Clans, Guilds, and Markets: Apprenticeship Institutions and Growth in the Pre-Industrial Economy*

David de la Croix — Univ. cath. Louvain

Matthias Doepke — Northwestern

Joel Mokyr — Northwestern

November 2015

Abstract

In the centuries leading up to the industrial revolution, Western Europe gradually pulled ahead of other world regions in terms of technological creativity, population growth, and income per capita. We argue that superior institutions for the creation and dissemination of productive knowledge help explain the European advantage. We build a model of technological progress in a pre-industrial economy that emphasizes the person-to-person transmission of tacit knowledge. The young learn as apprentices from the old. Institutions such as the family, the clan, the guild, and the market organize who learns from whom. We argue that medieval European institutions such as guilds and specific features such as journeymanship can explain the rise of Europe relative to regions that relied on the transmission of knowledge within extended families or clans.

*We thank Alice Fabre, Georgi Kocharkov, Guido Lorenzoni, Kiminori Matsuyama, and seminar participants at Arizona, Konstanz, the Minneapolis Fed, Northwestern, Princeton, St. Andrews, UCLA, Yale, Zurich, the OLG Days conference (Marseilles), the CESifo Summer Workshop, the Christmas meeting of Belgian economists, and the SED Annual Meeting for comments that helped substantially improve the paper. Financial support from the National Science Foundation (grant SES-0820409) and from the French speaking community of Belgium (grant ARC 15/19-063) is gratefully acknowledged.

de la Croix: IRES, Universit catholique de Louvain, Place Montesquieu 3, B-1348 Louvain-la-Neuve, Belgium (e-mail: david.delacroix@uclouvain.be). Doepke: Department of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208 (e-mail: doepke@northwestern.edu). Mokyr: Department of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208 (e-mail: j-mokyr@northwestern.edu)

1 Introduction

Historically, the intergenerational transmission of skills and more generally of "knowledge how" has been central to the functioning of all economies since the emergence of agriculture. This knowledge until recently was almost entirely "tacit" knowledge, in the standard sense used today in the economics of knowledge literature (Cowan and Foray, 1997; Foray, 2004, pp. 71–73, 96–98). Although economic historians have long recognized the importance of tacit knowledge for the functioning of the economy (Dunlop, 1911, 1912), it is only more recently that it has been explicitly connected with the literature on human capital and its role in the Industrial Revolution and the emergence of modern economic growth (Humphries, 2003, 2010; Kelly, Mokyr, and Ó Gráda, 2014). The literature on the economics of apprenticeship has focused on a number of topics we shall discuss in some detail below. Yet little has been done to analyze apprenticeship as a global phenomenon, organized in different modes.

In this paper, we examine the role of apprenticeship institutions that shaped the intergenerational transmission of knowledge for explaining economic growth in the pre-industrial era. We build a model of technological progress that emphasizes the person-to-person transmission of tacit knowledge from the old to the young (as in Lucas 2009 and Lucas and Moll 2014). Doing so allows us to go beyond the simplified representations of technological progress used in existing models of pre-industrial growth, such as Galor (2011). In our setup, a key role is played by institutions such as the family, the clan, the guild, and the market, which organize who learns from whom. We argue that the archetypes of modes of apprenticeship that we consider in the model, while abstract, can be mapped into real world institutions that were prevalent throughout history in different world regions.

We use the theory to address one of the central questions about pre-industrial growth, namely why Western Europe pulled ahead of other regions in terms of technological progress and growth in the centuries leading up to industrialization. In particular, we claim that medieval European institutions such as guilds and specific features such as journeymanship played a central role in generating a faster dissemination of new productive knowledge in Europe, compared to regions that relied on the transmission of knowledge within extended families or clans.[1]

Before developing a theory of the modes of institutional organization of apprenticeship and their implications for innovation, we start by underlining three key issues one should address in modeling. First, the central issue is the extent to which the mode of organizing the transmission of skills was consistent with technological progress. We take the view from the outset that all systems of apprenticeship are consistent with at least some degree of progress. Even when the system has strong conservative elements that administer rigid tests on the existing procedures and techniques in use, there is always a certain cumulative drift over time that can take place and raise productivity even in the most conservative systems as the result of learning by doing. That said, the rates at which innovation occurred within artisanal systems has differed dramatically over time, over different societies and even between different products. Differences in rates of technological progress may in principle have two different sources, namely the rate of original innovation and the speed of the dissemination of existing ