for all $i \in \mathscr{G}$. $\square$

The intuition for this result is the standard marginal analysis in economics, as depicted in Figure 2.3. Acquiring linguistic elements from L_0 means inventing new expressions that no other groups but i can understand. Doing so enhances i's ability to keep secrets from all other groups, which in turn raises i's expected payoff in all of its competitive games. The marginal benefit is therefore the sum of the probabilities of playing a competitive game with each other group in the region, weighted by the relative magnitude of competitive payoffs to cooperative ones, β. At the (interior) optimum, the marginal benefit must equal marginal cost of inventing new linguistic expressions. Observe that $|N^*|$ depends only on r, β and D, which are all exogenous parameters and constant over time. Therefore, $|N^*|$ is constant over time as well.

To obtain a similar characterization for $|E_i^*|$, we need to put some structure on the initial set of dialects, $\{L_i\}_{i\in\mathscr{G}}$. Specifically, we restrict attention to sets of dialects that are localized, as defined below.

Definition 2.3 (Localization)**.** The set of dialects $\{L_i\}_{i\in\mathscr{G}}$ is *localized* if the following condition is satisfied:

$$\ell \in L_i \setminus L_j \quad \Rightarrow \quad \ell \notin L_k, \qquad \forall i, j, k \in \mathscr{G} \text{ such that } d_{ij} < d_{ik}. \tag{2.10}$$

Localization means that if there is an element that a group j can learn from L_i, then that element cannot also be in the dialect of any of i's and j's mutually farther neighbors.[27] In other words, localization requires that whenever L_i and a distant neighbor's dialect L_k have a common element, then all the dialects of i's closer neighbors must also have that element.

Localization captures the intuition that it is very improbable for two distant languages to independently develop a common word-meaning pair. Evidence from historical linguistics suggests that the elements of pre-colonial languages at least roughly satisfy the localization property. When a linguistic group splits into two, the speakers of the two new dialects naturally tend to live close together. Studies of the Indo-European languages, for example, found that languages with a more recent common linguistic ancestor—e.g. Spanish and Portuguese, both of which share many common features of Latin, their "parent language"—are also geographically close to each other (Finegan, 2008).

According to Pagel et al. (2007), "[l]anguages, like species, evolve by a process of descent with modification". Linguistic divergences are also commonly represented as trees, similar to those relating biological species (Kruskal et al., 1971; Pagel, 2000; Esteban et al., 2012). Suppose that i and j have a more recent common linguistic ancestor than i, j, and k. An element that satisfies any one of the three following conditions satisfies the localization property: i) all three retain the element from a common ancestor, ii) an ancestor of both i and j split from k and then acquired it before i and j split from each other, or iii) any element acquired by a single group after all three had split.

There is also a theoretically appealing reason for our focus on symmetric and localized sets of dialects. As we show later in 2.2, if the region begins with a set of symmetric and localized dialects, then the set of dialects will always be symmetric and localized. If all groups in a neighborhood of some degree $d \in \{0, 1, \ldots, D\}$ speak the same language, then the set of dialects is localized. Localization, however, holds for a much more general set of dialects.

Assuming localization enables us to establish the order in which a group i learns from existing dialects. This order is closely connected to the magnitude of the marginal benefit of learning a subset of another dialect. Let $MB_i(d_{ij}, s) = 2^{d_{ij}-s}(p_{ij} + \beta q_{ij})$ denote the marginal

---

[27] It is worth noting that both i and j are equally distant from k when $d_{ij} < d_{ik}$. In Figure 2.1, suppose i, j, k are Sites 1, 2, and 3, respectively. Sites 1 and 2 are degree 1 neighbors with $d_{12} = d_{21} = 1$, and both are degree 2 neighbors of Site 3 with $d_{13} = d_{23} = 2$. Thus, $d_{12} < d_{13}$ implies that $d_{21} < d_{23}$.

benefit that i receives from learning a subset of elements that is shared by $2^{d_{ij}-s}$ neighbors of degree d_{ij}, where $s \in \{1, \ldots, d_{ij}\}$.[28]

Lemma 2.1 (Order of Learning)**.** *Suppose the set of dialects* $\{L_i\}_i$ *is localized. Then,* $MB_i(d_{ij}, s)$ *is decreasing in* d_{ij}*. As a result, it is always optimal for* i *to learn all of* $\bigcup_j (L_j \setminus L_i)$ *before learning anything from* $\bigcup_k (L_k \setminus L_i)$*, where* j, k *are such that* $d_{ij} < d_{ik}$*.*

Proof. From 2.3 and equation (Equation 2.2), it follows that after learning all the elements of all the languages where $d < d_{ij}$, learning a subset of elements $E_{ij} \subseteq L_j \setminus L_i$ is only going to affect i's payoff when it is interacting with neighbor(s) of degree d_{ij}. According to (Equation 2.9) and (Equation 2.7), therefore, given an arbitrary $d_{ij} \in \{1, \ldots, D\}$, $MB_i(d_{ij}, s)$ is at least

$$p_{ij} + \beta q_{ij} = \frac{r+\beta}{1+r}\pi(d_{ij}), \tag{2.11}$$

as is the case when $s = d_{ij}$, and at most

$$2^{d_{ij}-1}(p_{ij} + \beta q_{ij}) = 2^{d_{ij}-1}\left(\frac{r+\beta}{1+r}\pi(d_{ij})\right), \tag{2.12}$$

as is the case when $s = 1$. $s = 1$ is the case where all of i's degree d_{ij} neighbors know the element, so it will be useful in games with any of them. Observe that (Equation 2.11) and (Equation 2.12) are the same when $d_{ij} = 1$, because i only has one degree 1 neighbor. Observe also that the lowest marginal benefit of learning from a degree d_{ij} neighbor, i.e. $MB_i(d_{ij}, d_{ij})$, is higher than the highest marginal benefit of learning from a degree $d_{ij} + 1$ neighbor, i.e. $MB_i(d_{ij} + 1, 1)$:

$$\frac{r+\beta}{1+r}\pi(d_{ij}) - 2^{(d_{ij}+1)-1}\left(\frac{r+\beta}{1+r}\pi(d_{ij}+1)\right) \geq 0, \qquad \forall d_{ij} \in \{0, \ldots, D\}.$$

The inequality follows from (Equation 2.6). This completes the proof. □

---

[28] Localization, together with symmetry of dialects (see 2.4), imply that within each degree d_{ij} neighborhood, $MB_i(d_{ij}, s)$ has at most d_{ij} values, and s indicates the s^{th} highest value. Take $d_{ij} = 3$ for example. Group i has four neighbors of degree 3. Symmetry and localization require that any linguistic element i learns must be either (i) commonly shared by all four neighbors, (ii) shared by two groups (who are degree 1 neighbors with each other), or (iii) unique to one of the four groups. Hence, the three possible values of $MB_i(3, s)$ would be, from the highest to lowest, $MB_i(3, 1) = 4(p_{ij} + \beta q_{ij})$, $MB_i(3, 2) = 2(p_{ij} + \beta q_{ij})$, and $MB_i(3, 3) = (p_{ij} + \beta q_{ij})$. Bear in mind that some of these steps need not exist. For example, when $L_i = L_j$ (up to a subset of measure zero) for all $i, j \in \mathscr{G}$, $MB_i(d_{ij}, s) = 0$ for all d_{ij}. If not all dialects in the region are identical, however, symmetry and localization implies that the lowest step always exists.

2.1 allows us to partition the set of i's learnable elements, $\bigcup_k (L_k \setminus L_i)$, according to the marginal benefit they confer, and hence the neighborhood degrees. Let

$$P_i(d) = \begin{cases} \{\bigcup_k \left(L_k \setminus (\bigcup_{k'} L_{k'})\right) : d_{ik} = d \text{ and } d_{ik'} < d\} & \text{if } d \in \{1, \ldots, D\} \\ \varnothing & \text{if } d = 0 \end{cases}$$

describe the elements of this partition.

$P_i(d)$ is the set of learnable linguistic elements in the dialects of i's degree d neighbors, excluding the elements that are in the dialects of i's closer neighbors, of degrees less than d. Note that elements of i's own language are never learnable to i; since i is a degree 0 neighbor with itself, $P_i(d) \cap L_i = \varnothing$ for all d. Moreover, for any $d \neq d'$, $P_i(d)$ and $P_i(d')$ are disjoint.

It follows, therefore, that $|\bigcup_d P_i(d)| = \sum_d |P_i(d)|$. By 2.1, it also follows that, for any $A \subseteq P_i(d)$ and $A' \subseteq P_i(d')$ where $|A| = |A'|$ and $d < d'$, the marginal benefit of learning $|A|$ is strictly greater than that of learning $|A'|$. Moreover, each $P_i(d)$ can be further partitioned into d cells, based on how many neighbors of degree d share those elements.

By the definition of $MB_i(d_{ij}, s)$, the more neighbors sharing an element, the higher the marginal benefit that element confers. Therefore, assuming localization, we can order i's set of learnable elements by their associated marginal benefits: first by neighborhood degree d, and then by s within each $P_i(d)$. Such an ordering, together with 2.1, implies that as $|E_i|$ increases, marginal benefit decreases in a step-wise fashion.

Proposition 2.2 (Optimal size of E_i^*). *Let the set of dialects $\{L_i\}_i$ be localized. The optimal size of E_i^* is contained within the following interval:*

$$|E_i^*| \in \left[\sum_{k=0}^{d^*} |P_i(k)|, \quad \sum_{k=0}^{d^*+1} |P_i(k)|\right),$$

where $d^* \in \{0, \ldots, D\}$ is determined by

$$\frac{r+\beta}{1+r}\pi(d^*) \geq \sum_{k=0}^{d^*} |P_i(k)| \qquad \textit{and} \qquad \frac{r+\beta}{1+r}\pi(d^*+1) < \sum_{k=0}^{d^*+1} |P_i(k)|. \tag{2.13}$$

Our dynamic results in subsection 2.4.2 do not require that we know the exact value of $|E_i^*|$, only the degree of neighborhood d^* within which i chooses to learn all the remaining elements of its neighbors dialects.

Proof. First, observe that a unique $|E_i^*|$ exists. The size of acquired elements from existing dialects, $|E_i|$, takes a value from a compact set $\left[0, \sum_{k=0}^{D} |P_i(k)|\right]$, on which the objective function (Equation 2.9) is continuous. Hence a maximum exists. Since marginal benefit

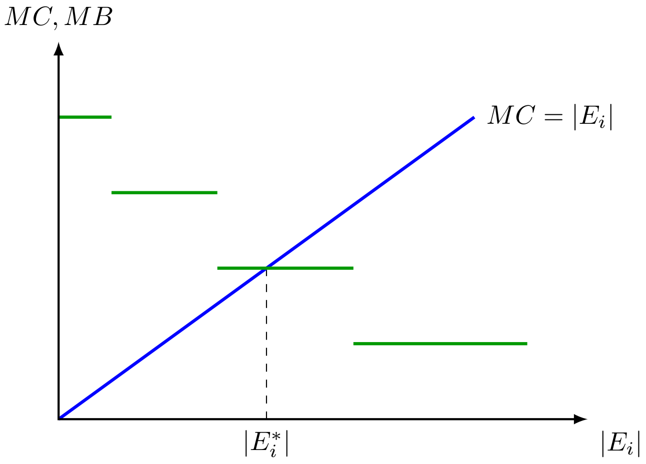


Figure 2.4: Optimal $|E_i^*|$

is weakly decreasing in $|E_i|$ while marginal cost is strictly increasing, uniqueness of $|E_i^*|$ is ensured. Figure 2.4 provides an illustration.

We can verify that $|E_i^*|$ is indeed bounded by the proposed interval through the first order condition. Notice that the benefit function is an increasing, piece-wise linear function that is not differentiable at a finite number of points, namely the points at which marginal benefit makes a discrete jump downwards. Nevertheless, for each value of $|E_i|$, there exists a set of (super-)derivatives for the benefit function, bounded by the left- and right-derivatives at $|E_i|$, and the set is a non-singleton at the points where the benefit function has a kink. The first order condition requires that at $|E_i^*|$, there exists a (super-)derivative of benefit function that is equal to the derivative of the cost function (the latter of which is uniquely defined at every $|E_i|$). Moreover, it has to be true that (i) at $|E_i| = \sum_{k=0}^{d^*} |P_i(k)|$, marginal benefit is weakly higher than marginal cost; and (ii) at $|E_i| = \sum_{k=0}^{d^*+1} |P_i(k)|$ marginal cost exceeds marginal benefit. But according to (Equation 2.13), d^* is chosen such that these two conditions are simultaneously satisfied. The marginal benefit of learning the last element in $P_i(d^*)$ is described by (Equation 2.11); for otherwise i would have learned it sooner. The marginal cost of learning this last element is $\sum_{k=0}^{d^*} |P_i(k)|$. Therefore, equation (Equation 2.13) ensures that d^* is chosen such that the marginal benefit of learning the last element in $P_i(d^*)$ is higher than the marginal cost. However, the marginal cost exceeds the marginal benefit if $P_i(d^*+1)$ is fully acquired.

We know $|E_i| = \sum_{k=0}^{d^*} |P_i(k)|$ when everything up to and including $P_i(d^*)$ is fully learned and $|E_i| = \sum_{k=0}^{d^*+1} |P_i(k)|$ when everything up to and including $P_i(d^* + 1)$ is fully learned. Therefore, $|E_i^*|$ must lie between these two values. □

Corollary 1. *At the optimum, group i will acquire (i) all the learnable elements from its neighbors with degree smaller than or equal to d^*; (ii) a proper subset of learnable elements from its degree d^*+1 neighbors; and (iii) no elements from its neighbors with degree greater than d^*+1.*

Proof. This follows directly from 2.1 and 2.2. □

Furthermore, if the region's dialects satisfy a symmetry condition, then 2.2 immediately implies that all groups learn the same measure of elements in a typical period.

Definition 2.4 (Symmetry)**.** The set of dialects $\{L_i^t\}_{i\in\mathscr{G}}$ is *symmetric* at t if the following two conditions are satisfied:

$$|L_i^t| = |L_j^t|, \qquad \forall i,j \in \mathscr{G} \tag{2.14}$$

$$|L_i^t \cap L_j^t| = |L_i^t \cap L_k^t|, \qquad \forall i,j,k \in \mathscr{G} \text{ such that } d_{ij} = d_{ik}. \tag{2.15}$$

Condition (Equation 2.14) requires that the measure of each dialect is the same. Condition (Equation 2.15) requires that the intersection of any dialect L_i^t with that of any equally distant neighbors has the same measure.

Corollary 2. *If the set of dialects $\{L_i\}_i$ is symmetric and localized, then the optimal size of E_i^* is the same for all groups in any given period, i.e. $|E_i^*| = |E^*|$ for all $i \in \mathscr{G}$.*

Proof. This follows directly from the properties of symmetry and localization of the set of dialects, and the symmetry of each group's decision. □

While 2.2 makes no requirement about which subset of the degree d^*+1 neighbors' dialects a group should learn, we make two assumptions about how groups learn from their d^*+1 neighbors' dialects. The first assumption ensures that the set of dialects stay symmetric and localized in the subsequent period, so that results derived in this section can be applied to analyze long-run properties of the region's languages in a dynamic context. The second assumption enables us to use the definition of same language (i.e. 2.2) consistently in the dynamic setting.[29]

Assumption 5. When a group is indifferent between learning from several dialects, it will learn the same measure of elements from each of them; the elements learned are chosen randomly from the set of learnable elements in those dialects.

---

[29] Without 6, we would have to modify the criterion for same language to a slightly more complicated version: $L_i^t \cup \left(\bigcup_{k=1}^{d^{t*}} E_{ik}\right) = L_j^t \cup \left(\bigcup_{k'=1}^{d^{t*}} E_{jk'}\right)$. The other aspects of the model are unaffected by this assumption.

Assumption 6. All groups within a degree d^* neighborhood learn the same subset of elements from their degree $d^* + 1$ neighbors.

Notice that the characteristics of the set of learnable elements may vary from period to period. Consequently, unlike $|N^*|$, $|E^*|$ need not be the same across time.

2.4.2 Long Run

In this subsection, we examine the long run implications of the model. Suppose that a period represents a human generation of approximately twenty years. Each new generation of group members inherits the dialect of the previous generation, and then may choose to learn from the neighbors' dialects as well as to invent novel expressions.

We assume that groups are myopic, so that each generation is only interested in payoffs in the current period.[30]

Myopic play implies that each group's decision in any given period is characterized in subsection 2.4.1. This allows us to analyze the trajectory of a region's linguistic composition over generations.

Definition 2.5. The *linguistic composition* in the region at time t is the set

$$\Lambda^t = \left\{ (\lambda, \widetilde{\mathscr{G}}_\lambda) : \lambda = \bigcup_{k \in \widetilde{\mathscr{G}}_\lambda} (L_k^t \cup E_k^t \cup N_k^t) \text{ where } \widetilde{\mathscr{G}}_\lambda \text{ is the set of groups speaking the same language } \lambda \right\}.$$

Recall that the criterion for "same language" is given in 2.2. We are interested in the steady state of a region's linguistic make-up; that is, a linguistic composition that is stable in the long run. We denote such a composition $\overline{\Lambda}$, wherein each language $\lambda \in \overline{\Lambda}$ is spoken by the same subset of groups $\widetilde{\mathscr{G}}_\lambda \subseteq \mathscr{G}$ in every subsequent period.[31] Henceforth, we use a superscript t on the variables to index time. In accordance with 2.2, we formally define a steady state as follows:

---

[30] This assumption frees us from considering repeated game effects. Group myopia over payoffs is not implausible. Group membership in small scale societies is fluid across generations. Most small scale societies are patrilocal or matrilocal, so adults expect that approximately half of their offspring, at maturity, will leave for another group.

[31] Note however that the languages themselves will not be the same; they will grow in size over time, according to the results in subsection 2.4.1.

Definition 2.6. The region's linguistic composition is in a steady state if for all t,

$$L_i^t \cup E_i^t = L_j^t \cup E_j^t \quad \Rightarrow \quad L_i^{t+1} \cup E_i^{t+1} = L_j^{t+1} \cup E_j^{t+1} \tag{2.16}$$

and

$$L_i^t \cup E_i^t \neq L_j^t \cup E_j^t \quad \Rightarrow \quad L_i^{t+1} \cup E_i^{t+1} \neq L_j^{t+1} \cup E_j^{t+1} \tag{2.17}$$

for all $i, j \in \mathscr{G}$.[32]

A linguistic composition is in a steady state if the following two conditions hold: (i) every pair of groups i, j that speak the same language in any period t will continue to speak the same language in $t+1$; and (ii) every pair i, j that do not speak the same language in t will not speak the same language in $t+1$. Therefore, a steady state of the linguistic composition is a set $\overline{\Lambda}$ of languages, each of which is spoken by the same subset $\widetilde{\mathscr{G}}_\lambda$ of groups over time.

Recall from subsection 2.3.3 that the order of events in a typical period t is as follows:

1. $\{|L_j^t \setminus L_i^t| : i, j \in \mathscr{G}\}$ is observed by all groups.

2. $\{E_i^t\}_i$ are chosen simultaneously.

    - The number of languages for period t, $\#(\Lambda^t)$, is counted at this point.[33]

3. $\{N_i^t\}_i$ are chosen simultaneously.

    - The set of dialects at the beginning of period $t+1$, $\{L_i^{t+1}\}_i = \{L_i^t \cup E_i^t \cup N_i^t\}_i$, is determined at this point.

4. A period game is drawn for every possible pair of groups according to (Equation 2.5) and (Equation 2.7).

5. The period games are played, and payoffs are determined based on $\{L_i^t, E_i^t, N_i^t\}_i$.

For the results in subsection 2.4.1 to apply in every period, it must be the case that the set of dialects is symmetric and localized at the beginning of every period. The following lemma shows that if $\{L_i^t\}_i$ satisfies symmetry and localization in the initial period, it will continue to do so in subsequent periods.

---

[32] In our model, the sizes of the languages are growing over time. Hence, the steady state of a linguistic composition is steady in the sense that each language in the composition has a stable speakership.

[33] $\#(\Lambda)$ is the number of elements in Λ.

Lemma 2.2. *Suppose* $\{L_i^t\}_i$ *is symmetric and localized. Then* $\{L_i^{t+1}\}_i$ *is also symmetric and localized.*

The proof of this lemma can be found in subsection B.2.1.

Before we state our main dynamic results, it helps to define an important value, $\overline{d}$. As we will state formally in 2.4 and 2.5, $\overline{d}$ is the threshold value of neighborhood distance. Groups with $d_{ij} \leq \overline{d}$ will speak the same language in the steady state, and groups with $d_{ij} > \overline{d}$ will speak different languages.

Definition 2.7. Define $\overline{d}$ as an element of $\{0, \ldots, D\}$ that satisfies the following condition:

$$\frac{r+\beta}{1+r}\pi(\overline{d}) \geq \sum_{k=0}^{\overline{d}} |P_i^t(k)| \qquad \text{and} \qquad \frac{r+\beta}{1+r}\pi(\overline{d}+1) < \sum_{k=0}^{\overline{d}+1} |P_i^t(k)|, \tag{2.18}$$

where

$$|P_i^t(k)| = \begin{cases} 2^{k-1}|N^*| & \text{if } k \in \{1, \ldots, \overline{d}\} \\ 0 & \text{if } k = 0 \end{cases}$$

and $|P_i^t(k)| \geq 2^{k-1}|N^*|$ if $k \geq \overline{d}+1$.

The particular neighborhood distance $\overline{d}$ has some special properties. The next proposition, for example, establishes that if $d_{ij} > \overline{d}$ for a pair of groups i, j, then their languages will, in some sense, become less similar over time. This is because the size of the set that j invents every period, $|N_j|$, will always be larger than the size of the set that i learns from j, $|E_{ij}^*|$. With each passing period, the set of elements that i speaks but j doesn't will become larger. Due to symmetry, the set of elements that j speaks but i doesn't will also become larger.

Proposition 2.3. *If* $d_{ij} > \overline{d}$ *then* $|E_{ij}^*| < |N_j|$. *Thus,* $|L_j \setminus L_i|$ *increases every period.*

The proof can be found in subsection B.2.2. 2.3 is a result regarding the trajectory of a pair of languages. From (Equation 2.5), $\pi(d)$ is the probability of a non-null interaction occurring between the groups. According to (2.6), $\pi(\cdot)$ is a decreasing function of d, so the further apart i and j are geographically, the less they interact. 2.3 establishes that if meaningful interactions between i, j are sufficiently infrequent, that is, if d is high enough, their dialects will become less and less similar over time. This process might look empirically similar to linguistic drift, as both types of divergence theoretically decrease with frequency of interaction.

The next proposition builds on the results of 2.3, establishing that if $d_{ij} \leq \overline{d}$, and i,j begin by speaking the same language, then their languages will continue to be the same. That is, in every period i will learn all of j's newly invented elements, and vice versa.

Proposition 2.4 (Existence and characterization of a steady state)**.** *There exists a steady state, $\overline{\Lambda}$, which is characterized by the following condition:*

$$L_i^t \cup E_i^t = L_j^t \cup E_j^t \quad \Leftrightarrow \quad d_{ij} \leq \overline{d}, \qquad \forall i, j \in \mathscr{G}, \tag{2.19}$$

where $\overline{d} \in \{0, \ldots, D\}$ is determined by (Equation 2.18).

The proof of this proposition can be found in subsection B.2.3. The proposition establishes that there is a steady state of the region's linguistic composition which can be fully described by a particular degree of neighborhood, $\overline{d}$. In this steady state, all groups that are members of the same $\overline{d}$ neighborhood speak the same language, and furthermore only groups which are members of the same $\overline{d}$ neighborhood speak the same language. In other words, if $d_{ij} \leq \overline{d}$ then $L_i \cup E_i = L_j \cup E_j$, and if $d_{ij} > \overline{d}$ then $L_i \cup E_i \neq L_j \cup E_j$. Note that $\overline{d}$ depends only the exogenous parameters of the model ($|N^*|$ is also a function of those parameters). Thus, $\overline{d}$ is constant over time.

Corollary 3 (Steady state number of languages)**.** *The number of languages in the steady state is $\#(\overline{\Lambda}) = G/2^{\overline{d}}$.*

Proof. This follows directly from condition (Equation 2.19). Each degree $\overline{d}$ neighborhood is inhabited by $2^{\overline{d}}$ groups, and these groups share a common language in the steady state. Since there are G groups, the number of languages in the steady state is therefore $G/2^{\overline{d}}$. □

Next we turn our attention to convergence, showing that for a general set of initial conditions, the linguistic composition converges towards the $\overline{\Lambda}$ described in 2.4.

Proposition 2.5. *If $\{L_i^0\}_{i \in \mathscr{G}}$ is symmetric and localized, then after a finite number of periods, the steady state $\overline{\Lambda}$ is reached.*

The proof of this proposition is presented in subsection B.2.4. This proposition says that, after finitely many periods, i will learn the entire set of learnable elements from its degree $\overline{d}$ neighbors and will therefore, by definition, speak the same language as each of them. It is worth noting the two special initial linguistic compositions where all groups have identical initial dialects or completely distinct ones both satisfy the symmetry and localization conditions.

Every period, i learns a larger set of the languages of its $d \leq \overline{d}$ neighbors than those groups collectively invent. In other words, $|E_i^*| > 2^{\overline{d}-1}|N^*|$. Suppose i does not initially speak the same language as its $d = 1$ neighbor, and $\overline{d} \geq 1$. In the first period, i begins "catching up" with its $d = 1$ neighbor, then later with its $d = 2$ neighbors, and so forth until it has learned the entire set of learnable elements from its neighbors of degree $d \leq \overline{d}$. At this point, $L_i \cup E_i = L_j \cup E_j$ holds for all of i's $d \leq \overline{d}$ neighbors.

From this point onwards, i learns at least $2^{\overline{d}-1}|N^*|$ elements every period, including all those invented in the last period by its neighbors of degree $d \leq \overline{d}$. i may learn some elements from its degree $\overline{d}+1$ neighbors, but as was shown in 2.3, not as many as this set of neighbors invents. For this reason, i never catches up with these neighbors.[34] From Corollary 3, $\#(\overline{\Lambda}) = G/2^{\overline{d}}$, therefore $G/2^{\overline{d}}$ is the steady state number of languages. Last but not least, since the steady state is uniquely determined by $\overline{d}$, we conclude that it is the unique steady state for the set of initial linguistic compositions that we consider.

2.4.3 Comparative Statics

Now we return to our original purpose: to show the effect of cooperative and competitive incentives on the number of languages in a region. In this section, we establish the key comparative static result for the number of languages in the steady state: $\#(\overline{\Lambda})$ is weakly decreasing in r, the ratio of the probabilities of cooperative and competitive interactions.

Furthermore, we examine the results of exogenous changes in the geographic locations of groups on the relationships between particular dialects. Specifically, we determine under what circumstances such a change would alter the steady state identity of the dialects of an arbitrary pair of groups.

Proposition 2.6. *The steady state number of languages, $\#(\overline{\Lambda})$, is weakly decreasing in r.*

In other words, the higher the regional ratio of the probability of cooperation vs. competition, the fewer languages there will be in the region in steady state.

Proof. According to Corollary 3, $\#(\overline{\Lambda}) = G/2^{\overline{d}}$. We show that $\#(\overline{\Lambda})$ is weakly decreasing in r by proving that $\overline{d}$ is weakly increasing in r. Recall that $\overline{d}$ is determined by (Equation 2.18). In particular, it must be the case that, for any $\overline{d} \in \{1, \ldots, D\}$, the first inequality

---

[34] The divergence of i's language from those of its higher degree neighbors actually occurs immediately, as a result of its choice of E_i in the first period.

of (Equation 2.18) can be rearranged as follows:

$$
\begin{aligned}
\frac{r+\beta}{1+r}\pi(\bar{d}) &\geq (2^{\bar{d}}-1)|N^*| \\
\frac{r+\beta}{1+r}\pi(\bar{d}) &\geq (2^{\bar{d}}-1)\frac{\beta}{1+r}\sum_{k=1}^{D}2^{k-1}\pi(k) && \text{by (2.8)} \\
(r+\beta) &\geq \frac{2^{\bar{d}}-1}{\pi(\bar{d})}\left[\beta\sum_{k=1}^{D}2^{k-1}\pi(k)\right]. && (2.20)
\end{aligned}
$$

Similarly, for any $\bar{d} \in \{0, \ldots, D-1\}$, the second inequality of (Equation 2.18) can be written as

$$(r+\beta) < \frac{2^{\bar{d}+1}-1}{\pi(\bar{d}+1)}\left[\beta\sum_{k=1}^{D}2^{k-1}\pi(k)\right]. \tag{2.21}$$

(Equation 2.20) and (Equation 2.21) together imply that

$$\frac{2^{\bar{d}}-1}{\pi(\bar{d})}\left[\beta\sum_{k=1}^{D}2^{k-1}\pi(k)\right] \leq r+\beta < \frac{2^{\bar{d}+1}-1}{\pi(\bar{d})}\left[\beta\sum_{k=1}^{D}2^{k-1}\pi(k)\right], \tag{2.22}$$

where the terms in the square brackets are independent of either r or $\bar{d}$. Note that the fraction $(2^k-1)/\pi(k)$ is strictly increasing in k, because $\pi(\cdot)$ is decreasing. Thus, when the increase in r is sufficiently large, $\bar{d}$ would also have to increase in order for (Equation 2.22) to hold. Likewise, when the decrease in r is large enough, $\bar{d}$ needs to decrease as well to satisfy (Equation 2.22). This completes the proof. □

Proposition 2.7. *The steady state number of languages, $\#(\bar{\Lambda})$, is not affected by multiplying the probability of a non-null interaction, $\pi(d)$, by any positive constant, $\phi > 0$.*

Proof. The statement follows trivially from (Equation 2.22): note that a constant ϕ multiplying $\pi(\cdot)$ gets canceled out immediately. □

This result shows that an exogenous change in the frequency of meaningful interactions in the region, as long as it does not affect either the relative frequency of cooperative vs. competitive interactions, r, nor the set of ratios $\pi(d)/\pi(d+1)$, $d \in \{1, \ldots, D-1\}$, will have no effect on the steady state number of languages in the region.

2.5 Evidence

2.5.1 Stylized Facts

Fact 1: Linguistic diversity is related to geography. The *Ethnologue*, a common reference for language classification, documents 7,105 human languages that are currently spoken worldwide (Lewis et al., 2013).[35] These languages are distributed very unevenly on earth.Papua New Guinea, for example, accounts for only 0.3% of the world's land mass, yet it is home to 12% of all living languages. Australia, by contrast, accounts for 5.1% of the world's land area, but hosts merely 3% of its languages.[36] Further examples of regions with high linguistic diversity include sub-Saharan west Africa and south-central Mexico (Lewis, 2009).[37] Examples of areas with low linguistic diversity include northern Asia, Australia, and Brazil. In previous studies, scholars have documented relationships between geography and linguistic diversity. The most notable among those studies, from an economics perspective, is by Michalopoulos (2012), who finds that variance in land quality and in altitude are positively and significantly correlated with linguistic diversity.

Scholars in several disciplines have documented empirical correlations between linguistic diversity and various ecological and geographic factors. Using historical data on native North American populations at the time of European contact, Mace and Pagel (1995) find a significant positive correlation between linguistic diversity and the diversity of mammal species, both of which, in turn, exhibit a pronounced negative latitudinal gradient. On a global scale, Harmon (1996) documents a positive correlation between linguistic and biological diversity. Also on a global scale, Nettle (1998) identifies climate as a key factor influencing global language distribution. In particular, he observes that areas with low rainfall and short growing

---

[35] The *Ethnologue* uses the ISO 639-3 standard, which classifies languages using a three-letter coding system. The basic criteria that the *Ethnologue* uses to identify languages (as opposed to, for example, dialects) are (i) mutual intelligibility between speakers of variants of a language; and (ii) existence of a common literature or of a common ethnolinguistic identity. (Lewis et al., 2013) The criterion for including a language in the *Ethnologue* is that it must be "known to have living speakers who learned [it] by transmission from parent to child as the primary language of day-to-day communication". (Lewis, 2009)

[36] These figures are the authors' calculation based on data from WolframAlpha (www.wolframalpha.com) and Lewis et al. (2013). Specifically, the world's land area is $1.4894 \times 10^8\,\mathrm{km}^2$, that of Papua New Guinea is $462,840\,\mathrm{km}^2$, and of Australia, $7.618 \times 10^6\,\mathrm{km}^2$. The number of languages currently in use is 836 for Papua New Guinea and 245 for Australia.

[37] Following the precedent of Michalopoulos (2012), we use the term *linguistic diversity* to refer to the number of languages in a region. Other terms, such as "language diversity" (Nettle, 1998; Pagel, 2000), and "density of language groups" (Mace and Pagel, 1995), are also used to refer to similar though not necessarily identical concepts.

season sustain fewer languages. Michalopoulos (2012) shows that contemporary linguistic diversity is related to geographic heterogeneity. He finds that variation in land quality (suitability for agriculture) and in elevation are both positively and significantly correlated with the number of languages in a particular region. Both Michalopoulos (2012) and Nettle (1998) find that average precipitation is significantly positively correlated with linguistic diversity.

Fact 2: languages change over time. Furthermore, human linguistic diversity changes over time. Pagel (2000) estimates that between 130,000 and 500,000 languages have been spoken on this planet, and that linguistic diversity peaked between 20,000 and 50,000 years ago. Despite this historical plurality, the number of languages is decreasing rapidly. Recent evidence suggests that we are on a trend towards the "hegemony" of a few dominant languages. It has been estimated that between 50% and 90% of the current languages will not survive to the next century (Hale et al., 1992; Austin and Sallabank, 2011), although there are also many examples of linguistic groups in the process of reviving their traditional languages (Bentahila and Davies, 1993).

Another puzzle is the diversity of trajectory between pairs of languages or dialects. A living language changes constantly, borrowing words from other languages, developing new ones, and abandoning others. Each language has many variants, or so-called dialects, over both space and time, such as Canadian English vs. Australian English, or Modern English vs. Middle English. Two languages can merge into one. When they merge, within a generation or two, the daughter language is referred to as a creole (Hall, 1966). A language may also split and diverge many times. All of the $1,200$ modern Malayo-Polynesian languages, for example, are believed to have descended from a common linguistic ancestor (Gray et al., 2009).

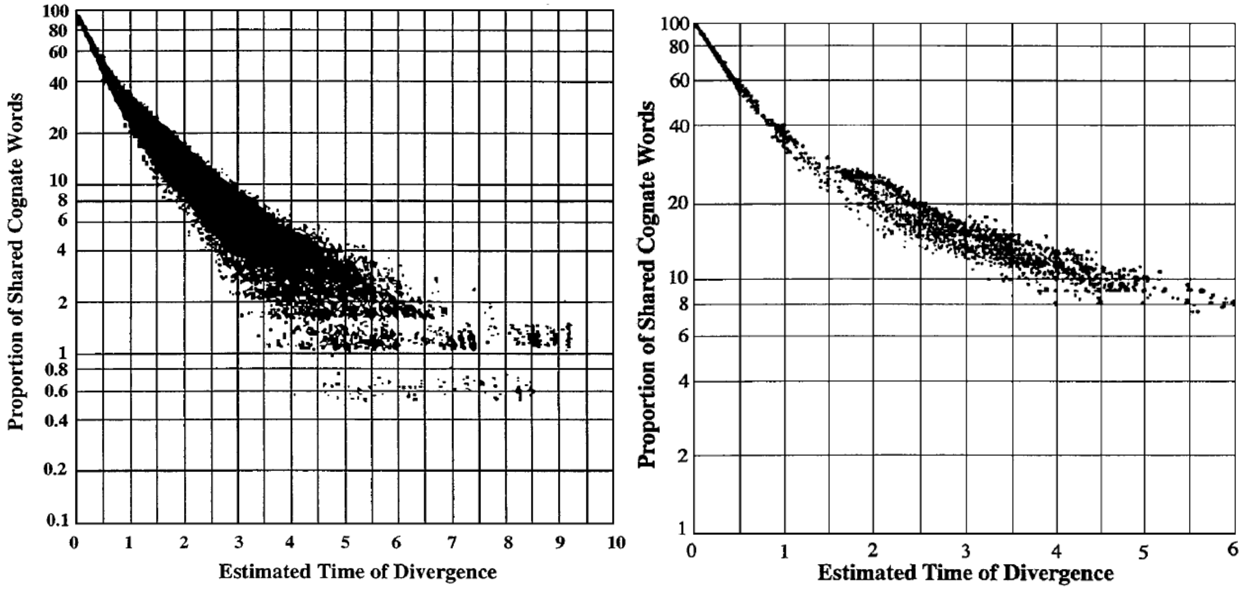
There are abundant examples of cooperative incentives resulting in linguistic change. During the colonial period, many linguistic groups developed new trading relationships with each other. As a result, hundreds of 'pidgin' languages emerged (Hall, 1966). Pidgins are informal linguistic hybrids that arise when two groups of formerly isolated peoples need to communicate extensively with one other. In subsequent generations, a pidgin may become a creole, which has a more formal set of vocabulary and grammatical rules and serves as a first language for many people. Some examples of creoles are the Chavacano language in the Philippines, Krio in Sierra Leone, and Tok Pisin in Papua New Guinea (Hall, 1966). Cooperative incentives can also cause linguistic change on a much smaller scale. The English words "ranch", "alligator", and "barbeque", for example, were borrowed from Spanish with only slight modification (Simpson et al., 1989).

There is a great deal of anecdotal evidence that conflict can lead to the invention of new words in the short run, and to linguistic divergence in the long run. As an example of the first case, Kulick (1992) reports that a group of Selepet speakers in Papua New Guinea changed their word for "no" from *bia* to *bune* to be distinct from a neighboring group. He also observed another Papua New Guinean linguistic group switching all of their masculine and feminine words—the words for mother and father, for example, switched in meaning. Kulick states that "people everywhere use language to monitor who is a member of their tribe." It is very difficult to acquire the ability to speak a non-native language, λ, perfectly. If a person manages to do so, then she is very likely to have a close relationship with members of the group whose native language is λ.

Most social groups innovate linguistically, as exemplified by nicknames, slang, and "inside" jokes. Such terms can be useful in situations of inter-group conflict: for example, teenagers vs. parents, police vs. criminals, high school A vs. high school B, or government vs. insurgents. In these cases, a group can gain a strategic advantage by using newly invented expressions to identify group membership and communicate without revealing information to outsiders. A well documented and extensive case of this type of linguistic invention began in East London during the 1840s (Partridge et al., 2008). Speakers of the English dialect Cockney began generating "rhyming slang" which is relatively easy for insiders to learn but incomprehensible to outsiders. 'Trouble and strife' means 'wife', for example, and 'Jack Jones' means 'alone'. Various hypotheses have been proposed to explain the appearance of Cockney rhyming slang. These include linguistic accident, a form of amusement, to confuse the authorities, or, as Hotten (1859) believed, by "chaunters and patterers", i.e. street traders, possibly to assist with collusion. Developing some group-specific language naturally facilitates the control of information, and thereby generates a competitive advantage.

On a larger scale, there are many examples of distinct languages that have emerged in high conflict situations. "Thieves' Cant", a term that refers collectively to dozens of dialects of English which arose during the 17th century in Great Britain, and were primarily spoken by criminal groups (Coleman, 2008). The Middle East (Goldschmidt and Davidson, 1991), the northern border area of Italy (Kaplan, 2000), Nigeria (Osaghae and Suberu, 2005; Suberu, 2001) and Papua New Guinea (Johnson and Earle, 2000) are all examples of regions with both long histories of conflict and according to the *Ethnologue*, very high current linguistic diversity.

There are also many cases where groups in a competitive environment invest resources in maintaining a distinct language. An important category is the revival of indigenous languages from near extinction, such as Halkomelem in Canada (Galloway, 2007), Welsh in



a Malayo-Polynesian Languages

b Indo-European Languages

Figure 2.5: Estimated Linguistic Divergence in Malayo-Polynesian and Indo-European Language Families (Pagel, 2000, p.399–400).

Wales (Aitchison and Carter, 2000), Xibe in China (Jang et al., 2011), and Basque in Spain (Gardner et al., 2000). Minority groups will often fiercely protect their languages against assimilation, as is the case for the QuÃ©bÃ©cois in Canada (d'Anglejan, 1984) and the Catalan in Spain (Roller, 2002). When facing conflict with sub-groups, a national government will sometimes try to eliminate the sub-group's language or dialect. Some examples of such elimination include the mandatory residential school system for aboriginal children in Canada in the 19th and 20th centuries (Milloy, 1999), and the establishment of the L'Académie française in 1635 with the goal of standardizing the French language (Rickard, 1989). Both policies arose during periods of conflict, the former over control of land and the latter over centralization of power.

Fact 3: Languages diverge rapidly when they first diverge. If drift is the only force affecting rates of linguistic divergence, there should be an upper bound on how quickly languages can diverge. Assuming a constant or increasing degree of isolation, the rate of divergence should be roughly constant over time. However, Figure 2.5, taken from Pagel (2000), show that rates of divergence are particularly high initially when a pair of groups undergoes a linguistic split, then decrease quickly, and finally flatten out as time progresses.

Note that the scale on the y axis is logarithmic, so the change over time in the rate of divergence is significantly more pronounced than it appears.

Although this is beyond the scope of our current model, we hypothesize that a single language would split into two during a period of high interior conflict. In this paper we assume a constant underlying probability of conflict to co-operation over time, but it is reasonable to suppose that this ratio could vary with changes in underlying variables, and apply to internal conflict as well as external conflict.

During a period of internal conflict then, each faction would find it beneficial to differentiate their language. As the pair of dialects became less mutually intelligible, the incentives to differentiate them further might reasonably diminish Furthermore, reversion to the mean suggests that if the initial probability of conflict is high enough to cause a linguistic group to split, this probability would likely decrease over time. In this case incentives to differentiate would decrease, and the pair of languages would differentiate more slowly. In both graphs, the rate of divergence between any pair of languages clearly starts high, decreases quickly initially and then tapers off slowly.

Under a drift theory, the rates of divergence should be roughly the same for pairs of groups that are completely isolated from one another. An interesting fact about Figure 2.5 is that there is a large range of divergence rates among Polynesian languages, and a somewhat lesser range of divergence rates among pairs of Indo-European languages. The higher variance in divergence rates among the Polynesian languages is consistent with strategic incentives affecting these rates. The Polynesian islands host a very large variety of political relationships, with some being intensely and chronically violent, and others almost perfectly peaceful (Younger, 2008).

2.5.2 Existing Explanations of Linguistic Diversity

Several theories have been put forth to account for the empirical relationship between linguistic diversity and its environmental covariates. Common among these theories is an informal model involving isolation and drift. We understand it to mean the following: random changes, or mutations, occur all time in language. If populations of a linguistic community become isolated, mutations in one group are not copied by the entire community, and the two populations slowly begin to speak different dialects, and eventually different languages. Drift in linguistics is therefore a phenomenon analogous to mutation and drift in population genetics; it is random and unconscious change that occurs in language over time. Once the dissimilarity passes some threshold of mutual unintelligibility, the dialects would be considered different languages.

We do not deny that there is random change in language over time, and that this process contributes to increased linguistic diversity. It is, however, insufficient to explain a large portion of it.

For each of the proposed explanations that follow, we would like to point out that it is relatively easy to make an argument that the geographic variable in question should have the opposite effect on isolation to the one proposed, and therefore the opposite effect on linguistic diversity.

Nettle (1998) and Pagel (2000) argue that regions with favorable geographic conditions, e.g. those that are conducive to steady food supply, tend to sustain small, self-sufficient groups which seldom interact with each other. Consequently, populations in those regions are separated into various linguistic communities. In unproductive territories, on the other hand, survival demands that people cooperate on a large scale, and so language there will likely be uniform.

Boyd and Richerson (1988) point out another channel through which geography may increase isolation and therefore linguistic diversity. They maintain that high geographic heterogeneity increases migration costs, and thus facilitates isolation of different groups. Then, due to the force of linguistic drift, the languages of these geographically separate communities become dissimilar over time. Our challenge to this theory is that high geographic heterogeneity also increases the incentives to either travel long distances to harvest resources that are locally scarce, or to trade for them, thus decreasing isolation.

Closest to our line of thinking is a short article by Pagel (2012). Pagel speculates that conflict over resources must be part of the reason why there are so many languages. He further mentions that humans are highly attuned to differences in speech, which helps them to identify the social group(s) to which individuals belong. (Our model formalizes and extends the suggestion given by Pagel (2012), that the multitude of languages exists to prevent people from understanding each other. He does not propose any explanation for why this might be the case.

2.5.3 Geography and conflict/co-op

The relationship between geography and conflict has received much attention from geographers and political scientists. Geographic variables that are found to affect regional levels of conflict include general resource abundance (Kratochwil, 1986; Fairhead, 2000; Le Billon, 2001; Renner, 2002; Smillie and Forskningsstiftelsen, 2002); high value resources such as gems, fuel, or narcotics (De Soysa, 2002; Fearon and Laitin, 2003; Buhaug and Lujala, 2005; Buhaug et al., 2009); presence of mountains (Fearon and Laitin, 2003; Buhaug and

Rød, 2006); forest cover (de Rouen and Sobek, 2004); and whether there is a rainy season (Buhaug and Lujala, 2005).[38]

There are several ways in which regional geography can affect potential gains from trade. Factors such as climate and soil differ across sites in a region, generating comparative advantages in the production of particular goods. The magnitudes of these comparative advantages, in turn, affect the potential for gains from trade. Secondly, geography affects the transaction costs of trading. In small scale societies the costs of travel, and therefore the costs of trade, depend on geographic variables including ruggedness of terrain, availability of freshwater, and the location of water bodies (Murrieta-Flores, 2010).

2.6 Discussion

Nothing can be more fundamental than language as a means of transmitting information between human beings. Yet languages themselves are complex, dynamic and imperfect. They can have a significant effect on economic variables. Understanding the mechanism behind linguistic change, therefore, is an interesting and worthwhile undertaking.

In this paper we contrast linguistic change due to incentives with the conventional explanation of change due to isolation and drift. While not explicitly modeling drift, the model does generate change that might be difficult to distinguish empirically from completely random drift. Consider groups i, j and k. i might make changes to its language because of strategic considerations with respect to k, but these changes also affect the relationship between the languages of i and j. Depending on the way linguistic similarity is measured, it might appear that the languages of i and j are 'drifting' apart. In addition, two groups that are sufficiently isolated from each other would not learn any of each others' language. If there exist closer groups with significant probability of conflict, these original two groups' language would 'drift' apart as they each invent new terms every period.

Both theories predict that the languages of a pair of geographically distant groups will become less similar over time. Further research is needed, particularly empirical research, to measure the importance of each mechanism in linguistic change.

The context of our model—small scale farming or foraging economies—may cause some readers to overlook its relevance to contemporary linguistic diversity. It is important, how-

---

[38] To our knowledge, no empirical study has been done to examine the effect of local variance in territorial quality and altitude on conflict. Since Michalopoulos finds these geographic variables to be positively correlated with linguistic diversity, we hypothesize that they would also be correlated conflict. Such evidence would link our theory to the empirical findings of Michalopoulos (2012); i.e. that variance in land quality and altitude are important determinant of linguistic diversity.

ever, to note that linguistic change is usually slow (Pagel et al., 2007). Sometimes a generation may not even be conscious about the changes occurring in their language. Anecdotal evidence of the English language suggests that it was already close to its present form before the advent of modern transportation and communication technology. A fluent English speaker today, for example, has only minimal difficulty in understanding Jonathan Swift's *Gulliver's Travels*, written in 1726, decades before the industrial revolution.

Despite our focus on small-scale societies, we do believe that cooperative and competitive incentives are important causes of linguistic change in the modern world. The factors affecting the type and frequency of interaction, however, are more complex and change quickly.

Chapter 3

The Economics of Early Warfare over Land[1]

---

[1] This chapter was previously published as Dow et al. (2017).

3.1 Introduction

Warfare is commonly defined as lethal conflict between organized groups. Here we investigate the incidence of early warfare among foragers and farmers in prehistory. Our focus is specifically on conflict over land, rather than other forms of lethal violence motivated by revenge, feuding, status competition, or the theft of portable wealth. We believe the insights derived from relatively simple economies of this kind are likely to have broader applicability to the study of warfare in general.

In our model, groups consist of individuals who maximize their expected food consumption. There is a production technology where food is obtained from labor and land, and a military technology where the probability of a successful attack on another group depends on the relative sizes of the competing groups. There are two sites or territories where food can be obtained, and their productivities may differ.

Groups are internally egalitarian in the sense that they share food equally and reach unanimous collective decisions. Attacks are motivated by a group's attempt to increase the quantity and/or quality of land it controls. To focus on cases in which wars are actually fought, we assume a group defends its land when it is attacked, because the alternative of fleeing to another location is even less attractive. A group that loses a war disappears, and the winning group spreads into the site or territory of the losing group.

Time periods are the length of a human generation. At the beginning of a period, each site has a population inherited from the past. Individual agents can move between sites, although this is generally costly. Once the final group sizes are determined, each group decides whether to attack the other. Finally, production takes place, agents have children, and fertility-driven Malthusian dynamics generate a new initial population for each site in the next period. The process then repeats.

We show that in this framework there are two necessary conditions for warfare. First, individual migration between sites must be costly enough. Second, there must be large enough technological or environmental shocks to the ratio of productivities across sites. Together, these conditions are sufficient.

The intuition is as follows. A large group with a poor site is tempted to attack a small group with a good site, because the large group has a high probability of winning and it gains valuable land if it does win. However, several forces tend to create a positive relationship between group size and site productivity, which typically deters attack. First, migration causes agents to congregate at high quality sites. Second, Malthusian fertility effects lead to higher populations at better sites. Third, warfare is self-limiting because it reallocates population toward better sites. These factors tend to promote peace. But war can occur if an exogenous shock creates a temporary imbalance between group size and productivity, and individual migration cannot remove this imbalance in the short run.

Our analysis suggests that warfare over land would have been rare among mobile foraging bands with exogamous marriage, where agents could respond to negative local shocks by exploiting kinship networks to join more prosperous groups. The analysis also suggests that warfare over land would have become more common after the emergence of sedentary foraging or farming groups with endogamous marriage and rigid membership criteria. Even in this case, exogenous shocks that changed the relative productivities of sites or territories would also have been necessary.

There is a small economic literature on conflict in prehistory. Baker (2003) develops a model where groups compete for access to land. In equilibrium there is no actual warfare. North et al. (2009) and Rowthorn and Seabright (2010) study the effects of early warfare on the emergence of institutions and agriculture, respectively. Bowles studies the effects of early warfare on the evolution of cooperation within social groups (Bowles, 2006; Choi and Bowles, 2007; Bowles, 2009; Bowles, 2012).

More generally, a number of economists and political scientists have developed theoretical models of conflict over resources. In economics, it is common to have a first stage where the distribution of power is determined by the choices of leaders, and a second stage with war or peace (see Garfinkel and Skaperdas, 2007; Acemoglu et al., 2012; Bhattacharya et al., 2015). In political science, the distribution of power is typically exogenous, and the focus is on the issue of whether bargaining can avert costly warfare (Fearon, 1995; Powell, 1996, 2006). Theorists from both disciplines agree that if complete contingent contracts were feasible, peace would prevail.

Fearon (1995) argues that asymmetric information, indivisibilities, or a lack of commitment devices can make it impossible to negotiate or enforce such contracts.

Relationships among pre-state groups of foragers or farmers are 'anarchic' in much the same way that relationships among states in the modern world are anarchic. Thus similar commitment problems arise. As Fearon (1995, 402) points out, repeated game effects cannot deter defections that will result in the death of one player. When a concession does not alter the prizes at stake or the probabilities of winning, groups will be tempted to renege after the concession is sunk. At best, a territorial concession by a weak group might reduce the benefit of an attack to a strong group and thus forestall it. However, compact sites with highly valued resources create indivisibility problems that are difficult to resolve in this way.

A key contribution of our paper is to show that even in this unpromising setting, powerful forces favor peace. In particular, we endogenize the probability of winning a conflict by having it depend on group sizes. The group sizes are determined in the short run by the migration decisions of individual agents, and in the long run by the effect of food income on fertility. Both mechanisms promote peace even without bargaining.

We also show that when individual mobility is costly enough, exogenous shocks can lead to warfare. Other authors have made related points. Powell (2006) finds that a rapid change in the distribution of power can induce warfare. In contrast to Powell, we focus on shocks to the value of the prizes at stake. Chassang and Miguel (2009) show that a temporary negative shock to labor productivity can cause warfare by reducing the current opportunity cost of fighting relative to the future value of the assets seized in war. We explore a different causal channel involving shocks to site-specific productivities, rather than aggregate shocks and intertemporal tradeoffs.

A growing literature supports the idea that environmental shocks can trigger war. Zhang et al. (2006, 2007) find a positive association between abnormally cold periods and domestic rebellions in China during AD 1000 - 1911. They argue that the causal mechanism involves Malthusian population growth during warm periods, followed by a contraction in food output in cold periods. Tol and Wagner (2010) find a similar pattern for most of Europe during AD 1000 - 1990, where unusually cold periods are associated with more violent conflict. Bai and Kung (2011) find that droughts are correlated with nomadic invasions of China between 220 BC and

1839 AD because drought had greater effects on food supply among pastoralists than among the farmers of central China. In a meta-analysis of econometric research using modern (post-1950) data, Burke et al. (2015) estimate that a one standard deviation rise in temperature increases intergroup conflict by 11.3%, with smaller but still substantial effects for drought or extreme rainfall.

Section 3.2 surveys the archaeological and anthropological evidence on warfare in early foraging and farming societies. Sections 3.3-3.5 develop the formal model through a process of backward induction. Section 3.3 treats both group sizes and site productivities as exogenous and derives conditions under which war or peace occurs. Section 3.4 moves back to the individual migration stage and characterizes locational equilibrium. Section 3.5 introduces our Malthusian fertility assumption and presents our necessary and sufficient conditions for warfare. We use the model to account for empirical evidence in section 3.6, and offer a few closing thoughts in section 3.7. Proofs of formal propositions are provided in an online appendix.

3.2 Evidence On Early Warfare

Our model assumes that groups are egalitarian with respect to consumption and decision-making. Small-scale foraging and farming societies often fit this description. Dow and Reed (2013) review archaeological evidence about foraging societies without stratification. Johnson and Earle (2000) provide ethnographic examples of foraging and farming groups with up to a few hundred members that do not have stratification (some do have war). These authors maintain that hereditary inequality and coercive leadership arise only in larger-scale chiefdoms and states, which are beyond the scope of this paper. According to Kelly (2013a, ch. 9), mobile foragers are generally egalitarian, apart from some gender and age distinctions, while sedentary foragers are more frequently stratified. Borgerhoff Mulder et al. (2009) show that inequality is associated with the inheritance of material assets, which is more common in agricultural than in hunter-gatherer societies.

Two archaeological markers are widely used to infer early warfare: skeletons showing signs of deadly force, and defensive structures or settlements in easily defended locations. Other evidence includes specialized weapons and artistic depictions. The latter are more relevant for chiefdoms or states than for smaller egalitarian groups.

Even when skeletons provide obvious evidence of violence, it can be difficult to distinguish warfare from individual homicide, execution, religious sacrifices, and non-warfare-related cannibalism. Uncertainties in the dating of skeletons can also make it difficult to distinguish cases in which many deaths occurred simultaneously (a massacre) from cases in which deaths occurred over decades or centuries (a cemetery).

Defensive structures and defensively located settlements, while evidence of the threat of warfare, are not evidence of actual warfare. They tend to deter groups that may be considering an attack. Such investments could have been undertaken in the belief that they would prevent war or reduce its likelihood. They could also have been motivated by a belief that in the event of an attack, the defenders would be more likely to prevail.

Ferguson (2013) reviews archaeological data on prehistoric warfare for Europe and the Near East. For Europe, the Upper Paleolithic shows negligible evidence of war, and barely any evidence of interpersonal violence. In the Mesolithic, from the onset of the Holocene around 11,600 BP (before present) until the arrival of agriculture, warfare is 'scattered and episodic'. This period is associated with increasing sedentism, more food storage, more distinctive group identities, and greater inequality. About 500-1000 years after the Neolithic transition to agriculture, warfare became widespread, and during the subsequent Copper, Bronze, and Iron Ages, warfare was the norm.

For the Near East, Ferguson's account begins with early Natufian society around 15,000 BP, which was pre-agricultural and similar to the European Mesolithic. There is no skeletal evidence of war in the southern Levant for the next 10,000 years. Likewise, there is no evidence of fortifications in this period. Peace came to an end in the southern Levant in the early Bronze Age around 3200 BC, a time that coincided with the formation of the Egyptian state. The apparent absence of war over this time span is remarkable in light of the intensive archaeological study the region has received; its history of climate shocks, technological innovation, and population fluctuation (Dow et al., 2009); and the presence of social stratification during some periods (Dow and Reed, 2013).

Ferguson reviews a substantial body of evidence indicating warfare in other parts of the Near East, including Anatolia and the northern Tigris area, at least as early as the Pottery Neolithic, beginning around 8400 BP. At some sites, warfare probably goes back to the Pre-Pottery Neolithic A, from 11,600 - 10,500 BP, when agriculture began to arise in southwest Asia. Warfare continued through the Copper and Bronze Ages.

The ethnographic literature provides direct evidence on warfare among hunter-gatherers. Modern hunter-gatherers may differ in important ways from those who lived thousands of years ago. In particular, contemporary hunter-gatherers generally live in extreme environments, and behavior in these societies has been influenced significantly by the outside world **(**Lee and Daly, 1999).

The Standard Cross Cultural Sample, or SCCS, is a representative sample of 186 well-documented and culturally independent pre-modern societies (Murdock and White, 1969, 2006). Two results from this data set stand out (Kelly, 2013a, 2013b). The first is that while there is evidence for warfare in all types of societies, the incidence of warfare is lower for egalitarian nomadic foragers than it is for non-egalitarian sedentary foragers. The second (first reported in Keeley, 1996) is that population pressure is correlated with warfare while population density is not, where population pressure refers to population relative to total food, and population density refers to population relative to total land.

Fry and Soderberg (2013) use the SCCS to study 148 episodes of lethal violence in 21 mobile forager bands. They conclude that "most incidents of lethal aggression may be classified as homicides, a few others as feuds, and a minority as war." (p. 270). Kelly (2000, ch. 2) observes that in a sample of 25 foraging societies drawn from the SCCS, war is strongly associated with segmented social structures that create group identities. Ember and Ember (1992) use the SCCS to study the effect of ecological variables on war. They find that by far the best predictor of warfare is a history of natural disasters, and that chronic scarcity has no independent effect on warfare.

Kelly (2013a) makes some additional observations about foraging societies that are particularly relevant to warfare over land. All such societies have ways of assigning individuals to specific tracts of land and allowing them to gain access to others (pp. 154-161). In some instances, individuals can move flexibly across social groups. There is a widespread but not universal pattern among foragers in which "connections to land are social and permeable rather than geographic and rigid" (p. 156). Perimeter defense is rare, but the need to ask permission to use another group's territory is common (it is usually granted). When foragers become sedentary, however, population is usually so high that residential movement is not possible without displacing another group; "war appears when mobility is not an option" (p. 205).

3.3 Production and Warfare

Consider two sites, i = A, B. Each agent is risk neutral and maximizes expected food consumption. Agents regard death as equivalent to being alive without food, where both yield zero utility (see section 7). Each agent is endowed with one unit of labor time. Agents are of negligible size relative to the total population at a site.

The production function for food is

$$(1) \qquad Y_i = s_i n_i^{\alpha} \qquad \text{with } 0 < \alpha < 1 \qquad\qquad i = A, B$$

where $s_i > 0$ is the productivity or quality of site i and n_i is the population of site i at the time of production. Site qualities are affected by permanent geographic features such as lakes, rivers, good soil, and diverse local ecosystems, as well as transient factors such as weather. The agents at a site share food equally, so in the absence of warfare each agent at site i receives the average product

$$(2) \qquad y_i = s_i n_i^{\alpha-1} \qquad\qquad i = A, B$$

The two sites together form a region with total population $n_A + n_B = N > 0$.

We will develop a model with the following steps in each period: (i) Malthusian dynamics determine an initial population for each site; (ii) individual agents can migrate between sites; and (iii) the resulting groups decide whether or not to attack. We proceed by backward induction. This section provides results for step (iii), treating group sizes (n_A, n_B) and site qualities (s_A, s_B)

as exogenous. Section 4 studies the location decisions at step (ii), and section 5 describes the population dynamics at step (i).

At step (iii), a war occurs if at least one group chooses to attack. Given that a war takes place, the probability that group i wins is

$$(3) \qquad p_i = n_i/(n_A + n_B) \qquad \text{or} \qquad p_i = n_i/N \qquad\qquad i = A, B$$

The probability of victory for a group is equal to its share in total population. This seems plausible for small-scale societies without specialized warriors or weapons. The military technology in (3) assigns no advantage to the attackers or defenders. Our functional form is the simplest of those considered by Garfinkel and Skaperdas (2007).

We make several other simplifying assumptions (see section 7 for a discussion). The payoffs are 'winner take all' in the sense that the winning group keeps its own site and gains the site of the opposing group. The winners suffer no casualties and spread their population across the two sites to maximize total food output. The losing group is killed or driven away (see Keeley, 1996) and receives a zero payoff.

When group i with population n_i wins the contest, its total food output is

$$(4) \qquad H(n_i) = \max \{s_A L_A^\alpha + s_B L_B^\alpha \text{ subject to } L_A \geq 0, L_B \geq 0, L_A + L_B = n_i\}$$

Due to the constant output elasticity, the equalization of marginal products across sites in (4) also equalizes the average products across sites. Each member of the winning group is indifferent between sites and there is no need for side payments among the winners.

We express H using the following notation. Define σ as in (5a) and φ as in (5b) below. Then write $H(n_i)$ as in (5c) and the optimal labor allocation as in (5d):

$$(5a) \qquad \sigma \equiv s_A/s_B \in (0, \infty)$$

$$(5b) \qquad \varphi(s_A, s_B) \equiv s_A/[1 + \sigma^{-1/(1-\alpha)}]^\alpha + s_B/[1 + \sigma^{1/(1-\alpha)}]^\alpha$$

$$(5c) \qquad H(n_i) = \varphi(s_A, s_B) n_i^\alpha$$

(5d) $L_A(n_i; \sigma) = n_i/[1 + \sigma^{-1/(1-\alpha)}]$ and $L_B(n_i; \sigma) = n_i/[1 + \sigma^{1/(1-\alpha)}]$

If a war occurs and group i wins, each member of this group has the food consumption

(6) $h(n_i) = H(n_i)/n_i = \varphi(s_A, s_B)n_i^{\alpha-1}$

When group $j \neq i$ attacks, group i is indifferent between attacking or not because a war occurs in either case. When group j does not attack, it is optimal for group i to attack if $p_ih(n_i) > s_in_i^{\alpha-1}$ where the left side is expected food per person from war and the right side is food per person from peace. A dominance argument implies that

(7) peace occurs if and only if $p_ih(n_i) \leq s_in_i^{\alpha-1}$ for i = A, B

We next consider how war and peace are determined by the group sizes (n_A, n_B) and site qualities (s_A, s_B). Using (6) and (7), the functions defined in (8) give the lower bounds on the probability of victory that would motivate an attack by each group.

(8) $x_A(\sigma) \equiv s_A/\varphi(s_A, s_B) \in (0, 1)$
$x_B(\sigma) \equiv s_B/\varphi(s_A, s_B) \in (0, 1)$

These functions depend only on the ratio $\sigma \equiv s_A/s_B$ and have the following features.

Lemma 1 (x_A and x_B as functions of σ).
(9a) $x_A(\sigma) + x_B(\sigma) > 1$ for all $\sigma > 0$
(9b) $x_A'(\sigma) > 0$ and $x_B'(\sigma) < 0$ for all $\sigma > 0$
(9c) $x_A(0) = 0$ and $x_A(\infty) = 1$
(9d) $x_B(0) = 1$ and $x_B(\infty) = 0$

Our first proposition uses the lemma to show that the occurrence of war or peace depends only on the ratio of the site qualities and the ratio of the group sizes.

<u>Proposition 1</u> (war and peace). Fix $s_A/s_B \equiv \sigma \in (0, \infty)$ and $n_A/n_B \in (0, \infty)$.

(a) If $0 < n_A/N < 1 - x_B(\sigma)$ then group B attacks and there is war.

(b) If $1 - x_B(\sigma) \leq n_A/N \leq x_A(\sigma)$ then neither group attacks and there is peace.

(c) If $x_A(\sigma) < n_A/N < 1$ then group A attacks and there is war.

Groups A and B never attack simultaneously. When average products are equal across sites (that is, $s_A n_A^{\alpha-1} = s_B n_B^{\alpha-1}$), case (b) holds with strict inequalities and there is peace.

These results are shown in Figure C.1. The productivity ratio σ is on the horizontal axis. Group A's population share n_A/N, which gives the probability that A wins a war if one occurs, is on the vertical axis. For a fixed productivity ratio σ, a war occurs if either A or B has a large enough population share. The boundary cases $n_A/N = 0$ and $n_A/N = 1$ lead to warfare in the trivial sense that the grand coalition attacks an empty site and wins with certainty. Peace occurs for intermediate population allocations. The dashed curve indicates the locus where average products are equal across sites, which is always in the interior of the peace region. The utility functions of the individual agents are

$$(10) \qquad \text{war:} \quad u_A(n_A, n_B) = [\varphi(s_A, s_B)/N]n_A^{\alpha} \quad \text{and} \quad u_B(n_A, n_B) = [\varphi(s_A, s_B)/N]n_B^{\alpha}$$

$$\text{peace:} \quad u_A(n_A, n_B) = s_A n_A^{\alpha-1} \quad \text{and} \quad u_B(n_A, n_B) = s_B n_B^{\alpha-1}$$

In wartime, an agent prefers to be in a larger group because this increases the probability of victory. In peacetime, an agent prefers to be in a smaller group because this increases the average product of labor.

3.4 Individual Mobility

We now move back to the migration stage and examine the location decisions of the individual agents. We assume an initial population allocation $m = (m_A, m_B) \geq 0$ with $m_A + m_B = N$. This starting point is exogenous here but will be endogenized in section 5. We want to know how individual migration determines a final allocation $n = (n_A, n_B) \geq 0$, which then leads to group decisions about war or peace as in section 3. Agents observe the site qualities (s_A, s_B) and ignore their own influence on the sizes of the groups at each site. Each agent takes the locations of all other agents as given and correctly anticipates whether war or peace will occur.

An agent initially located at site i who moves to site $j \neq i$ only receives a fraction $\eta \in (0, 1]$ of the utility u_j available to non-movers at $j \neq i$. The penalty $1-\eta$ may reflect a preference to live near family and friends, or to remain in a group with familiar customs and beliefs. It may also reflect a stigma imposed on outsiders by members of the other group. We use a multiplicative migration cost because under this assumption, only the relative group sizes will be important in our Malthusian model in section 5.

Equilibrium is defined so that no agent wants to change sites.

Definition 1 (locational equilibrium). Fix $\sigma > 0$ and $N > 0$. Consider allocations $n = (n_A, n_B) \geq 0$ with $n_A + n_B = N$. We call n a *locational equilibrium* (LE) in any of the following cases: $n = (N, 0)$; $n = (0, N)$; or $n > 0$ with

$$(11) \qquad u_A(n) \geq \eta u_B(n) \qquad \text{and} \qquad u_B(n) \geq \eta u_A(n)$$

The boundary cases (N, 0) and (0, N) are always equilibria because an agent who moves to an empty site is attacked by all other agents and loses with certainty, resulting in zero utility. In such boundary cases, only one of the inequalities in (11) holds because utility is positive at one site and zero at the other. In interior cases, both inequalities must hold.

The next task is to identify population distributions that lead to particular war or peace outcomes (as in Proposition 1) while simultaneously forestalling migration (as in Definition 1). For example, suppose we want an allocation leading to a war in which B attacks A, while ensuring that no individual agents flee their current locations. This is clearly true for $n = (0, N)$, which yields a trivial war and is also a locational equilibrium.

The interesting cases involve interior allocations. Now to have a war in which B attacks, we must satisfy both Proposition 1(a) and (11). For a given productivity ratio σ, Proposition 1(a) requires $n_A/n_B < (1-x_B)/x_B$. We obtain the utility levels for warfare from (10) and substitute these into (11) to obtain $\eta^{1/\alpha} \leq n_A/n_B \leq (1/\eta)^{1/\alpha}$. Combining these, we require $\eta^{1/\alpha} \leq n_A/n_B \leq \min \{(1/\eta)^{1/\alpha}, (1-x_B)/x_B\}$, where n_A/n_B must be strictly less than $(1-x_B)/x_B$. More generally, we obtain the following restrictions on the ratio n_A/n_B.

Proposition 2 (interior locational equilibria). Fix $\sigma > 0$ and $N > 0$. An allocation $n > 0$ is an LE if and only if it satisfies one of the following (mutually exclusive) conditions.

(a) B attacks: $n_A/n_B \in LE_B \equiv [\eta^{1/\alpha}, \min \{(1/\eta)^{1/\alpha}, (1-x_B)/x_B\}]$

(b) peace: $n_A/n_B \in LE_P \equiv [\max \{(\sigma\eta)^{1/(1-\alpha)}, (1-x_B)/x_B\}, \min \{(\sigma/\eta)^{1/(1-\alpha)}, x_A/(1-x_A)\}]$

(c) A attacks: $n_A/n_B \in LE_A \equiv [\max \{\eta^{1/\alpha}, x_A/(1-x_A)\}, (1/\eta)^{1/\alpha}]$

We require $n_A/n_B < (1-x_B)/x_B$ strictly in (a), and we require $n_A/n_B > x_A/(1-x_A)$ strictly in (c).

Figure C.2 shows a case where LE_B exists and LE_A does not (the interpretation of the arrows will be discussed later). Group size ratios n_A/n_B in the interval $(0, (1-x_B)/x_B)$ lead to an attack by B due to Proposition 1(a). To keep agents at A from fleeing to B or vice versa, we also require $\eta^{1/\alpha} \leq n_A/n_B \leq (1/\eta)^{1/\alpha}$ as explained above. In the case depicted in Figure C.2, the upper bound on LE_B is given by $(1/\eta)^{1/\alpha}$, but in other situations it could be given by $(1-x_B)/x_B$ instead.

The peace interval LE_P in Proposition 2(b) is always non-empty. When $\eta = 1$ so there is no mobility cost, LE_P contains the single point $n_A/n_B = \sigma^{1/(1-\alpha)}$ at which the average products are equal across sites. As η falls and mobility costs rise, the interval LE_P grows (see Figure C.2). When η is close to zero, every peaceful allocation in Proposition 1(b) is an LE and the interval LE_P in Figure C.2 becomes $[(1-x_B)/x_B, x_A/(1-x_A)]$.

The warfare intervals LE_B and LE_A in Propositions 2(a) and 2(c) may or may not exist. From Proposition 2(a), the interval LE_B is non-empty if and only if $\eta^{1/\alpha} < (1-x_B)/x_B$. Using the properties of $x_B(\sigma)$ from (9), this requires a sufficiently large σ, so site A is a sufficiently valuable prize for group B. When $\eta = 1$ so there is no mobility cost, this reduces to $x_B(\sigma) < 1/2$. Similarly, Proposition 2(c) implies that LE_A is non-empty if and only if $x_A/(1-x_A) < (1/\eta)^{1/\alpha}$. This requires a sufficiently small σ, so site B is a sufficiently valuable prize for group A. When $\eta = 1$, this reduces to $x_A(\sigma) < 1/2$.

Given a fixed value of σ, and assuming the relevant intervals exist, LE_B and LE_A both expand as mobility costs rise. This occurs simply because when migration is more expensive, agents are less inclined to move despite an impending war. When the warfare intervals exist, they may either be adjacent to the peace interval or separated from it by an interval of non-LE allocations. However, the ratios n_A/n_B in the warfare interval LE_B are always less than those in the peace interval LE_P because group B must be strong enough to make an attack worthwhile. Likewise, the ratios in LE_A always exceed those in LE_P.

Now consider a given initial (pre-migration) allocation $m \geq 0$. We want to know how the decisions of individual agents lead to a final (post-migration) allocation $n \geq 0$. When $m = (N, 0)$ or $(0, N)$, we already have a locational equilibrium from Definition 1 and no agents change sites. The same is true for any interior allocation $m > 0$ such that m_A/m_B is in one of the three LE intervals described in Proposition 2.

We therefore focus on cases where m is interior but not an LE. This implies that one of the conditions in (11) must be violated. Because the other condition in (11) must be satisfied, agents can only move in one direction. For example, if $u_B(m) < \eta u_A(m)$ then agents migrate from B to A and no agents migrate from A to B. In this situation we study allocations $n = (n_A, n_B) \geq 0$ with $n_A + n_B = N$ such that $n_A \geq m_A$. Migration stops when the system reaches an allocation where $u_B(n) = \eta u_A(n)$. This implies that $u_A(n) \geq \eta u_B(n)$ also holds, so we have an LE. If there is no interior $n_A \geq m_A$ at which this occurs, all agents go to site A, and the final allocation is $n_A = N$ and $n_B = 0$. If the initial allocation has $u_A(m) < \eta u_B(m)$, the logic is the same with the roles of A and B reversed. This reasoning leads to the following definition.

Definition 2 (initial and final allocations). Fix $\sigma > 0$ and $N > 0$. Also fix an initial (pre-migration) allocation $m \geq 0$. The final (post-migration) allocation n is as follows.

(a) If m is an LE then $n = m$. Otherwise:

(b) If $u_B(m) < \eta u_A(m)$ then n is the allocation with the smallest $n_A \geq m_A$ at which $u_B(n) = \eta u_A(n)$. When this condition cannot be satisfied, we set $n = (N, 0)$.

(c) If $u_A(m) < \eta u_B(m)$ then n is the allocation with the smallest $n_B \geq m_B$ at which $u_A(n) = \eta u_B(n)$. When this condition cannot be satisfied, we set $n = (0, N)$.

The intuition behind Definition 2 can be clarified by considering the directions in which agents flow. Suppose the system is not yet in locational equilibrium. When war is expected, the utility functions are those for war in (10), and agents flow toward the site with the larger current population. This process is self-reinforcing and can only end with a boundary equilibrium or a transition to the peace interval. When peace is expected, the utility functions are those for the peace case in (10), and agents flow toward the site with the higher average product. This reduces the difference in average products between the sites, and migration stops when this gap is small enough relative to the mobility cost. If the average product is initially higher at site A, migration stops at the lower bound of the peace interval LE_P in Proposition 2, and if the average product is initially higher at site B, migration stops at the upper bound of the peace interval LE_P.

For completeness, we need to consider the cases where LE_B is empty or not, and where LE_A is empty or not. The formal results are as follows.

<u>Proposition 3</u> (migration). Fix $\sigma > 0$ and $N > 0$. Choose an arbitrary initial allocation $m \geq 0$. The final allocation n is well defined and exactly one of the following cases applies.

(a) If $m = (0, N)$, $m = (N, 0)$, $m_A/m_B \in LE_B$, $m_A/m_B \in LE_P$, or $m_A/m_B \in LE_A$, then $n = m$. In the rest of the proposition, we limit attention to interior allocations $m > 0$.

(b) Suppose LE_B is empty. If m_A/m_B is below $(1-x_B)/x_B$ then $n = (0, N)$. If m_A/m_B is at least $(1-x_B)/x_B$ and below the lower bound of LE_P then n_A/n_B is equal to the lower bound of LE_P.

(c) Suppose LE_B is non-empty. If m_A/m_B is below the lower bound of LE_B then $n = (0, N)$. If m_A/m_B is strictly between the upper bound of LE_B and the lower bound of LE_P then n_A/n_B is equal to the lower bound of LE_P.

(d) Suppose LE_A is empty. If m_A/m_B is above $x_A/(1-x_A)$ then $n = (N, 0)$. If m_A/m_B is above the upper bound of LE_P and does not exceed $x_A/(1-x_A)$ then n_A/n_B is equal to the upper bound of LE_P.

(e) Suppose LE_A is non-empty. If m_A/m_B is above the upper bound of LE_A then $n = (N, 0)$. If m_A/m_B is strictly between the upper bound of LE_P and the lower bound of LE_A then n_A/n_B is equal to the upper bound of LE_P.

Figure C.2 shows the migration process for the case where LE_B exists and LE_A does not, where the arrows indicate the direction of migration. If m_A/m_B is below LE_B, group B will attack, we have $u_A(m) < \eta u_B(m)$, and agents flow from site A to site B. Using the war utilities in (10), this migration process raises u_B and lowers u_A, reinforcing the initial inequality. Population continues to flow in the same direction until all agents are at site B and we have $n = (0, N)$.

Next, suppose m_A/m_B is between LE_B and LE_P. If $m_A/m_B < (1-x_B)/x_B$ so group B will attack, we use the war utilities from (10). The fact that m_A/m_B is above LE_B implies $u_B(m) < \eta u_A(m)$, so agents flow from B to A. As long as B will attack, migration raises u_A and lowers u_B so the preceding inequality continues to hold. Once the system reaches $n_A/n_B = (1-x_B)/x_B$, we use the peace utilities from (10). Now migration from site B to site A lowers u_A and raises u_B. This continues until $u_B(n) = \eta u_A(n)$ at the lower bound of LE_P. Migration then stops and we have a peaceful LE.

Finally, suppose m_A/m_B is above LE_P. When peace is expected, population flows toward B and migration ends at the upper bound of LE_P. When A is expected to attack, population flows toward A until all agents are at site A and we have $n = (N, 0)$.

Several implications follow from Proposition 3. First, if the initial allocation m is not an LE, the system must move to one of four final allocations: (N, 0), (0, N), the lower bound of LE_P, or the upper bound of LE_P. None is consistent with non-trivial war. Thus, non-trivial warfare can only arise if the system starts from an initial population allocation m that is already in LE_B or LE_A, and mobility costs keep all agents at their initial sites. If the initial allocation m yields peace but is not an LE, the final allocation will also yield peace, and migration reduces the gap in average products across the sites.

The interval LE_P is stable in the sense that a small deviation to a point that is not an LE causes migration flows that bring the system back to LE_P. The same is true for the boundary cases (N, 0) and (0, N). But if the warfare intervals LE_A and LE_B exist they are unstable, in the sense that a small deviation to a point that is not an LE will lead either to a boundary case or an allocation in the peace interval LE_P. Points in the interior of LE_B or LE_A are neutrally stable because a small deviation leaves the system in LE_B or LE_A.

When mobility cost is zero ($\eta = 1$), three locational equilibria always exist: the boundary cases with trivial war, and the peaceful allocation with equal average products. If $x_B(\sigma) < 1/2$ then LE_B contains the single point $n_A = n_B = N/2$ and it is empty otherwise. A symmetric result holds for LE_A. At most one of these two interior warfare equilibria can exist. If either one does, it is unstable. Thus a positive mobility cost is a necessary condition for a stable equilibrium with non-trivial warfare.

We now pause to summarize the results so far. For an exogenously fixed ratio of site productivities $\sigma = s_A/s_B$, Proposition 1 derived the group size ratios n_A/n_B that lead to an attack by B, peaceful coexistence, or an attack by A. We then added the requirement that individual agents who foresee these outcomes not change sites. The group size ratios n_A/n_B consistent with this individual rationality condition were given by the intervals LE_B, LE_P, and LE_A in Proposition 2.

Next, we examined migration when the initial population allocation is not in one of these intervals. Proposition 3 showed that if the system is not already at a locational equilibrium with non-trivial warfare (LE_B or LE_A), then individual migration cannot take the system to such an equilibrium. Instead, it either pushes the system to a boundary case with trivial warfare, or to the boundary of the peace interval LE_P. In section 5, we study how migration costs and productivity shocks can interact to yield non-trivial wars.

3.5 Malthusian Dynamics

This section embeds the model from sections 3 and 4 in a Malthusian framework where population growth depends on food income. Dow and Reed (2013) use a similar framework to study inequality in small-scale societies. Ashraf and Galor (2011) provide evidence supporting the application of a Malthusian model to pre-industrial societies.

Denote periods by $t = 0, 1, \ldots$, where a period is one human generation. For an individual agent, let ρy be the number of that agent's children who survive to be adults, where y is the agent's food income and $\rho > 0$ is a constant. This captures the idea that adults with more food income are more fertile, have healthier children, or both. Adults in period t produce food, have children, and then die at the end of the period. The children become adults in period t+1 at the site where their parents produced food.

Let there be some sequence of productivity ratios $\sigma^t = s_A^t/s_B^t$ for t = 0, 1, . . These are given exogenously by technological or environmental factors. Each period starts with an initial allocation of agents $m^t = (m_A^t, m_B^t)$ inherited from the past. The agents can then move from their birth site to the other site as described in section 4. These moves (if any) give the final allocation $n^t = (n_A^t, n_B^t)$. Once the final allocation has been determined, the two groups decide whether to attack, and payoffs are obtained as in section 3.

Suppose first that peace occurs in period t. This can happen only when the final allocation $n^t > 0$ is interior. The food income of an agent at site A is $y_A^t = s_A^t (n_A^t)^{\alpha-1}$ and the food income of an agent at site B is $y_B^t = s_B^t (n_B^t)^{\alpha-1}$. The Malthusian linkage between food and surviving children gives $m_A^{t+1} = \rho y_A^t n_A^t$ and $m_B^{t+1} = \rho y_B^t n_B^t$. Thus

(12) peace in period t yields $$m_A^{t+1}/m_B^{t+1} = \sigma^t(n_A^t/n_B^t)^\alpha$$

Now suppose instead that a war occurs in period t and the winner is $W \in \{A, B\}$. Group W allocates its population n_W^t across sites to maximize total food output as in (4), while the opposing group disappears. Let the resulting number of agents at each site be L_A^t and L_B^t where $L_A^t + L_B^t = n_W^t$. Due to the fact that the production function from (1) has constant output elasticity, the equalization of marginal products across sites implies equalization of average products. Thus every member of W has the same food income $y^t = H(n_W^t; s_A^t, s_B^t)/n_W^t$. Using $m_A^{t+1} = \rho y^t L_A^t$ and $m_B^{t+1} = \rho y^t L_B^t$ along with (5d) implies that

(13) war in period t yields $$m_A^{t+1}/m_B^{t+1} = (\sigma^t)^{1/(1-\alpha)}$$

Notice that (13) applies both when $n^t > 0$ so the war is non-trivial, and also when the war is trivial because $n^t = (N^t, 0)$ or $(0, N^t)$.

<u>Lemma 2</u> (group size ratio). Fix $\sigma^t \in (0, \infty)$ and choose any interior allocation $m^t > 0$. The pair $(\sigma^t, m_A^t/m_B^t)$ suffices to determine the final allocation n^t; whether war or peace occurs in period t; and the group size ratio m_A^{t+1}/m_B^{t+1}, where $m^{t+1} > 0$ is interior.

The lemma shows that the initial group size ratio m_A^0/m_B^0 and the sequence of productivity ratios $\{\sigma^t\}$ determine the subsequent group size ratios m_A^t/m_B^t for all $t \geq 1$. This information also suffices to determine whether war or peace occurs in each period. Accordingly, our results do not depend on the absolute productivity levels (s_A^t, s_B^t), the absolute population levels N^t, or the identities of the winners in particular conflicts.

Now suppose we are given σ^t and m_A^t/m_B^t for period t. As in Lemma 2, we can compute m_A^{t+1}/m_B^{t+1}. We want to know the conditions under which a non-trivial war will occur in period t+1. This question is readily answered using Proposition 3. Such a war occurs if and only if the ratio m_A^{t+1}/m_B^{t+1} is in the set LE_B or LE_A. If neither is true, then either the ratio is in LE_P or m^{t+1} is not an LE. In the former case, we have peace. In the latter case, migration leads to a peaceful outcome or to formation of the grand coalition followed by a trivial war. The following proposition formalizes these ideas.

<u>Proposition 4</u> (war and peace with Malthusian dynamics). Fix $\sigma^t \in (0, \infty)$ and $m_A^t/m_B^t \in (0, \infty)$. Compute $m_A^{t+1}/m_B^{t+1} \in (0, \infty)$ as in Lemma 2. There is a non-trivial war in period t+1 if and only if the following two conditions are satisfied:

(a) $\eta^{1/\alpha} \leq m_A^{t+1}/m_B^{t+1} \leq (1/\eta)^{1/\alpha}$ and

(b) $\sigma^{t+1} \notin [\sigma_A^{t+1}, \sigma_B^{t+1}]$

where the bounds σ_A^{t+1} and σ_B^{t+1} are determined using m_A^{t+1}/m_B^{t+1} to compute m_A^{t+1}/N^{t+1}, and implicitly defining σ_A^{t+1} and σ_B^{t+1} using $m_A^{t+1}/N^{t+1} \equiv x_A(\sigma_A^{t+1}) \equiv 1 - x_B(\sigma_B^{t+1})$. These ratios have the property $\sigma_A^{t+1} < \sigma^t < \sigma_B^{t+1}$.

The necessity of (a) is straightforward. We have already established that any time individual agents migrate from one site to another, a non-trivial war is impossible. Thus we are limited to cases in which no migration occurs in period t+1. If a war is expected, the utility functions are obtained from (10). The absence of migration implies that these utilities must satisfy (11). Putting these conditions together, the initial population ratio in period t+1 (which is also the final population ratio) must satisfy Proposition 4(a). In this situation, the mobility costs incorporated in the parameter η are large enough to prevent migration despite an impending war.

The necessity of (b) has the following intuition. If we use Malthusian dynamics to compute m_A^{t+1}/m_B^{t+1} and take this ratio as given, we know the initial population share m_A^{t+1}/N^{t+1} for group A. Because we have ruled out migration as explained in connection with Proposition 4(a), this must also be the final population share n_A^{t+1}/N_A. As shown in Figure C.3, this fixes a horizontal line that determines an interval of productivity ratios σ^{t+1} consistent with peace. The latter is the interval $[\sigma_A^{t+1}, \sigma_B^{t+1}]$ in Proposition 4(b). To have a war, the productivity ratio σ^{t+1} must fall outside this interval, so $\sigma^{t+1} \notin [\sigma_A^{t+1}, \sigma_B^{t+1}]$.

The key ideas from Proposition 4 can now be summarized. Malthusian dynamics generate a population allocation at the beginning of period t+1. If a war is to occur, this allocation cannot be altered by migration. Moreover, given the probabilities of winning determined by the population shares, the productivity ratio must be high or low enough that an attack is attractive for one group or the other. The conditions in Proposition 4 are sufficient because for given initial group sizes, a war does occur when migration is costly enough and the productivity ratio is extreme enough.

Proposition 4 has the following important implication.

Corollary. If $\sigma^{t+1} = \sigma^t$ then there is peace in period t+1, regardless of whether there was war or peace in period t.

This follows from the fact that the productivity ratio σ^t is an interior point of the interval $[\sigma_A^{t+1}, \sigma_B^{t+1}]$ as indicated in Figure C.3. In order for war to occur, σ^{t+1} must involve a large enough shock relative to σ^t. When $\sigma^{t+1} = \sigma^t$, Proposition 4 already rules out a non-trivial war. The corollary strengthens this to exclude trivial wars involving the grand coalition.

The corollary shows that when the productivity ratio is constant over some time interval, there can be at most one war during this interval. If there is a war, it occurs at the start of the interval, and peace prevails thereafter. The reason is that war equalizes average products across sites, and hence equalizes population growth rates across sites. This maintains the existing group size ratio, which maintains the equality of the average products. Thus, peace continues. A series of wars therefore requires a series of shocks.

A useful thought experiment is to shut down individual migration by having η close to zero. This guarantees that (a) in Proposition 4 is satisfied. However, (b) is also necessary for war. In a stationary productivity environment, Malthusian fertility effects ensure that (b) never holds. Even without migration, this is enough to preserve peace.

Any exogenous variable that can alter the relative productivities of the sites can serve as a trigger for warfare. One such variable is production technology. To cause a series of wars, technical innovations must repeatedly change the productivity ratio. A candidate for such a process is the diffusion of agriculture across Europe, which was associated with increased warfare (see section 2).

Climate provides another source of exogenous shocks. Such shocks can easily change the relative productivities of sites or territories. For example, decreased rainfall may reduce productivity substantially at sites far from lakes or rivers but only slightly at sites with good access to surface water. Historical linkages between climate shocks and warfare are well documented (see section 1).

We close this section with some remarks on the relationship between our theory and that of Malthus (1798 [2004]). Our framework is 'Malthusian' only in the sense that we rely on a positive linkage between food per capita and fertility (or a negative linkage with child mortality). This is the operative check on population growth. We do not rely on mortality from warfare as a check on population. In fact, our model predicts that in a stationary productivity environment, regional population will converge to an equilibrium level without warfare.

Some anthropologists argue that warfare will occur in a demographic steady state. For example, Harris (1977, 60) suggests that reproductive pressure leads to both warfare and female infanticide, and that infanticide is "a savage but uniquely effective solution to the Malthusian dilemma". In contrast, we believe fertility effects can restrain population growth even without any mortality effects derived from warfare or infanticide.

3.6 Empirical Implications

Three mechanisms tend to promote peace in our model: mobility between groups, Malthusian fertility effects, and reallocation of population after a war. All three assign larger populations to better sites. The resulting positive association between group size and site quality tends to deter attacks. A productivity shock can trigger a war by creating a temporary imbalance between group size and site quality, provided that the imbalance cannot be undone through individual migration.

Our theory helps explain Ferguson's (2013) finding that in Europe, warfare was rare or non-existent in the Upper Paleolithic, sporadic in the Mesolithic, and common in the Neolithic and later. The small foraging bands of the Upper Paleolithic probably had fluid social boundaries, with easy individual mobility across groups through exogamous marriage and kinship networks. At very low population densities it would also have been difficult to exclude outsiders from a territory, leading to open access conditions (Dow and Reed, 2013). Moreover, the foraging bands of the Upper Paleolithic almost certainly had Malthusian fertility dynamics. Together, these factors would make warfare unlikely.

The shift to the Mesolithic and the Neolithic was associated with rising sedentism (Dow and Reed, 2015) and social segmentation (Kelly, 2000, ch. 4). This would make it harder for individuals to move among groups. In addition, community sizes were rising. Research on recent small-scale societies suggests that this probably made endogamous marriage more common (Dow et al., 2016), reducing kinship linkages across groups and increasing mobility costs for individuals. In combination with the shocks resulting from technological and environmental change, this would have made warfare more common.

Our results are also consistent with stylized facts from anthropology. The model accounts for the empirical finding that warfare is a function not of population density but rather of pressure on food resources. Individual migration or Malthusian dynamics can create high densities at good sites without leading to war. However, the model supports the idea that warfare occurs when a group has many mouths to feed and few resources with which to feed them. The model correctly predicts that war will be more common when shocks from natural disasters or climate change are larger and more frequent. And finally, our prediction that higher individual mobility costs make warfare more likely is consistent with ethnographic evidence that warfare occurs more often in societies with strict enforcement of group identities.

3.7 Concluding Remarks

We have made a number of assumptions about warfare that could be questioned. In many cases, altering our assumptions would make war less likely. For example, one could assume that winning groups suffer injuries or deaths; that winning groups may not gain full control over the losing group's land; that attackers are risk averse rather than risk neutral; that groups preparing an attack incur an opportunity cost in the form of lost food output; that defenders have advantages due to familiarity with local geography; that the defenders can build fortifications; that defenders may flee rather than stand and fight; that potential attackers may have kinship connections with defenders; and so on. Because our assumptions often stack the deck in favor of war, this strengthens our view that war was rare among small egalitarian groups in prehistory.

Another assumption that could be modified involves the utility function. We treat death as equivalent to being alive without food. But having no food may be better than death (perhaps more food will arrive tomorrow, and one has a chance of surviving until then). This 'happy to be alive' utility function makes the prospect of death seem worse in relative terms and would discourage warfare, much as risk aversion would. On the other hand, a minimum amount of food is necessary for life, so positive amounts of food below this subsistence threshold would be useless. This 'nothing to lose' utility function would make agents more inclined to attack when a food shortage looms, much as risk seeking would. Furthermore, it captures some additional Malthusian intuition about the causes of war. The question is whether these ideas have empirical implications. For example, the 'happy to be alive' case might apply when agents can smooth consumption by storage or pooling, while the 'nothing to lose' case might apply when expected food is low, variance is high, and smoothing is infeasible.

In principle, our framework applies both to hunter-gatherers and early farmers. One could argue that farmers make greater site-specific investments in land clearance, irrigation systems, terracing, and so on, in addition to actual planting of crops, and thus are more inclined to defend their sites. Our model already assumes that groups defend their sites when attacked, and such investments reinforce the incentive to do so. But if mobile hunter-gatherers are attacked, they may be more likely to flee, encouraging an attack. One could also argue that farmers and sedentary foragers who stay in a single location for many years have greater opportunities to make defensive investments that deter attack. Nevertheless, the evidence from section 2 suggests that warfare is more common among farmers than among mobile hunter-gatherers. This appears to be true even for egalitarian farming groups, although warfare undoubtedly increases when such groups develop into stratified chiefdoms or states.

We close with a remark on institutions. As discussed in section 2, the southern Levant enjoyed about 10,000 years during which warfare was almost non-existent. From the perspective of our theory, this might seem surprising, given the many environmental and technological shocks affecting the region in this period, including an early transition to agriculture. One possible explanation is that commitment problems were solved via third-party enforcement. However, this would have required something resembling a state, and there is no evidence for state formation during the period in question.

Our theory suggests an alternative: perhaps institutions of a more decentralized kind helped to transfer population across communities in response to shocks. Examples could include exogamous marriage, shared language, shared religion, or shared norms. Regional networking among local elites could also have facilitated individual migration when relative productivities changed. This is a possibility even for an agrarian society where warfare might otherwise have been expected. More generally, we believe good institutions can help to promote peace even when third party enforcement is absent.

Bibliography

Daron Acemoglu, Mikhail Golosov, Aleh Tsyvinski, and Pierre Yared. A dynamic theory of resource wars. *The Quarterly Journal of Economics*, 127(1):283–331, 2012.

J.W. Aitchison and H. Carter. *Language, economy, and society: The changing fortunes of the Welsh language in the twentieth century.* University of Wales Press, 2000.

G.A. Akerlof. The market for" lemons": Quality uncertainty and the market mechanism. *The quarterly journal of economics*, pages 488–500, 1970.

Kurt W Alt, Marion Benz, Wolfgang Müller, Margit E Berner, Michael Schultz, Tyede H Schmidt-Schultz, Corina Knipper, Hans-Georg K Gebel, Hans J Nissen, and Werner Vach. Earliest evidence for social endogamy in the 9,000-year-old-population of basta, jordan. *Plos One*, 8(6):e65649, 2013.

James E Anderson and Eric Van Wincoop. Trade costs. Technical report, National Bureau of Economic Research, 2004.

Cemal Eren Arbatlı, Quamrul H Ashraf, Oded Galor, and Marc Klemp. Diversity and conflict. *Econometrica*, 88(2):727–797, 2020.

Quamrul Ashraf and Oded Galor. Dynamics and stagnation in the malthusian epoch. *American Economic Review*, 101(5):2003–41, 2011.

Peter K. Austin and Julia Sallabank. *The Cambridge handbook of endangered languages.* Cambridge University Press, Cambridge, 2011.

Ying Bai and James Kai-sing Kung. Climate shocks and sino-nomadic conflict. *Review of Economics and Statistics*, 93(3):970–981, 2011.

Matthew J Baker. An equilibrium conflict model of land tenure in hunter-gatherer societies. *Journal of Political Economy*, 111(1):124–173, 2003.

Robin L Bennett, Arno G Motulsky, Alan Bittles, Louanne Hudgins, Stefanie Uhrich, Debra Lochner Doyle, Kerry Silvey, C Ronald Scott, Edith Cheng, Barbara McGillivray, et al. Genetic counseling and screening of consanguineous couples and their offspring: recommendations of the national society of genetic counselors. *Journal of genetic counseling*, 11(2):97–119, 2002.

Lee Benshoof and Randy Thornhill. The evolution of monogamy and concealed ovulation in humans. *Journal of Social and Biological Structures*, 2(2):95–106, 1979.

A. Bentahila and E.E. Davies. Language revival: restoration or transformation? *Journal of Multilingual & Multicultural Development*, 14(5):355–374, 1993.

Sourav Bhattacharya, Joyee Deb, and Tapas Kundu. Mobility and conflict. *American Economic Journal: Microeconomics*, 7(1):281–319, 2015.

Andreas Blume and Oliver Board. Language barriers. *Econometrica*, 81(2):781–812, 2013a. ISSN 1468-0262. doi: 10.3982/ECTA9183. URL `http://dx.doi.org/10.3982/ECTA9183`.

Andreas Blume and Oliver Board. Intentional vagueness. *Erkenntnis*, pages 1–45, 2013b. ISSN 0165-0106. doi: 10.1007/s10670-013-9468-x. URL `http://dx.doi.org/10.1007/s10670-013-9468-x`.

Andreas Blume, Yong-Gwan Kim, and Joel Sobel. Evolutionary stability in games of communication. *Games and Economic Behavior*, 5:547–575, 1993.

Samuel Bowles. Group competition, reproductive leveling, and the evolution of human altruism. *science*, 314(5805):1569–1572, 2006.

Samuel Bowles. Did warfare among ancestral hunter-gatherers affect the evolution of human social behaviors? *Science*, 324(5932):1293–1298, 2009.

Samuel Bowles. Warriors, levelers, and the role of conflict in human social evolution. *Science*, 336(6083):876–879, 2012.

Robert Boyd and Peter J. Richerson. *Culture and the Evolutionary Process.* University of Chicago Press, Chicago, 1988.

H. Buhaug and P. Lujala. Accounting for scale: Measuring geography in quantitative studies of civil war. *Political Geography*, 24(4):399–418, 2005.

H. Buhaug and J.K. Rød. Local determinants of african civil wars, 1970–2001. *Political Geography*, 25(3):315–335, 2006.

H. Buhaug, S. Gates, and P. Lujala. Geography, rebel capability, and the duration of civil conflict. *Journal of Conflict Resolution*, 53(4):544–569, 2009.

Marshall Burke, Solomon M Hsiang, and Edward Miguel. Climate and conflict. 2015.

D. Charlesworth and B. Charlesworth. Inbreeding depression and its evolutionary consequences. *Annual review of ecology and systematics*, 18:237–268, 1987.

Deborah Charlesworth and John H Willis. The genetics of inbreeding depression. *Nature reviews genetics*, 10(11):783–796, 2009.

Sylvain Chassang, Gerard Padró i Miquel, et al. Economic shocks and civil war. *Quarterly Journal of Political Science*, 4(3):211–28, 2009.

Haiyun Chen. *Essays on the Economics of Linguistic Diversity and Preference for Surprise.* PhD thesis, Simon Fraser University, Burnaby, B.C., Canada, 2017.

Jung-Kyoo Choi and Samuel Bowles. The coevolution of parochial altruism and war. *science*, 318(5850):636–640, 2007.

Julie Coleman. *A History of Cant and Slang Dictionaries: Volume III: 1859-1936: Volume III: 1859-1936*, volume 3. OUP Oxford, 2008.

Alison d'Anglejan. Language planning in quebec: An historical overview and future trends. *Conflict and Language Planning in Quebec*, 5:29, 1984.

G.M. Dawson. *Notes on the Shuswap people of British Columbia.* Dawson Brothers, 1891.

Karl R. de Rouen and D. Sobek. The dynamics of civil war duration and outcome. *Journal of Peace Research*, 41(3):303–320, 2004.

I. De Soysa. Paradise is a bazaar? Greed, creed, and governance in civil war, 1989-99. *Journal of Peace Research*, 39(4):395–416, 2002.

Klaus Desmet, Ignacio Ortuño-Ortín, and Romain Wacziarg. The political economy of linguistic cleavages. *Journal of Development Economics*, 97(2):322–338, 2012.

G.K. Dow, C.G. Reed, and N. Olewiler. Climate reversals and the transition to agriculture. *Journal of Economic Growth*, 14(1):27–53, 2009.

Gregory K Dow and Clyde G Reed. The origins of inequality: Insiders, outsiders, elites, and commoners. *Journal of Political Economy*, 121(3):609–641, 2013.

Gregory K Dow and Clyde G Reed. The origins of sedentism: Climate, population, and technology. *Journal of Economic Behavior & Organization*, 119:56–71, 2015.

Gregory K Dow, Clyde G Reed, and Simon Woodcock. The economics of exogamous marriage in small-scale societies. *Economic Inquiry*, 54(4):1805–1823, 2016.

Gregory K Dow, Leanna Mitchell, and Clyde G Reed. The economics of early warfare over land. *Journal of Development Economics*, 127:297–305, 2017.

Carol R Ember and Melvin Ember. Resource unpredictability, mistrust, and war: A cross-cultural study. *Journal of Conflict Resolution*, 36(2):242–262, 1992.

Joan Esteban, Laura Mayoral, and Debraj Ray. Ethnicity and conflict: An empirical study. *American Economic Review*, 102(4):1310–42, September 2012. doi: 10.1257/aer.102.4.1310.

J. Fairhead. The conflict over natural and environmental resources. *War, hunger, and displacement. The origins of humanitarian emergencies*, 1:147–178, 2000.

James D Fearon et al. Rationalist explanations for war. *International organization*, 49:379–379, 1995.

J.D. Fearon and D.D. Laitin. Ethnicity, insurgency, and civil war. *American Political Science Review*, 97(1):75–90, 2003.

R Brian Ferguson. The prehistory of war and peace in europe and the near east. *War, peace, and human nature: the convergence of evolutionary and cultural views*, pages 191–240, 2013.

Edward Finegan. *Language: Its Stucture and Use.* Thomson Wadsworth, 5 edition, 2008.

Douglas P Fry and Patrik Söderberg. Lethal aggression in mobile forager bands and implications for the origins of war. *Science*, 341(6143):270–273, 2013.

B. Galloway. Language revival programs of the Nooksack Tribe and the Stó:lō Nation. *Be of Good Mind: Essays on the Coast Salish*, page 212, 2007.

N. Gardner, M. Serralvo, and C Williams. Language revitalization in comparative context: Ireland, the Basque Country and Catalonia. In C Williams, editor, *Language Revitalization: Policy and Planning in Wales*, pages 311–355. University of Wales Press, Cardiff, 2000.

Michelle R. Garfinkel and Stergios Skaperdas. Economics of conflict: An overview. In Todd Sandler and Keith Hartley, editors, *Handbook of Defense Economics*, volume 2, pages 649–709. Elsevier, 2007.

Arthur Goldschmidt and Lawrence Davidson. *A concise history of the Middle East*. Westview Press, 1991.

Russell D Gray, Alexei J Drummond, and Simon J Greenhill. Language phylogenies reveal expansion pulses and pauses in pacific settlement. *Science*, 323(5913):479–483, 2009.

Michael Gurven, Hillard Kaplan, and Maguin Gutierrez. How long does it take to become a proficient hunter? implications for the evolution of extended development and long life span. *Journal of human evolution*, 51(5):454–470, 2006.

Ken Hale, Michael Krauss, Lucille J. Watahomigie, Akira Y. Yamamoto, Colette Craig, LaVerne Masayesva Jeanne, and Nora C. England. Endangered languages. *Language*, 68 (1):1–42, 1992. ISSN 00978507. URL `http://www.jstor.org/stable/416368`.

Robert Anderson Hall. *Pidgin and creole languages*, volume 7. Cornell University Press, Ithaca, 1966.

David Harmon. Losing species, losing languages: Connections between biological and linguistic diversity. *Southwest Journal of Linguistics*, 15:89–108, 1996.

Marvin Harris. I977. cannibals and kings: The origins of cultures. *Vintage, New York*.

Brian Hayden. Feasting in prehistoric and traditional societies. *Food and the status quest: an interdisciplinary perspective*, pages 127–147, 1996.

Agnar Helgason, Snæbjörn Pálsson, Daníel F Guðbjartsson, Kári Stefánsson, et al. An association between the kinship and fertility of human couples. *Science*, 319(5864):813–816, 2008.

John C Hotten. *A Dictionary of Modern Slang, Cant, and Vulgar Words*. Hotten, 1859.

T. Jang, K. Lim, and T. Payne. Xibe language revitalization and documentation project. 2011.

A. Johnson and T. Earle. *The evolution of human societies: from foraging group to agrarian state*. Stanford University Press, 2000.

David Kaplan. Conflict and compromise among borderland identities in northern italy. *Tijdschrift voor economische en sociale geografie*, 91(1):44–60, 2000.

Hillard Kaplan, Kim Hill, Rowe V Cadelina, Brian Hayden, David Charles Hyndman, Richard J Preston, Eric Alden Smith, David E Stuart, and David R Yesner. Food sharing among ache foragers: Tests of explanatory hypotheses [and comments and reply]. *Current anthropology*, 26(2):223–246, 1985.

Hillard S Kaplan, Eric Schniter, Vernon L Smith, and Bart J Wilson. Risk and the evolution of human exchange. *Proceedings of the Royal Society B: Biological Sciences*, 279(1740): 2930–2935, 2012.

Lawrence H Keeley. *War before civilization*. OUP USA, 1996.

Raymond Case Kelly. *Warless societies and the origin of war*. University of Michigan Press, 2000.

Robert L Kelly. *The lifeways of hunter-gatherers: the foraging spectrum*. Cambridge University Press, 2013a.

Robert L Kelly. From the peaceful to the warlike: ethnographic and archaeological insights into hunter-gatherer warfare and homicide. *War, peace, and human nature: The convergence of evolutionary and cultural views*, pages 151–167, 2013b.

Devra G Kleiman. Monogamy in mammals. *The Quarterly review of biology*, 52(1):39–69, 1977.

F. Kratochwil. Of systems, boundaries, and territoriality: An inquiry into the formation of the state system. *World Politics*, 39(1):27–52, 1986.

Joseph B. Kruskal, Isidor Dyen, and Paul Black. The vocabulary method of reconstructing language trees: Innovations and large-scale applications. In F. R. Hodson, D. G. Kendall, and P. Tautu, editors, *Mathematics in the Archeological and Historical Sciences*, pages 361–380. Edinburgh University Press, Edinburgh, 1971.

D Kulick. Language shift and cultural reproduction: socialization, self, and syncretism in a papua new guinean village. 1992.

Shrikant Kuntla, Srinivas Goli, TV Sekher, and Riddhi Doshi. Consanguineous marriages and their effects on pregnancy outcomes in india. *International Journal of Sociology and Social Policy*, 2013.

R.C. Lacy, A. Petric, and M. Warneke. Inbreeding and outbreeding in captive populations of wild animal species. *The natural history of inbreeding and outbreeding*, pages 352–374, 1993.

Edward P. Lazear. Culture and language. *Journal of Political Economy*, 107(S6):pp. S95–S126, 1999. ISSN 00223808. URL `http://www.jstor.org/stable/10.1086/250105`.

P. Le Billon. Angola's political economy of war: The role of oil and diamonds, 1975–2000. *African Affairs*, 100(398):55–80, 2001.

Paul L Leberg and Brigette D Firmin. Role of inbreeding depression and purging in captive breeding and restoration programmes. *Molecular Ecology*, 17(1):334–343, 2008.

Richard B Lee, Richard Heywood Daly, Richard Daly, et al. *The Cambridge encyclopedia of hunters and gatherers*. Cambridge University Press, 1999.

M. Paul Lewis, editor. *Ethnologue: Languages of the World*. SIL International, Dallas, Texas, 16th edition, 2009. URL `http://www.ethnologue.org/web.asp`.

M. Paul Lewis, Gary F. Simons, and Charles D. Fennig, editors. *Ethnologue: Languages of the World*. SIL International, Dallas, Texas, 17th edition, 2013. URL `http://www.ethnologue.com`.

Barton L. Lipman. Language and economics. In Nicola Dimitri, Marcello Basili, and Itzhak Gilboa, editors, *Cognitive Processes and Economic Behavior*, chapter 7, pages 75–93. Routledge, London, 2003.

Ruth Mace and Mark Pagel. A latitudinal gradient in the density of human languages in north america. *Proceedings: Biological Sciences*, 261(1360):117–121, 1995. URL `http://www.jstor.org/stable/50055`.

]barrus2004thomas Thomas Robert Malthus. Thomas malthus: An essay on the principle of population. In Philip Appleman, editor, *An Essay on the Principle of Population*. Norton, 2nd edition, 1798[2004].

Ruth McQuillan, Niina Eklund, Nicola Pirastu, Maris Kuningas, Brian P McEvoy, Tõnu Esko, Tanguy Corre, Gail Davies, Marika Kaakinen, Leo-Pekka Lyytikäinen, et al. Evidence of inbreeding depression on human height. *PLoS Genetics*, 8(7):e1002655, 2012.

Stelios Michalopoulos. The origins of ethnolinguistic diversity. *American Economic Review*, 102(4):1508–39, 2012. doi: 10.1257/aer.102.4.1508. URL `http://www.aeaweb.org/articles.php?doi=10.1257/aer.102.4.1508`.

J. Milloy. When a language dies. *Index on Censorship*, 28(4):54–64, 1999.

Monique Borgerhoff Mulder, Samuel Bowles, Tom Hertz, Adrian Bell, Jan Beise, Greg Clark, Ila Fazzio, Michael Gurven, Kim Hill, Paul L Hooper, et al. Intergenerational wealth transmission and the dynamics of inequality in small-scale societies. *science*, 326(5953): 682–688, 2009.

George P Murdock and Douglas R White. Standard cross-cultural sample. *Ethnology*, 8(4): 329–369, 1969.

George P Murdock and Douglas R White. Standard cross-cultural sample: On-line edition. 2006.

Patricia Murrieta-Flores. Travelling in a prehistoric landscape: Exploring the influences that shaped human movement. In *Making History Interactive. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 37th International Conference (Williamsburg, Virginia, United States of America 2009). British Archaeological Reports, International Series*, volume 2079, pages 258–276, 2010.

Daniel Nettle. Explaining global patterns of language diversity. *Journal of Anthropological Archaeology*, 17(4):354–374, 1998. doi: 10.1006/jaar.1998.0328. URL `http://www.sciencedirect.com/science/article/pii/S0278416598903282`.

Douglass C North, John Joseph Wallis, and Barry R Weingast. Violence and the rise of open-access orders. *Journal of democracy*, 20(1):55–68, 2009.

Eghosa E Osaghae and Rotimi T Suberu. *A History of Identities, Violence and Stability in Nigeria*. Centre for Research on Inequality, Human Security and Ethnicity, University of Oxford, 2005.

Mark Pagel. The history, rate and pattern of world linguistic evolution. In Chris Knight, and Miahael Studdert-Kennedy, and James R. Hurford, editor, *The Evolutionary Emergence of Language*, pages 391–416. Cambridge University Press, Cambridge, 2000.

Mark Pagel. War of words: the language paradox explained. *New Scientist*, 216(2894):38–41, 2012.

Mark Pagel, Quentin D Atkinson, and Andrew Meade. Frequency of word-use predicts rates of lexical evolution throughout indo-european history. *Nature*, 449(7163):717–720, 2007.

Eric Partridge, Tom Dalzell, and Terry Victor. *The Concise New Partridge Dictionary of Slang and Unconventional English*. Psychology Press, 2008.

Robert Powell. Stability and the distribution of power. *World Pol.*, 48:239, 1995.

Robert Powell. War as a commitment problem. *International organization*, pages 169–203, 2006.

Anne Pusey and Marisa Wolf. Inbreeding avoidance in animals. *Trends in Ecology & Evolution*, 11(5):201–206, 1996.

Katherine Ralls, Jonathan D Ballou, and Alan Templeton. Estimates of lethal equivalents and the cost of inbreeding in mammals. *Conservation biology*, 2(2):185–193, 1988.

M. Renner. The anatomy of resource wars. *World Watch Paper*, 162, 2002.

P. Rickard. *A history of the French language*. Routledge, 1989.

Arthur J Robson. Efficiency in evolutionary games: Darwin, nash and the secret handshake. *Journal of Theoretical Biology*, 144(3):379–396, 1990.

Arthur J Robson. The biological basis of economic behavior. *Journal of economic literature*, 39(1):11–33, 2001.

Arthur J Robson. Evolution and human nature. *Journal of Economic Perspectives*, 16(2): 89–106, 2002.

Elisa Roller. When does language become exclusivist? linguistic politics in catalonia. *National Identities*, 4(3):273–289, 2002.

Mark R Rosenzweig. Risk, private information, and the family. *The American Economic Review*, 78(2):245–250, 1988a.

Mark R Rosenzweig. Risk, implicit contracts and the family in rural areas of low-income countries. *The Economic Journal*, 98(393):1148–1170, 1988b.

Mark R Rosenzweig and Oded Stark. Consumption smoothing, migration, and marriage: Evidence from rural india. *Journal of political Economy*, 97(4):905–926, 1989.

Robert Rowthorn and Paul Seabright. Property rights, warfare and the neolithic transition. 2010.

Ariel Rubinstein. *Economics and Language: Five Essays.* Cambridge University Press, Cambridge, 2000.

Mark A Schneider and Lewellyn Hendrix. Olfactory sexual inhibition and the westermarck effect. *Human Nature*, 11(1):65–91, 2000.

John Andrew Simpson, Edmund SC Weiner, et al. *The Oxford English Dictionary*, volume 2. Clarendon Press Oxford, 1989.

I. Smillie and F. Forskningsstiftelsen. *Dirty Diamonds: Armed Conflict and the Trade in Rough Diamonds.* Fafo Institute for Applied Social Science, 2002.

Camilla Stoltenberg, Per Magnus, Anders Skrondal, and Rolv Terje Lie. Consanguinity and recurrence risk of stillbirth and infant death. *American Journal of Public Health*, 89(4): 517–523, 1999.

Rotimi T Suberu. *Federalism and ethnic conflict in Nigeria.* United States Inst of Peace Press, 2001.

Miikka Tallavaara, Miska Luoto, Natalia Korhonen, Heikki Järvinen, and Heikki Seppä. Human population dynamics in europe over the last glacial maximum. *Proceedings of the National Academy of Sciences*, 112(27):8232–8237, 2015.

Richard SJ Tol and Sebastian Wagner. Climate change and violent conflict in europe over the last millennium. *Climatic Change*, 99(1):65–79, 2010.

Pierre L Van den Berghe. Human inbreeding avoidance: Culture in nature. *Behavioral and Brain Sciences*, 6(1):91–102, 1983.

Dirk Van Vuren and Lynne A Isbell. Differential costs of locational and social dispersal and their consequences for female group-living primates. *Behaviour*, 133(1-2):1–36, 1996.

Robert Walker, Kim Hill, Hillard Kaplan, and Garnett McMillan. Age-dependency in hunting ability among the ache of eastern paraguay. *Journal of Human Evolution*, 42(6):639–657, 2002.

Karl Wärneryd. Cheap talk, coordination, and evolutionary stability. *Games and Economic Behavior*, 5:532–546, 1993.

Glenn E Weisfeld, Tiffany Czilli, Krista A Phillips, James A Gall, and Cary M Lichtman. Possible olfaction-based mechanisms in human kin recognition and inbreeding avoidance. *Journal of experimental child psychology*, 85(3):279–295, 2003.

Edward Westermarck. The history of human marriage, 3 vols. 1921.

Arthur P Wolf. *Sexual attraction and childhood association: A Chinese brief for Edward Westermarck*. Stanford University Press, 1995.

S.M. Younger. Conditions and mechanisms for peace in precontact polynesia. *Current Anthropology*, 49(5):927–934, 2008.

David D Zhang, Chi Yung Jim, George CS Lin, Yuan-Qing He, James J Wang, and Harry F Lee. Climatic change, wars and dynastic cycles in china over the last millennium. *Climatic Change*, 76(3):459–477, 2006.

David D Zhang, Jane Zhang, Harry F Lee, and Yuan-qing He. Climate change and war frequency in eastern china over the last millennium. *Human Ecology*, 35(4):403–414, 2007.

Appendix A

Chapter 1 Appendix

A.1 The long-run distribution of fitness depression

This appendix proves that $\lim_{\epsilon\to 0}\lim_{t\to\infty} P\left[\delta_{it}=0\right]\to 0$. The order of the limits is important; without letting the population reach a steady state for a given ϵ there is no guarantee that the limit is meaningful. For example $\lim_{\epsilon\to 0} P\left[\delta_{i0}=0\right]$ does not converge and is simply $P\left[\delta_{i0}=0\right]$.

Lemma A.1. *$\forall\ \lambda$ s.t. $F\left(\bar{\theta}\left(\lambda\right)\right)\in[0,1)$, the distribution of fitness depression converges to a unique steady state regardless of initial conditions: $\lim_{t\to\infty}\mathbf{p_t}\left(\lambda,\epsilon\right)=\mathbf{p}\left(\lambda,\epsilon\right)$.*

For $F\left(\bar{\theta}\left(\lambda\right)\right)=0$, the Lemma follows directly from the same line of argument as Proposition 1.2. The population has no fitness depression at all after $D+1$ periods, and so $\mathbf{p_t}\left(\lambda,\epsilon\right)$ converges to a steady state, degenerate, distribution.

For $F\left(\bar{\theta}\left(\lambda\right)\right)\in(0,1)$, recall that $\epsilon=D^{-1}$. Let $p_{j,t}\left(\lambda,\epsilon\right)=P\left[\delta_{it}=Dj|\lambda,\epsilon\right]$. We can then define $\mathbf{p_t}\left(\lambda,\epsilon\right)=\left[p_{0,t}\left(\lambda,\epsilon\right),p_{1,t}\left(\lambda,\epsilon\right),\ldots,p_{D,t}\left(\lambda,\epsilon\right)\right]$. I proceed by defining the transition matrix for fitness depression, $M\left(\lambda,\epsilon\right)$.

I begin with the transition dynamics of fitness depression at the population level, outside the steady state. $\forall j\in\{1,...D\}$:

$$p_{j,t}=\begin{array}{l}\frac{2(1-(j-1)\epsilon)\left(E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\right)}{F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)+2(1-(j-1)\epsilon)\left(E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\right)}p_{j-1,t-1}+\\ \frac{F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)}{F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)+2(1-(j+1)\epsilon)\left(E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\right)}p_{j+1,t-1}\end{array}\tag{A.1}$$

That is, the fraction of groups with fitness depression level $j\epsilon$ at time t is equal to the sum of a) the fraction of groups with fitness depression level $(j-1)\,\epsilon$ at time t-1 who marry endogamously (after accounting for fitness depression), multiplied by the proportion of groups with fitness depression level $(j-1)\,\epsilon$ at time $t-1$and b) the fraction of groups with fitness depression level $(j+1)\,\epsilon$ at time t-1 who marry exogamously, multiplied by the proportion groups with fitness depression level $(j+1)\,\epsilon$ at time $t-1$. For $j=0$, the formula is slightly different:

$$p_{0,t}=\begin{array}{l}\frac{2\epsilon\left(E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\right)}{F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)+2\epsilon\left(E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\right)}p_{0,t-1}+\\ \frac{F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)}{F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)+2(1-\epsilon)\left(E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\right)}p_{1,t-1}\end{array}\tag{A.2}$$

Together, equations (A.1) and (A.2) allow me to define the a transition matrix for fitness depression. Let

$$A(\lambda,\epsilon,j)=\frac{F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)}{F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)+2\left(1-(j-1)\epsilon\right)\left(E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\right)}$$

$$B(\lambda,\epsilon,j)=\frac{2\left(1-(j-1)\epsilon\right)\left(E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\right)}{F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)+2\left(1-(j-1)\epsilon\right)\left(E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\right)}$$

$A(\lambda,\epsilon,j)$is the probability of moving down one level of fitness depression (less depressed), while $B(\lambda,\epsilon,j)$ is the probability of a group moving up one level of fitness depression.

For completeness, $M(\lambda,\epsilon)$ is the $D+1\times D+1$ matrix:

$$M(\lambda,\epsilon)=$$

$$\begin{bmatrix} A(\lambda,\epsilon,0) & B(\lambda,\epsilon,0) & 0 & \cdots & 0 & 0 & 0 & \cdots \\ A(\lambda,\epsilon,1) & 0 & B(\lambda,\epsilon,1) & \cdots & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots \\ 0 & 0 & 0 & \cdots & A(\lambda,\epsilon,j) & 0 & B(\lambda,\epsilon,j) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Provided that $F\left(\bar{\theta}(\lambda,\epsilon)\right)\in(0,1)$, the probability of a group moving down one level of fitness depression, $A(\lambda,\epsilon,j)$, is strictly positive, which implies that the probability of moving from fitness depression level j to $j-k$ in k periods is strictly positive $\forall j\in\{0..D\},k\in\{1,...,j\}$. Because the probability of a group moving down one level, $A(\lambda,\epsilon,j)$, is also strictly positive, the same argument applies to moves from fitness depression level j to $j-k$ in k periods for $k\in\{1,...,D-j\}$. Finally, the probability of moving up one level and then down one level within two periods is strictly positive, so the probability of returning to fitness depression level j in 2 periods is also strictly positive. It may appear that groups can only return to a fitness depression level j in k periods if k is even, but because groups can get "stuck" at $\delta_{it}=0$ they can also return if k is odd provided $k>2j+1$. Together, these results imply that for some $k>2D+1$, $(M(\lambda,\epsilon))^k$ has only positive entries, and so $M(\lambda,\epsilon)$ is a regular stochastic matrix. This in turn implies that $\mathbf{p}(\lambda,\epsilon)$ converges to a unique steady state, $\mathbf{p}(\lambda,\epsilon)$.

Lemma A.2. $\lim_{\epsilon\to 0}\lim_{t\to\infty}P[\delta_{it}=0]\to 0$

In what follows, $\bar{\theta}$is used as shorthand for $\bar{\theta}(\lambda)$.The rest of the proof divides into two cases depending whether or not $F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\geq\frac{1}{2}E[\theta]$. Intuitively, this condition describes whether the mass of groups becoming more inbred is greater than the mass become less inbred, before accounting for losses due to fitness depression. Towards a contradiction, suppose first that the inequality holds weakly. Rearranging, I get:

$$\frac{2\left(E\left[\theta\right]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF\left(\theta\right)\right)}{F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF\left(\theta\right)}\leq 1$$

I now turn to the steady state population wide level of fitness depression:

$$\delta\left(\lambda,\epsilon\right)=E\left[\delta_i|\lambda,\epsilon\right]$$

$$\delta\left(\lambda,\epsilon\right)=\sum_{j=0}^{D}p_j\epsilon j$$

where p_j is the $j+1$th element of $\mathbf{p}\left(\lambda,\epsilon\right)$. In the steady state, the fraction of groups moving from state i to state $i+1$ must be equal to the fraction moving from i to $i-1$.

$$\frac{F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)}{2(1-j\epsilon)\left(E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\right)+F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)}p_j=$$
$$\frac{2(1-(j-1)\epsilon)\left(E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\right)}{2(1-(j-1)\epsilon)\left(E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\right)+F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)}p_{j-1}$$

Rearranging:

$$p_j=\begin{array}{c}\dfrac{2(1-j\epsilon)\left(E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\right)+F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)}{2(1-(j-1)\epsilon)\left(E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\right)+F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)}\times\\ \dfrac{2(1-(j-1)\epsilon)\left(E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\right)}{F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)}p_{j-1}\end{array}$$

Let $\psi=\max_{j\in\{1\ldots D\}}\frac{2(1-j\epsilon)\left(E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\right)+F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)}{2(1-(j-1)\epsilon)\left(E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\right)+F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)}\frac{2(1-(j-1)\epsilon)\left(E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\right)}{F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)}$. Because the two fractions are strictly less than 1 for all j, ψ must also be strictly less than 1. [1]

Therefore,.$\forall j>0$, $p_j<\psi p_{j-1}$.We can now turn to $\lim_{D\to\infty}\delta\left(\lambda,D^{-1}\right)$, where $\delta\left(\lambda,D^{-1}\right)=\lim_{t\to\infty}\mathrm{E}\left[\delta_{it}|\lambda,D^{-1}\right]$

$$\lim_{D\to\infty}\delta\left(\lambda,D^{-1}\right)=\lim_{D\to\infty}\sum_{j=0}^{D}p_j j\epsilon=\lim_{D\to\infty}\sum_{j=0}^{D}p_j\frac{j}{D}$$

---

[1] This is maximized at the smallest possible value of j, 1

$$< \lim_{D\to\infty} \sum_{j=0}^{D} \frac{j}{D} \psi^j p_0$$

$$= p_0 \lim_{D\to\infty} \sum_{j=0}^{D} j\psi^j \lim_{D\to\infty} \frac{1}{D}$$

$$= p_0 \left(\frac{\psi}{(\psi-1)^2} \right) \lim_{D\to\infty} \frac{1}{D}$$

$$= p_0 \left(\frac{\psi}{(\psi-1)^2} \right) \times 0 = 0$$

Because fitness depression cannot be negative, it must be the case that:

$$\lim_{D\to\infty} \delta\left(\lambda, D^{-1}\right) = 0$$

But if population level fitness depression is zero, even if all other groups were marrying exogamously, a group would only want to marry exogamously if $\bar{\theta} < \frac{1}{2}E[\theta]$ (See section X). This result, assumption (X) (that the c.d.f. of F is concave), and the requirement for a PBNE, implies that $F\left(\bar{\theta}\right)\int_0^{\bar{\theta}} \theta dF(\theta) < \frac{1}{2}E[\theta]$, which contradicts our original assumption that $F\left(\bar{\theta}\right)\int_0^{\bar{\theta}} \theta dF(\theta) > \frac{1}{2}E[\theta]$.

The only alternative is that $F\left(\bar{\theta}\right)\int_0^{\bar{\theta}} \theta dF(\theta) < \frac{1}{2}E[\theta]$. Towards a contradiction, assume $\lim_{D\to\infty} p_0\left(\lambda, D^{-1}\right) > 0$.

Choose $k = \left\lceil (p_0)^{-1} \right\rceil$ [2]

We want to show that $\lim_{D\to\infty} p_0\left(\lambda, D^{-1}\right) > 0$ implies that $\lim_{D\to\infty} \sum_{j=0}^{k} p_j \to K > 1$, which violates the definition of $\mathbf{p}(\lambda, \epsilon)$ as a vector of probabilities. Recall that:

$$p_j = \begin{array}{c} \frac{2(1-j\epsilon)\left(E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\right)+F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)}{2(1-(j-1)\epsilon)\left(E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\right)+F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)} \times \\ \frac{2(1-(j-1)\epsilon)\left(E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\right)}{F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)} p_{j-1} \end{array}$$

Taking the limit with fixed j:

---

[2] This implies that $D \geq \left\lceil (p_0)^{-1} \right\rceil$, but because I take the limit as $D \to \infty$, this implication is not restrictive.

$$\lim_{D\to\infty} p_j = \lim_{\epsilon\to 0} \begin{array}{c} \frac{2(1-j\epsilon)\left(E[\theta]-x\left(\bar{\theta}\right)\right)+F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)}{2(1-(j-1)\epsilon)\left(E[\theta]-x\left(\bar{\theta}\right)\right)+F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)}\times \\ \frac{2(1-(j-1)\epsilon)\left(E[\theta]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)\right)}{F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF(\theta)} p_{j-1} \end{array}$$

$$\lim_{D\to\infty} p_j = \frac{2\left(E\left[\theta\right]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF\left(\theta\right)\right)}{F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF\left(\theta\right)} p_{j-1}$$

But the fraction must be greater than zero by the assumption made above. So

$$\lim_{D\to\infty} p_j > p_{j-1}$$

And this is true $\forall\ k \leq j$. Therefore, by induction:

$$\lim_{D\to\infty} p_j = \frac{2\left(E\left[\theta\right]-F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF\left(\theta\right)\right)}{F\left(\bar{\theta}\right)\int_0^{\bar{\theta}}\theta dF\left(\theta\right)} p_{j-1} > p_0$$

Summing the first k probabilities, we get:

$$\lim_{D\to\infty} \sum_{j=0}^{k} p_j > (k+1)\, p_0$$

$$\lim_{D\to\infty} \sum_{j=0}^{k} p_j > \left(\left\lceil (p_0)^{-1} \right\rceil + 1\right) p_0 > 1$$

Therefore $\lim_{D\to\infty} p_0 = 0$. This completes the proof.

Appendix B

Chapter 2 Appendix

B.1 Justification for the Reduced Form Expected Payoffs

In this section we provide a justification for the two reduced form expected payoffs, (Equation 2.2) and (Equation 2.3), described in subsection 2.3.3. In particular, we show how communication in games of cooperative and competitive natures can lead to the specific forms of these payoff functions.

Our argument is developed in a framework of two-player extensive form games, whose payoff structure is of either cooperative or competitive nature. The game begins with nature choosing a state and one randomly selected player to observe it. This player, referred to as the sender, then decides whether to transmit a message regarding the state. If a message is sent, then the other group, called the receiver, has an opportunity to choose an action based on the sender's message.

Each group is endowed with a set of messages that refer to a subset of the states. Elements in the intersection of the two spaces are understood by both groups, while those unique to one group's message space are understood only by that group. We derive a group's expected payoffs in the subgame perfect equilibria, and show that they are related to the sizes of intersection and relative complement of the two message spaces in the way specified by (Equation 2.2) and (Equation 2.3).[1]

B.1.1 General Environment

Suppose at the beginning of each period t, a two dimensional state (τ, a) is drawn to determine a game $\Gamma^t(\tau, a)$ played by groups i and j. The parameter $\tau \in \{\text{coop}, \text{comp}\}$ determines the strategic nature of the interaction, and $a \in A = [0, 1]$ determines the optimal action(s) for each group. The two parameters are drawn independently of each other: $\tau = \text{coop}$ with probability $\frac{p}{p+q}$ and $\tau = \text{comp}$ with probability $\frac{q}{p+q}$; a is drawn uniformly from A. Let σ_k, $k = i, j$, denote group k's action in $\Gamma^t(\tau, a)$. The payoffs in $\Gamma^t(\text{coop}, a)$ are

$$\pi_i(\sigma_i, \sigma_j) = \pi_j(\sigma_i, \sigma_j) = \begin{cases} 1 & \text{if } \sigma_i = \sigma_j = a \\ 0 & \text{otherwise,} \end{cases} \qquad \forall a \in A; \tag{B.1}$$

---

[1] While we interpret a group's message space to be its dialect, the framework proposed in this section does not integrate perfectly with the model presented in the paper. Specifically, in this section, the measure of the intersection of the message spaces is assumed to be smaller than the measure of the action space. For this to hold every period in the main sections of the paper, the action space would have to grow exogenously every period at a rate no smaller than the growth rate of the intersections.

and the payoffs in $\Gamma^t(\text{comp}, a)$ are

$$\pi_i(\sigma_i, \sigma_j) = \begin{cases} 2 & \text{if } \sigma_i = a \neq \sigma_j \\ -2 & \text{if } \sigma_i = \sigma_j = a \end{cases} \quad \text{and} \quad \pi_j(\sigma_i, \sigma_j) = -\pi_i(\sigma_i, \sigma_j), \qquad \forall a \in A. \tag{B.2}$$

Let $M_k = [x_k, y_k]$, where $0 \leq x_k < y_k \leq 1$ and $k = i, j$, be the set of messages that group k can use to refer to actions in A. Note that $M_i \cap M_j$ need not be empty, and that M_i and M_j need not have the same size. Each message $m \in M_k$ refers to one and only one action $a \in A$. More precisely, let $\mu_k : M_k \to A$ be a measure-preserving injection that maps elements in M_k to those in A. Thus, the meaning (or referent) of a message $m \in M_k$ is given by $\mu_k(m)$. Further, let μ_i, μ_j satisfy the property that $\mu_i(m) = \mu_j(m)$ for all $m \in M_i \cap M_j$, so that messages in the intersection of the two message spaces refer to the same actions. Denote by m_k^a a message sent by k that refers to action a; thus $m_k^a = \mu_k^{-1}(a)$. We say that group j *understands* a message m_i sent by group i if and only if $m_i \in M_i \cap M_j$.[2]

Let $\widetilde{\Gamma}^t$ be an extensive form game based on $\Gamma^t(\tau, a)$:

1. With equal probability, nature chooses either i or j to be the sender.
2. The sender, say i, observes the states (τ, a).
    - If the state a is such that there exists no $m_i \in M_i$ with $\mu_i(m_i) = a$, the game ends and both i and j get zero payoff.
    - If there exists $m_i \in M_i$ with $\mu_i(m_i) = a$, then i decides whether to engage the receiver, j, in $\Gamma^t(\tau, a)$.
        - If i chooses not to engage, then the game ends with both groups getting zero payoff.
        - If i chooses to engage, then it sends m_i^a and chooses $\sigma_i = a$ in $\Gamma^t(\tau, a)$. Both τ and m_i^a are then observable to j.
3. Based on τ and m_i^a, j chooses an action $\sigma_j \in A$.
4. The payoffs are determined by σ_i and σ_j according to (Equation B.1) and (Equation B.2).

The game $\widetilde{\Gamma}^t$ can be interpreted as follows. The leader of group i discovers an interaction opportunity characterized by (τ, a) that would involve another group j. Group i's optimal

---

[2] We assume that messages in $M_i \cap M_j$ is common knowledge.

payoff is a function of a. If group i does not have the requisite language (i.e. m_i^a) to communicate a among its members, then it cannot take advantage of this opportunity. Hence there is no change in either group's payoff. Suppose i has a message m_i^a that refers to a. The the leader of i can still decide whether to take advantage of the opportunity based on the anticipated response of j. If she decides not to, then there is no change in either group's payoff. If she decides to take the opportunity, then she sends m_i^a, which serves as a coordination signal for members of group i. Group j can also observes m_i^a with some probability, and for simplicity, we assume this probability to be 1. Based on whether j understands m_i^a and the observed nature of the interaction, j responds (optimally).

In the next two subsections, we focus on the case where the sender has a message for the drawn state a, and analyze a group's expected payoff in a subgame perfect equilibrium (SPE).

B.1.2 Cooperative Interaction

Suppose $\tau = \text{coop}$, and i is the sender. The following is an SPE of $\widetilde{\Gamma}^t$:

- i always engages, sends m_i^a and chooses $\sigma_i = a$;[3]
- j chooses $\sigma_j = a$ if $m_i^a \in M_i \cap M_j$, and chooses uniformly from A if $m_i^a \in M_i \setminus M_j$.

In this SPE, the payoff of both groups are positive whenever a is such that $m_i^a \in M_i \cap M_j$. This event occurs with probability $|M_i \cap M_j|$, since $|A| = 1$. Therefore, the expected payoff of i is

$$|M_i \cap M_j|.$$

In fact, it is easy to verify that all SPEs (except a subset of measure zero of them) have the same derived expected payoffs. If we interpret $M_i = L_i \cup E_{ij}$, then (Equation 2.2) is derived.

B.1.3 Competitive Interaction

Suppose $\tau = \text{comp}$, and i is the sender. The following is an SPE of $\widetilde{\Gamma}^t$:

- i engages, sends m_i^a, and chooses $\sigma_i = a$ if and only if a is such that $m_i^a \in M_i \setminus M_j$;
- j chooses $\sigma_j = a$ if $m_i^a \in M_i \cap M_j$, and chooses uniformly from A if $m_i^a \in M_i \setminus M_j$.

---

[3] Note that this is a weakly dominant strategy for i.

In this SPE, the sender guarantees a non-negative payoff. The payoff is positive whenever a is such that $m_i^a \in M_i \setminus M_j$. As a result, the *ex ante* (before the identity of the sender is determined) expected payoff of group i is

$$\frac{1}{2}\Big[\Pr\Big(\{a : m_i^a \in M_i \setminus M_j\}\Big)(2) + \Pr\Big(\{a : m_j^a \in M_j \setminus M_i\}\Big)(-2)\Big] \\ = |M_i \setminus M_j| - |M_j \setminus M_i|.$$

Again, all SPEs (except a subset of measure zero of them) have the same derived expected payoffs. If we interpret $M_k = L_k \cup E_k \cup N_k$, $k = i, j$, then (Equation 2.3) is derived.

B.2 Proofs

B.2.1 Proof of 2.2

Proof. Let $\{L_i^t\}_i$ be symmetric and localized. We want to show that $\{L_i^{t+1}\}_i$ is also symmetric and localized, namely, it satisfies conditions (Equation 2.14), (Equation 2.15), and (Equation 2.10).

It is obvious that $\{L_i^{t+1}\}_i$ satisfies condition (Equation 2.14) by 2.1 and Corollary 2.

Next, to see that $\{L_i^{t+1}\}_i$ satisfies condition (Equation 2.15), consider i, j, k such that $d_{ij} = d_{ik}$. If $d_{ij} = d_{ik} \le d^{t*}$, where d^{t*} is determined by (Equation 2.13) for every $t \ge 1$, then we know by Corollary 1 that $L_i^t \cup E_i^t = L_j^t \cup E_j^t = L_k^t \cup E_k^t$. From the definition of L^{t+1}, it follows that $|L_i^{t+1} \cap L_j^{t+1}| = |L_i^{t+1} \cap L_k^{t+1}|$. Suppose $d_{ij} = d_{ik} > d^{t*}$. There are two cases to consider. First, if $d_{ij} = d_{ik} > d^{t*} + 1$, then i will learn nothing from either j or k (nor would the latter two learn anything from i). So $|L_i^{t+1} \cap L_j^{t+1}| = |L_i^{t+1} \cap L_k^{t+1}| = |L_i^t \cap L_j^t| = |L_i^t \cap L_k^t|$. Second, suppose $d_{ij} = d_{ik} = d^{t*} + 1$. Observe that

$$|L_i^{t+1} \cap L_j^{t+1}| = |L_i^t \cap L_j^t| + |E_{ij}^t| + |E_{ji}^t| + |E_{i-j}^t \cap E_{j-i}^t|$$
$$|L_i^{t+1} \cap L_k^{t+1}| = |L_i^t \cap L_k^t| + |E_{ik}^t| + |E_{ki}^t| + |E_{i-k}^t \cap E_{k-i}^t|,$$

where $E_{m-n} = E_m \setminus E_{mn}$. By 5, $|E_{ij}^t| = |E_{ik}^t|$. Since the decision problem is symmetric, it must be that $|E_{ji}^t| = |E_{ki}^t|$ as well. Moreover, that both j and k are i's degree $d^* + 1$ neighbors implies that j and k are within the same d^* neighborhood. By 6, all groups within a d^* neighborhood learn the same subset of elements from their degree $d^* + 1$ neighbors. It follows that $E^t_{j-i} = E^t_{k-i}$ up to a subset of measure zero. Therefore, $|E^t_{i-j} \cap E^t_{j-i}| = |E^t_{i-k} \cap E^t_{k-i}|$.[4] As a result, condition (Equation 2.15) holds in $t + 1$.

Lastly, condition (Equation 2.10) says that if an element ℓ is learnable by i from a degree d neighbor, then it cannot be in the dialect of a neighbor of degree greater than d. Suppose, for contradiction, that $\{L^{t+1}_i\}_i$ violates condition (Equation 2.10), i.e. there exists an element $\widehat{\ell}$ such that

$$\widehat{\ell} \in L^{t+1}_j \setminus L^{t+1}_i \quad \wedge \quad \widehat{\ell} \in L^{t+1}_k$$

for some i, j, k such that $d_{ij} < d_{ik}$. Since $\{L^t_i\}_i$ is localized and $L^{t+1}_k = L^t_k \cup E^t_k \cup N^t_k$, it must be the case that either $\widehat{\ell} \in E^t_{jk} \setminus E^t_{ik}$ or $\widehat{\ell} \in E^t_k$. $\widehat{\ell} \in E^t_{jk} \setminus E^t_{ik}$ contradicts the assumption that when $d_{ik} = d_{jk}$, i and j learn the same set from k.

$\widehat{\ell} \in E^t_k$ and $\widehat{\ell} \in L^{t+1}_j \setminus L^{t+1}_i$ imply that

$$\widehat{\ell} \in (E^t_k \cap (L^t_j \cup E^t_j)) \quad \wedge \quad \widehat{\ell} \notin (E^t_i \cap (L^t_j \cup E^t_j)).$$

But this is inconsistent with 2.1 and 2.2. If $d_{ij} \leq d^{t*}$, then there is no element in $L^t_j \cup E^t_j$ that i does not know, hence contradicting the second conjunct. If $d_{ij} > d^{t*}$, k will not choose to learn anything from j. This is true because when $d_{ij} < d_{ik}$, we have $d_{ki} = d_{kj}$. By $d_{ij} < d_{ik}$ and $d_{ij} \leq d^{t*}$, it must be the case that $d_{ik} > d^{t*} + 1$. Combining this fact with 2.1, k will not learn anything from i, and so neither will k learn from j. Hence we have a contradiction with the first conjunct. Therefore, condition (Equation 2.10) must hold for $\{L^{t+1}_i\}_i$.

This completes the proof. □

B.2.2 Proof of 2.3

To prove this proposition, it is helpful to first introduce the following lemma:

Lemma B.1. *Let $\{L^t_i\}_{i \in \mathscr{G}}$ be symmetric and localized. Then, $d^{t*} \leq \overline{d}$, where d^{t*} is determined by (Equation 2.13) for every $t \geq 1$.*

Proof. First, observe that, for all $t \geq 1$, all $i \in \mathscr{G}$,

$$\sum_{k=1}^{d} |P^t_i(k)| \geq (2^d - 1)|N^*|, \qquad \forall d \in \{1, \dots, D\}. \tag{B.3}$$

---

[4] We do not actually need 6 to prove that $|E^t_{i-j} \cap E^t_{j-i}| = |E^t_{i-k} \cap E^t_{k-i}|$; 5 alone is enough, although the exposition will be more complicated.

This is true because, from 2.1, we know that each group invents a measure of $|N^*|$ new linguistic elements in every $t-1$ ($t = 1, 2, \dots$). Hence in t, the set of learnable elements within a degree d neighborhood for group i, $\bigcup_{k=1}^{d} P_i^t(k)$, must contain at least, and potentially more than, a measure of $(2^d - 1)|N^*|$ elements. Recall that d^{t*} and $\overline{d}$ are determined by conditions (Equation 2.13) and (Equation 2.18), respectively. According to (Equation 2.18), $|P_i^t(d)| = 2^{d-1}|N^*|$ for $d \leq \overline{d}$, and so (Equation B.3) holds with equality for $d \leq \overline{d}$. Therefore, conditions (Equation 2.13) and (Equation 2.18) imply that $d^{t*} \leq \overline{d}$ for all $t \geq 1$. □

This lemma says that regardless of initial linguistic composition, no group will ever learn the all the learnable elements of all of its $\overline{d}+1$ degree neighbors.

Now we are ready to prove 2.3 itself.

Proof. First, consider i, j such that $d_{ij} > \overline{d}+1$. From B.1, $d^{t*} \leq \overline{d}$. Trivially, $d^{t*}+1 \leq \overline{d}+1$. Therefore, if $d_{ij} > \overline{d}+1$, then $d_{ij} > d^* + 1$. According to Corollary 1, i will never learn any elements of j's language. Every period, $|L_j \setminus L_i|$ increases by exactly $|N^*|$.

Next, consider i, j such that $d_{ij} = \overline{d}+1$ and $d^* < \overline{d}$. Since $d^*+1 < d_{ij}$, by Corollary 1 i will not learn any element of j's language.

Next consider i, j such that $d_{ij} = \overline{d}+1$ and $d^* = \overline{d}$. From Equation 2.8, $|P_i(d)| \geq 2^{d-1}|N^*|$ for all d. From Corollary 1, i learns $|E_i^*| - \sum_{k=0}^{d^*} |P_i(k)|$ from its d^*+1 neighbors. The amount that i learns from its degree $\overline{d}+1$ neighbors, $|E_i^*| - \sum_{k=0}^{\overline{d}} |P_i(k)|$, is maximized when $|P_i(d)| = 2^{d-1}|N^*|$ for all $d \leq \overline{d}$. Since $d^* = \overline{d}$ it must be true by 2.2 that $|E_i^*| < \sum_{k=0}^{\overline{d}+1} 2^{k-1}|N^*|$. According to 5, i learns an equal measure of each of its degree $\overline{d}+1$ neighbor's languages, so$i$ will not learn more than $|N^*|$ of any single one of its degree $\overline{d}+1$ neighbors' languages. Therefore $|E_{ij}^*| < |N_j|$ and $|L_j \setminus L_i|$ increases every period. □

B.2.3 Proof of 2.4

Proof. According to 2.6, we need to show that if condition (Equation 2.19) holds for t then it also holds for $t+1$.

Suppose (Equation 2.19) is true for t. We know from 2.1 that between t and $t+1$, each group invents a set of new elements of measure $|N^*|$. Therefore, within a degree $\overline{d}$ neighborhood, a measure of $2^{\overline{d}}|N^*|$ new elements will have been invented by the end of period t. For $\overline{d} = 0$, it is trivially true that (Equation 2.19) holds for $t+1$, because a dialect is always the same as itself. Consider $\overline{d} \geq 1$. At $t+1$, dialects within a $\overline{d}$ neighborhood will stay as one language—i.e. $L_i^{t+1} \cup E_i^{t+1} = L_j^{t+1} \cup E_j^{t+1}$ for all i, j such that $d_{ij} \leq \overline{d}$—if and only if the marginal benefit at $|E^*| = (2^{\overline{d}} - 1)|N^*|$ outweighs the marginal cost. But this is true by how $\overline{d}$ is determined in (Equation 2.18). Therefore, condition (Equation 2.19) holds for $t+1$. $\square$

B.2.4 Proof of 2.5

One additional lemma will be useful in proving 2.5.

Lemma B.2. *For any $t \geq 1$ and $\overline{d} \geq 1$, we have*

$$d^{t*} < \overline{d} \quad \Rightarrow \quad |E^{t*}| > (2^{\overline{d}} - 1)|N^*|,$$

and

$$d^{t*} = \overline{d} \quad \Rightarrow \quad |E^{t*}| \geq (2^{\overline{d}} - 1)|N^*|.$$

Proof. Let $|\overline{E}|$ denote the measure of elements learned by a group in the steady state $\overline{\Lambda}$. From 2.4, we know that $|\overline{E}|$ is at least $(2^{\overline{d}} - 1)|N^*|$. Consider the first implication. Suppose $d^{t*} < \overline{d}$, so that the region's linguistic composition is *out of* the steady state. Then the marginal benefit of learning elements on the interval $\left[(2^{\overline{d}} - 1)|N^*|, |\overline{E}|\right]$ is at least $\frac{r+\beta}{1+r}\pi(d^{t*})$, which is strictly higher than the marginal benefit of the elements on the same interval if a group were learning *in* the steady state. This is illustrated in Figure B.1. In the top panel, the linguistic composition if out of steady state ($d^* = 2$); in the bottom panel, the linguistic composition is in steady state ($\overline{d} = 3$).[5] On the interval $\left[(2^{\overline{d}} - 1)|N^*|, |\overline{E}|\right]$, the marginal benefit in the top panel is higher than that in the bottom panel. Since the marginal cost of learning is the same both in and out of the steady state, it follows that a group must learn a strictly larger measure of elements out of steady state. As a consequence, $|E^{t*}| > |\overline{E}| \geq (2^{\overline{d}} - 1)|N^*|$ and the first implication is established. The second implication follows trivially from the fact that when $d^{t*} = \overline{d}$, the lower bound in 2.2 applies to both $|E^{t*}|$ and $|\overline{E}|$. But since the latter is bounded below by $(2^{\overline{d}} - 1)|N^*|$ according to 2.4, so must $|E^{t*}|$. $\square$

Now we are ready to prove 2.5 itself.

Proof of 2.5. If $\overline{d} = 0$, then by (Equation 2.13) and (Equation 2.18) we have $d^{t*} = 0$ for all $t \geq 1$. Consequently, the steady state $\overline{\Lambda}$ is achieved at the end of the first period at the latest. Suppose $\{L_i^0\}_i$ is such that $L_i^0 = L_j^0$ for all $i, j \in \widetilde{\mathscr{G}} \subseteq \mathscr{G}$. Then, at the beginning of $t = 1$, each L_k will differ by at least a measure of $|N^*|$ elements. Since $d^{t*} = 0$ for all t, we

---

[5] To simplify the drawing, we assumed that the marginal benefit of learning has only one step for each d_{ij}. See footnote 28 for more detail.

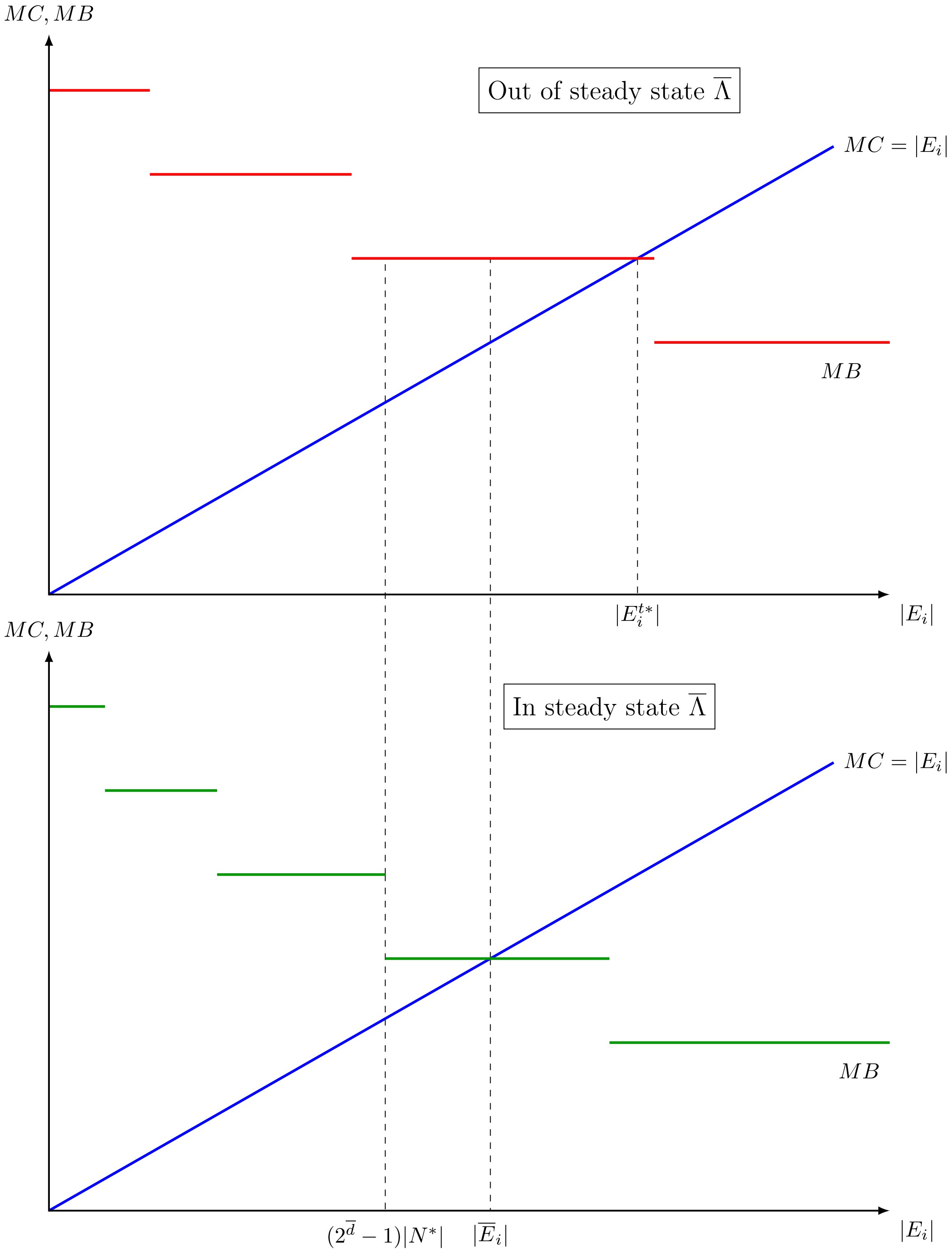


Figure B.1: Marginal Benefit of Learning in and out of Steady State must have $|E_i^{1*}| < |N^*|$. Thus, for all $t \geq 2$, it must be the case that $L_i^t \neq L_j^t$ for all $i, j \in \mathscr{G}$. If, on the other hand, not all of the initial set of dialects are identical and $d^{t*} = 0$, then *a fortiori*, $L_i^t \neq L_j^t$ for any $i, j \in \mathscr{G}$ and any $t \geq 1$.

For the remainder of the proof, we assume $\overline{d} \geq 1$, and the proof proceeds as follows: First we show that after the initial period, all pairs of groups i and j with $d_{ij} > \overline{d}$ will never speak the same language. Then, we show that all groups within a degree $\overline{d}$ neighborhood will eventually share a common language as defined in 2.2.

Let $\{L_i^0\}_i$ be symmetric and localized, and consider a pair of groups i, j with $d_{ij} > \overline{d}$. By B.1, we know that $d_{ij} > d^{t*}$. This implies that, for each t, the measure of i's set of learnable elements is strictly greater than the measure of the set that i actually learns: $\sum_{k=0}^{d_{ij}} |P_i^t(k)| > |E_i^{t*}|$. According to 2.1, therefore, it is never optimal for i to learn the entire set of $L_j^t \setminus L_i^t$ for every $t \geq 0$. As a consequence, we have

$$d_{ij} > \overline{d} \quad \Rightarrow \quad L_i^t \cup E_i^t \neq L_j^t \cup E_j^t, \qquad \forall t \geq 1.$$

Next, consider a pair of groups i, j for which $d_{ij} \leq \overline{d}$. In each period t, i's closest $2^{\overline{d}} - 1$ neighbors invent a measure of $(2^{\overline{d}} - 1)|N^*|$ new linguistic elements. B.2 shows that when d^{t*} is less than (or equal to) $\overline{d}$, the size of the set of elements learned by i is strictly (or weakly) greater than the number of elements newly invented by the closest $2^{\overline{d}} - 1$ neighbors. Since each L_i is of a finite measure, and the rate of linguistic convergence (i.e. size of acquired elements from existing dialects) is higher than the rate of linguistic divergence (i.e. size of invented elements) within a degree $\overline{d}$ neighborhood, dialects within such a neighborhood will attain $L_i \cup E_i = L_j \cup E_j$ in finitely many periods and will remain that way thereafter. As a result, there exists a $T < \infty$ such that for all $t \geq T$, we have $L_i^t \cup E_i^t = L_j^t \cup E_j^t$ for all i, j with $d_{ij} \leq \overline{d}$. This completes the proof. □

B.3 Figures

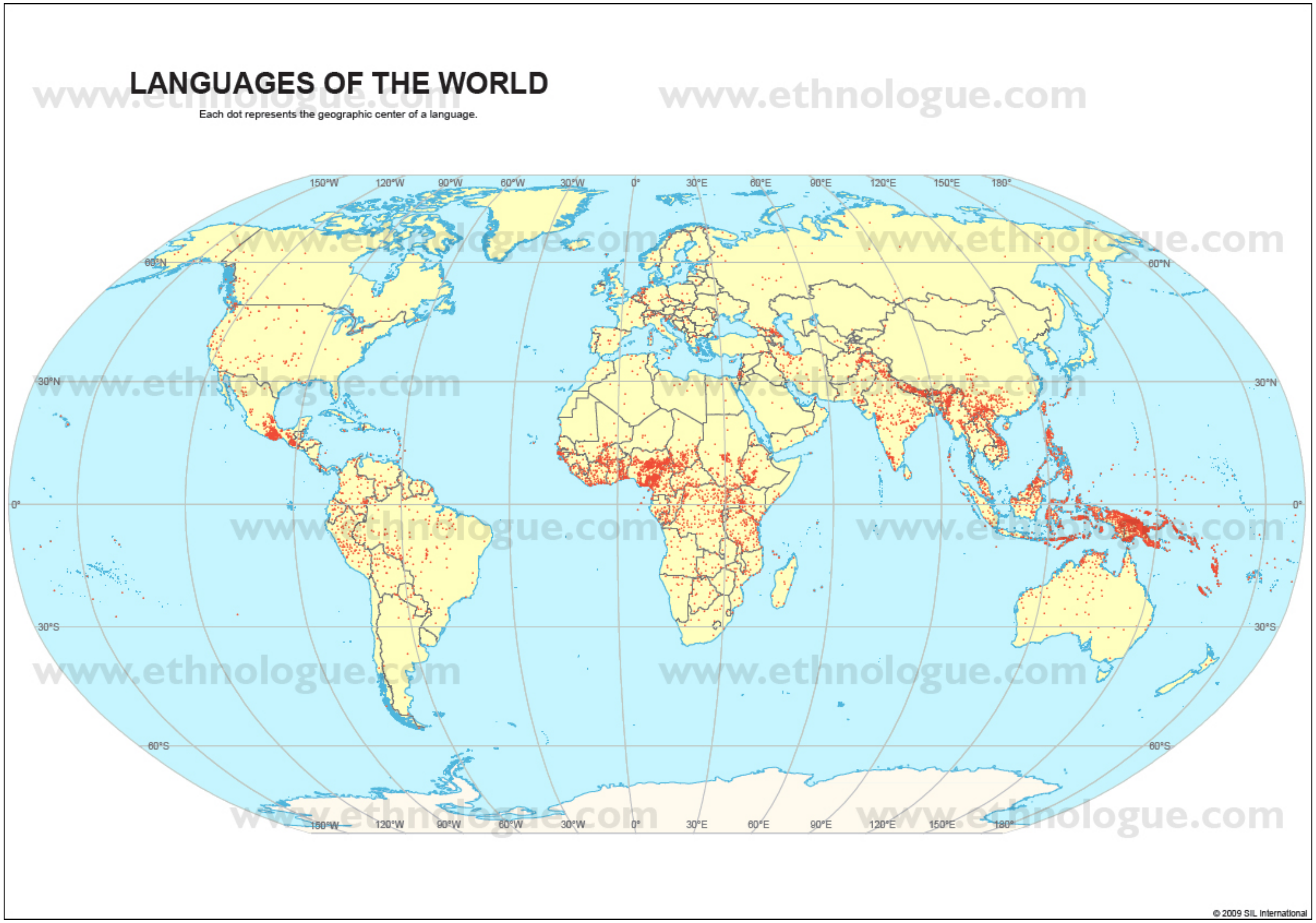


Figure B.2: Distribution of Languages in the World
Each red dot indicates the geographic center of a language. Source: Lewis (2009, `http://archive.ethnologue.com/16/show_map.asp?name=World&seq=10`).

Appendix C

Chapter 3 Appendix

C.1 Proofs of Formal Propositions

Lemma 1 (x_A and x_B as functions of σ).

(a) From (8) $x_A(\sigma) + x_B(\sigma) = (s_A + s_B)/\varphi(s_A, s_B)$, where $\varphi(s_A, s_B) \equiv s_A/[1 + \sigma^{-1/(1-\alpha)}]^\alpha + s_B/[1 + \sigma^{1/(1-\alpha)}]^\alpha$ from (5b). Using $\sigma \equiv s_A/s_B > 0$ from (5a), some algebra shows that $x_A(\sigma) + x_B(\sigma) > 1$ iff the following expression is positive:

$$\sigma[(2 + \sigma^{-1/(1-\alpha)} + \sigma^{1/(1-\alpha)})^\alpha - (1 + \sigma^{1/(1-\alpha)})^\alpha]$$
$$+ (2 + \sigma^{-1/(1-\alpha)} + \sigma^{1/(1-\alpha)})^\alpha - (1 + \sigma^{-1/(1-\alpha)})^\alpha$$

This is true because each line is strictly positive.

(b) From (8) we have $x_A(\sigma) \equiv s_A/\varphi(s_A, s_B)$. From (4) and (5c) this can be rewritten as

$x_A(\sigma) = s_A / \max \{s_A L_A^\alpha + s_B L_B^\alpha$ subject to $L_A \geq 0, L_B \geq 0, L_A + L_B = 1\}$ or

$x_A(\sigma) = 1 / \max \{L_A^\alpha + (1/\sigma) L_B^\alpha$ subject to $L_A \geq 0, L_B \geq 0, L_A + L_B = 1\}$

The right hand side is increasing in σ due to the envelope theorem. Thus $x_A'(\sigma) > 0$ for all $\sigma > 0$. A similar argument shows that $x_B'(\sigma) < 0$ for all $\sigma > 0$.

(c) Write $x_A(\sigma) = 1 / \max \{L_A^\alpha + (1/\sigma) L_B^\alpha$ subject to $L_A \geq 0, L_B \geq 0, L_A + L_B = 1\}$ as in (b). Because multiplication of the objective function by the constant $(1/s_A)$ has no effect on the solution, the optimal L_A and L_B in the denominator are given by (5d) for $n = 1$. Making this substitution, it can be shown that the denominator of $x_A(\sigma)$ approaches infinity as $\sigma \to 0$ and approaches 1 as $\sigma \to \infty$. This gives $x_A(0) = 0$ and $x_A(\infty) = 1$.

(d) The proof parallels (c).

Proposition 1 (war and peace).

Attack is a dominant strategy for A iff $p_A h(n_A) > s_A n_A^{\alpha-1}$ or equivalently $n_A/N > x_A(\sigma)$. Attack is a dominant strategy for B iff $p_B h(n_B) > s_B n_B^{\alpha-1}$ or equivalently $n_B/N > x_B(\sigma)$. When neither inequality holds, there is peace as in (7). If both A and B attack, then $n_A/N + n_B/N > x_A(\sigma) + x_B(\sigma) > 1$ where the second inequality is obtained from (9a). This is impossible because $n_A + n_B = N$. Thus A and B cannot both attack. The results in (a), (b), and (c) follow from the first three sentences above.

To show that equality of marginal products implies peace, fix $N > 0$ and $\sigma > 0$. Write $L_A^* = N/[1 + \sigma^{-1/(1-\alpha)}]$ and $L_B^* = N/[1 + \sigma^{1/(1-\alpha)}]$ as in (5d), where $L_A^* + L_B^* = N$. This is the unique labor allocation that equates marginal products across sites, and it is also the unique allocation that equates average products across sites. Let the total food output from (L_A*, L_B*) be

$$Y^* = s_A(L_A^*)^{\alpha} + s_B(L_B^*)^{\alpha}$$
$$= H(N) = \max \{s_A L_A^{\alpha} + s_B L_B^{\alpha} \text{ subject to } L_A \geq 0, L_B \geq 0, L_A + L_B = N\}.$$

Because the average products are equal, we have $Y^*/N = s_A(L_A^*)^{\alpha-1} = s_B(L_B^*)^{\alpha-1}$. Peace is strictly better than war for group A when $Y^*/N = s_A(L_A^*)^{\alpha-1} > (L_A^*/N)h(L_A^*) = H(L_A^*)/N$. This holds because $L_A^* < N$ gives $H(L_A^*) < H(N) = Y^*$. The proof is similar for B. This shows that peace is strictly better for each group, so case (b) holds with strict inequalities.

<u>Proposition 2</u> (interior locational equilibria).

For any given allocation n there are three possibilities: (a) B attacks; (b) there is peace; or (c) A attacks. Using Proposition 1, we consider each case in turn.

(a) B attacks iff $n_A/N < 1\text{-}x_B$ or equivalently $n_A/n_B < (1\text{-}x_B)/x_B$. The utility functions are those from the warfare case in (10). The inequalities in (11) yield $\eta^{1/\alpha} \leq n_A/n_B \leq (1/\eta)^{1/\alpha}$. Together this gives Proposition 2(a).

(b) There is peace iff $1\text{-}x_B \leq n_A/N \leq x_A$ or equivalently $(1\text{-}x_B)/x_B \leq n_A/n_B \leq x_A/(1\text{-}x_A)$. The utility functions are those from the peace case in (10). The inequalities in (11) yield $(\sigma\eta)^{1/(1-\alpha)} \leq n_A/n_B \leq (\sigma/\eta)^{1/(1-\alpha)}$. Together this gives Proposition 2(b).

(c) A attacks iff $x_A < n_A/N$ or equivalently $x_A/(1\text{-}x_A) < n_A/n_B$. The utility functions are those from the warfare case in (10). The inequalities in (11) yield $\eta^{1/\alpha} \leq n_A/n_B \leq (1/\eta)^{1/\alpha}$. Together this gives Proposition 2(c).

<u>Proposition 3</u> (migration).

(a) Follows from Definition 1, the construction of LE_B, LE_P, and LE_A in Proposition 2, and Definition 2(a).

(b) Suppose LE_B is empty and consider two possibilities:

(i) $m_A/m_B < (1\text{-}x_B)/x_B$, which yields war; and

(ii) $(1\text{-}x_B)/x_B \leq m_A/m_B$ with m_A/m_B below LE_P, which yields peace.

In case (i), the utility functions for war in (10) give $u_A(n) = [\varphi(s_A, s_B)/N]n_A^{\alpha}$ and $u_B(n) = [\varphi(s_A, s_B)/N]n_B^{\alpha}$ for all n with $n_A \leq m_A$. The fact that LE_B is empty implies $(1-x_B)/x_B \leq \eta^{1/\alpha}$ so we have $n_A/n_B < \eta^{1/\alpha}$ for all such n. This gives $u_A(n) < \eta u_B(n)$ in (11) for all such n. From Definition 2(c) we obtain the final allocation n = (0, N).

In case (ii), the utility functions for peace in (10) give $u_A(n) = s_A n_A^{\alpha-1}$ and $u_B(n) = s_B n_B^{\alpha-1}$ for all n with $n_A \geq m_A$ and n_A/n_B below LE_P. From the construction of LE_P, this implies $\eta u_A(n) > u_B(n)$ for all such n. The allocation n with the smallest $n_A \geq m_A$ such that $\eta u_A(n) = u_B(n)$ is the one where n_A/n_B equals the lower bound of LE_P.

(c) Suppose LE_B is non-empty and consider two possibilities:

(i) m_A/m_B is below the lower bound of LE_B, which yields war;

(ii) m_A/m_B is between the upper bound of LE_B and the lower bound of LE_P, which may yield either war or peace.

In case (i), the argument is the same is in (b)(i) above, except that now $n_A/n_B < \eta^{1/\alpha}$ follows from the fact that all allocations under consideration have n_A/n_B below the lower bound of LE_B.

In case (ii), suppose $(1/\eta)^{1/\alpha} < m_A/m_B < (1-x_B)/x_B$, which yields war. The utility functions for war in (10) give $u_A(n) = [\varphi(s_A, s_B)/N]n_A^{\alpha}$ and $u_B(n) = [\varphi(s_A, s_B)/N]n_B^{\alpha}$ for all n with $n_A \geq m_A$ and $n_A/n_B < (1-x_B)/x_B$. From $(1/\eta)^{1/\alpha} < m_A/m_B \leq n_A/n_B$, at any such n we have $u_B(n) < \eta u_A(n)$. This includes the allocation m. From Definition 2(b), no such n can be a final allocation. Now consider n with $n_A/n_B \geq (1-x_B)/x_B$ where n_A/n_B is below the lower bound of LE_P. Any such n yields peace. At $n_A/n_B = (1-x_B)/x_B$ the function $u_B(n)$ is continuous while $u_A(n)$ has an upward jump. This maintains $u_B(n) < \eta u_A(n)$. From the construction of LE_P, the allocation n with the smallest $n_A \geq m_A$ such that $u_B(n) = \eta u_A(n)$ is the one where n_A/n_B equals the lower bound of LE_P. If instead the initial allocation has $m_A/m_B \geq (1-x_B)/x_B$ and m_A/m_B is below the lower bound of LE_P, we repeat the last part of the argument above.

(d) The argument is symmetric to case (b) above.

(e) The argument is symmetric to case (c) above.

Lemma 2 (group size ratio).

Fix $\sigma^t \in (0, \infty)$ and $m_A^t/m_B^t \in (0, \infty)$. We proceed in the following steps.

(a) Recall the definitions of $x_A(\sigma)$ and $x_B(\sigma)$ in (8). From Proposition 2, the pair $(\sigma^t, m_A^t/m_B^t)$ determines whether m_A^t/m_B^t is in one of the sets LE_B, LE_P, or LE_A. If it is, then from Definition 2(a) we have $n_A^t/n_B^t = m_A^t/m_B^t$.

(b) If m_A^t/m_B^t is not in one of the sets LE_B, LE_P, or LE_A, this and the fact that m^t is interior implies that one of the four cases (b)-(e) in Proposition 3 applies. From Proposition 2, σ^t determines whether LE_B is empty or non-empty and likewise for LE_A. The ratios $(\sigma^t, m_A^t/m_B^t)$ together determine which of (b)-(e) in Proposition 3 applies, and also the final allocation n^t, where n^t must be $(N^t, 0)$, $(0, N^t)$, the lower bound of LE_P, or the upper bound of LE_P.

(c) If step (a) applies with $m_A^t/m_B^t \in LE_P$, or step (b) applies and n^t is the lower or upper bound of LE_P, there is peace in period t. This follows because all ratios in the LE_P interval satisfy the conditions for peace in Proposition 1 by construction. We then obtain m_A^{t+1}/m_B^{t+1} from (12), where the new allocation m^{t+1} is interior.

(d) If step (a) applies with $m_A^t/m_B^t \in LE_B$ or $m_A^t/m_B^t \in LE_A$, there is a non-trivial war in period t and we obtain m_A^{t+1}/m_B^{t+1} from (13), where m^{t+1} is interior. If step (b) applies and n^t is $(N^t, 0)$ or $(0, N^t)$, there is a trivial war in period t, and again we obtain m_A^{t+1}/m_B^{t+1} from (13), where m^{t+1} is interior.

Proposition 4 (war and peace with Malthusian dynamics).

From Proposition 3, a non-trivial war occurs in period t+1 iff $m_A^t/m_B^t \in LE_B$ or $m_A^t/m_B^t \in LE_A$. In all other cases, either $n^{t+1} = (N^{t+1}, 0)$ or $n^{t+1} = (0, N^{t+1})$ so there is a trivial war; or $n_A^{t+1}/n_B^{t+1} \in LE_P$ so there is peace.

(a) Proposition 2(a) shows that a necessary condition for $m_A^{t+1}/m_B^{t+1} \in LE_B$ is $\eta^{1/\alpha} \leq m_A^{t+1}/m_B^{t+1} \leq (1/\eta)^{1/\alpha}$. Proposition 2(c) shows that the same condition is necessary for $m_A^{t+1}/m_B^{t+1} \in LE_A$.

(b) Proposition 2(a) shows that a necessary condition for $m_A^{t+1}/m_B^{t+1} \in LE_B$ is $m_A^{t+1}/m_B^{t+1} < [1-x_B(\sigma^{t+1})]/x_B(\sigma^{t+1})$. Proposition 2(c) shows that a necessary condition for $m_A^{t+1}/m_B^{t+1} \in LE_A$ is $m_A^{t+1}/m_B^{t+1} > x_A(\sigma^{t+1})/[1-x_A(\sigma^{t+1})]$.

When the necessary condition in (a) is combined with the necessary condition for LE_B in (b), by Proposition 2 this suffices for $m_A^{t+1}/m_B^{t+1} \in LE_B$. The result for LE_A is the same.

We want to show that one of the inequalities in (b) holds iff $\sigma^{t+1} \notin [\sigma_A^{t+1}, \sigma_B^{t+1}]$ as in Proposition 4(b). The solutions for σ_A^{t+1} and σ_B^{t+1} exist and are unique due to the continuity of x_A and x_B; the monotonicity of these functions from (9b); and the limit properties of these functions from (9c) and (9d). We next show $\sigma_A^{t+1} < \sigma_B^{t+1}$. Suppose $\sigma_A^{t+1} = \sigma_B^{t+1}$. From (9a) we have $x_A(\sigma_A^{t+1}) + x_B(\sigma_B^{t+1}) > 1$, which contradicts the definition of σ_A^{t+1} and σ_B^{t+1}. Suppose $\sigma_A^{t+1} > \sigma_B^{t+1}$. From (9b), x_A is an increasing function, so this gives $x_A(\sigma_A^{t+1}) + x_B(\sigma_B^{t+1}) > x_A(\sigma_B^{t+1}) + x_B(\sigma_B^{t+1}) > 1$, which again contradicts the definition of σ_A^{t+1} and σ_B^{t+1}. Thus $\sigma_A^{t+1} < \sigma_B^{t+1}$.

We now establish $\sigma_A^{t+1} < \sigma^t < \sigma_B^{t+1}$. Using the monotonicity of x_A and x_B in (9b), this holds iff $x_A(\sigma_A^{t+1}) < x_A(\sigma^t)$ and $1 - x_B(\sigma^t) < 1 - x_B(\sigma_B^{t+1})$. From the definitions of σ_A^{t+1} and σ_B^{t+1}, these inequalities hold iff $1 - x_B(\sigma^t) < m_A^{t+1}/N^{t+1} < x_A(\sigma^t)$. We will show that the latter pair of inequalities is always satisfied.

(i) Suppose there is a (trivial or non-trivial) war in period t. This implies $L_A^t/L_B^t = (\sigma^t)^{1/(1-\alpha)}$ where (L_A^t, L_B^t) is obtained from (5d). From (13), $m_A^{t+1}/m_B^{t+1} = (\sigma^t)^{1/(1-\alpha)}$. We know L_A^t/L_B^t equalizes average products at the productivity ratio σ^t so the same is true for m_A^{t+1}/m_B^{t+1}. Proposition 1 gives $1-x_B(\sigma^t) < m_A^{t+1}/N^{t+1} < x_A(\sigma^t)$.

(ii) Suppose there is peace in period t. From Proposition 3, this implies $n_A^t/n_B^t \in LE_P$. First consider the case in which $n_A^t/n_B^t > (\sigma^t)^{1/(1-\alpha)}$ so n_A^t/n_B^t exceeds the group size ratio that equalizes average products in period t. By (12), $m_A^{t+1}/m_B^{t+1} = \sigma^t(n_A^t/n_B^t)^\alpha$. This gives $n_A^t/n_B^t > m_A^{t+1}/m_B^{t+1} > (\sigma^t)^{1/(1-\alpha)}$. Due to $n_A^t/n_B^t \in LE_P$ and the fact that $(\sigma^t)^{1/(1-\alpha)}$ is in the interior of LE_P in period t, m_A^{t+1}/m_B^{t+1} is in the interior of the set LE_P defined by σ^t in period t. Proposition 2(b) then gives $1-x_B(\sigma^t) < m_A^{t+1}/N^{t+1} < x_A(\sigma^t)$. A parallel argument yields the same result for the case in which $n_A^t/n_B^t < (\sigma^t)^{1/(1-\alpha)}$. The only other case is $n_A^t/n_B^t = (\sigma^t)^{1/(1-\alpha)}$, which gives $m_A^{t+1}/m_B^{t+1} = (\sigma^t)^{1/(1-\alpha)}$. Again, m_A^{t+1}/m_B^{t+1} is in the interior of the set LE_P defined by σ^t and Proposition 2(b) gives $1-x_B(\sigma^t) < m_A^{t+1}/N^{t+1} < x_A(\sigma^t)$.

This concludes the proof that $\sigma_A^{t+1} < \sigma^t < \sigma_B^{t+1}$.

When $\sigma^{t+1} < \sigma_A^{t+1}$, the monotonicity of x_A gives $x_A(\sigma^{t+1}) < x_A(\sigma_A^{t+1}) \equiv m_A^{t+1}/N^{t+1}$ or $x_A(\sigma^{t+1})/[1-x_A(\sigma^{t+1})] < m_A^{t+1}/m_B^{t+1}$. When $\sigma^{t+1} > \sigma_B^{t+1}$, the monotonicity of x_B gives $1-x_B(\sigma^{t+1}) > 1-x_B(\sigma_B^{t+1}) \equiv m_A^{t+1}/N^{t+1}$ or $[1-x_B(\sigma^{t+1})]/x_B(\sigma^{t+1}) > m_A^{t+1}/m_B^{t+1}$. When $\sigma_A^{t+1} \le \sigma^{t+1} \le \sigma_B^{t+1}$, we have $[1-x_B(\sigma^{t+1})]/x_B(\sigma^{t+1}) \le m_A^{t+1}/m_B^{t+1} \le x_A(\sigma^{t+1})/[1-x_A(\sigma^{t+1})]$. Thus one of the inequalities in (b) of the proof holds iff $\sigma^{t+1} \notin [\sigma_A^{t+1}, \sigma_B^{t+1}]$ as in Proposition 4(b).

Corollary.

From Proposition 4, $\sigma_A^{t+1} < \sigma^t = \sigma^{t+1} < \sigma_B^{t+1}$ implies that there cannot be a non-trivial war in period t+1 regardless of whether there is war or peace in period t. A trivial war can be ruled out using (i) and (ii) in the proof of Proposition 4 and substituting $\sigma^t = \sigma^{t+1}$ to show that $m_A^{t+1}/m_B^{t+1} \in LE_P$ for period t+1. Proposition 3(a) then yields peace in period t+1.

C.2 Figures

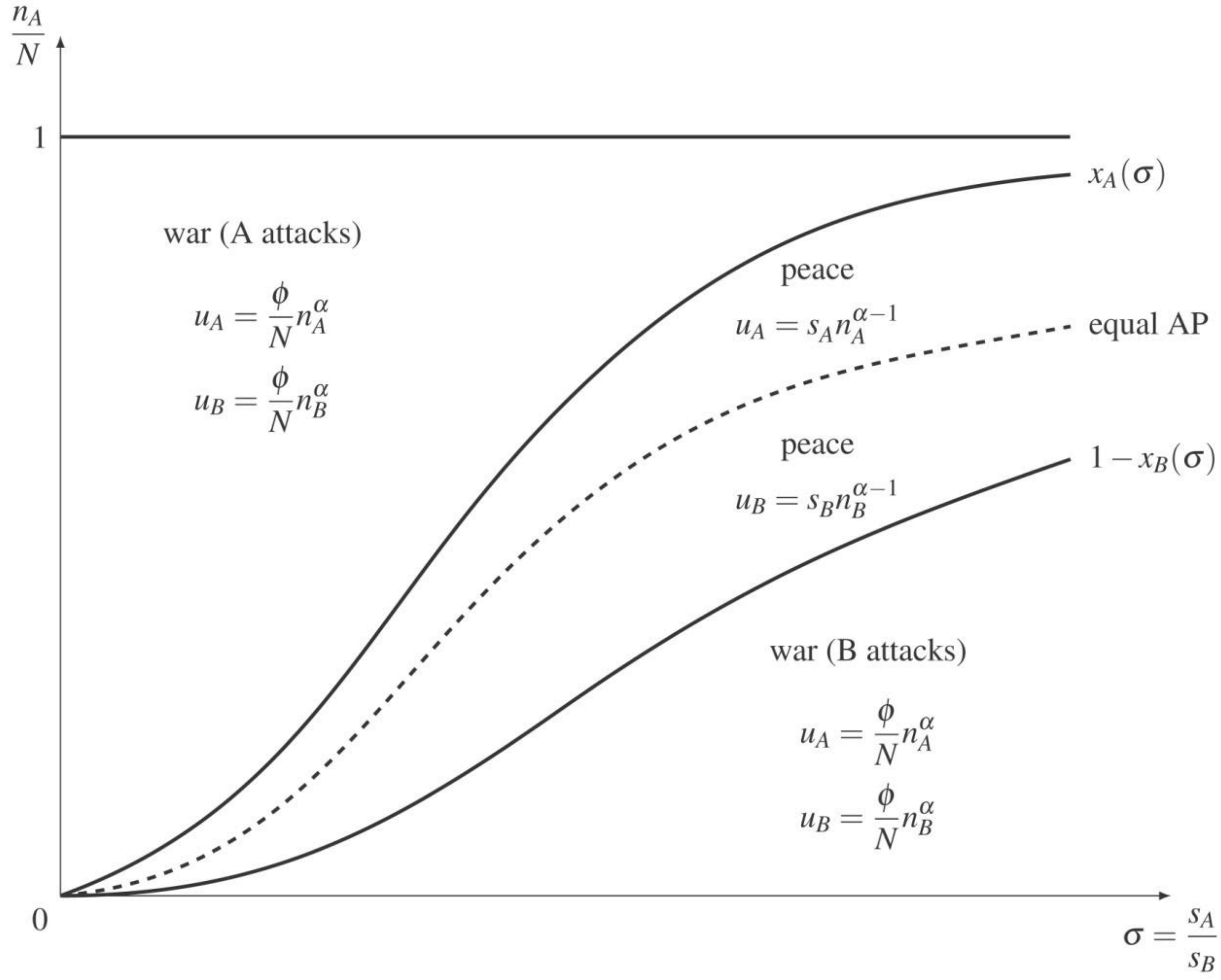


Figure C.1: War and Peace Determined by the Ratio of Site Productivities and the Ratio of Group Sizes
(AP = average product)

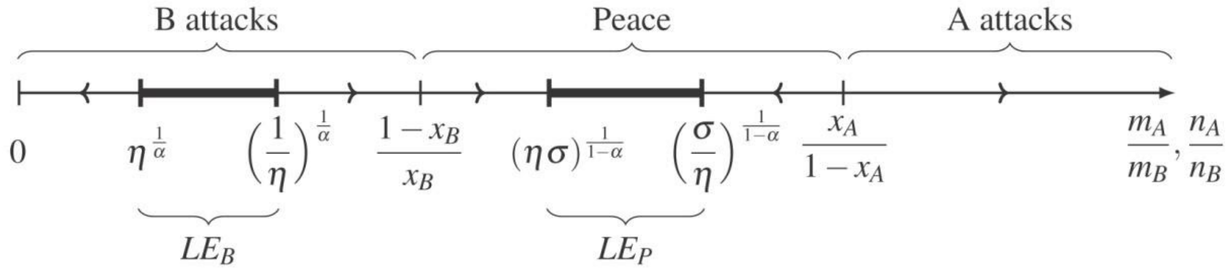


Figure C.2: Migration and Locational Equilibrium (Example where LEB Exists But LEA Does Not)

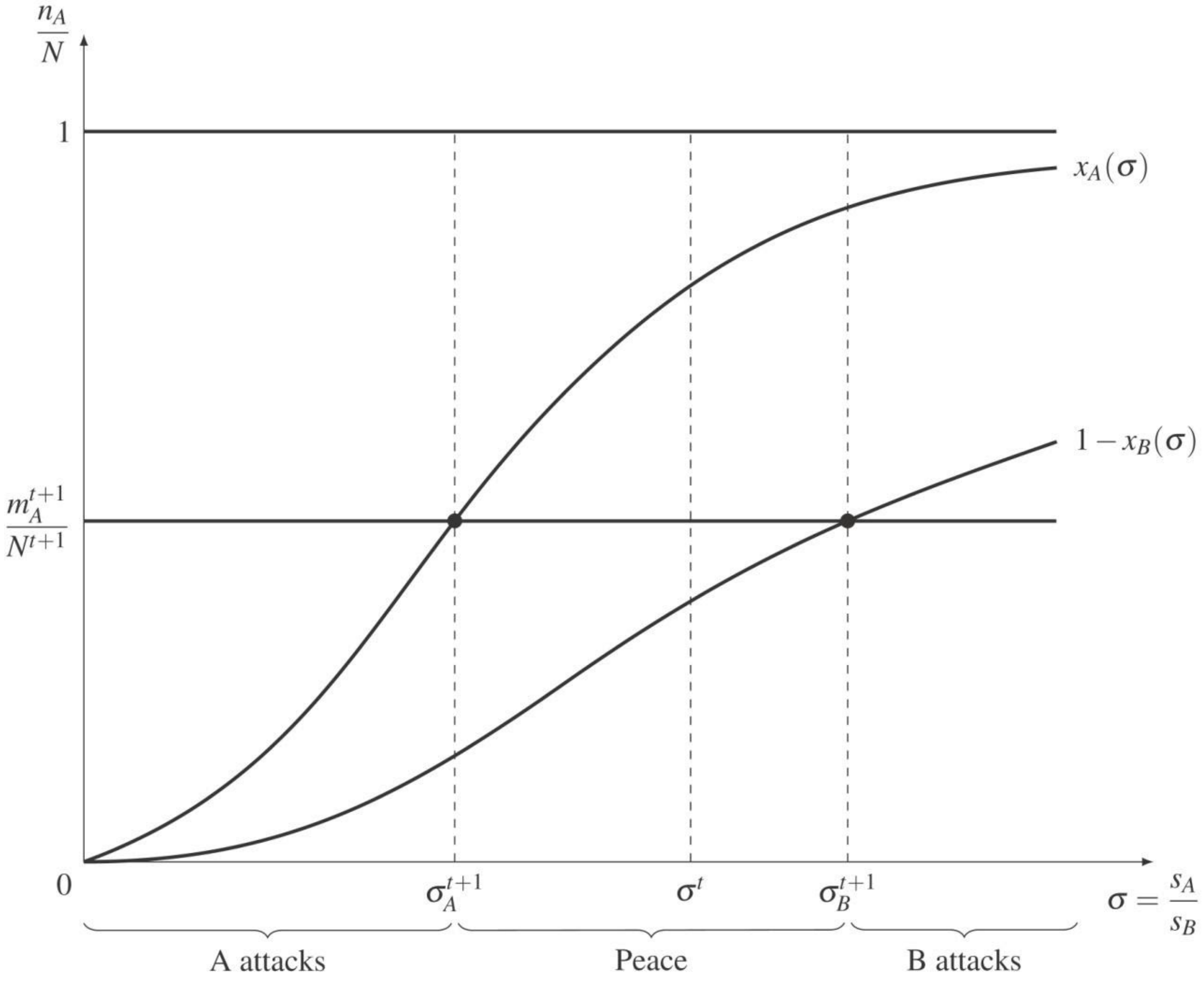


Figure C.3: Productivity Ratios for War and Peace in Period t+1 Given Initial Group Sizes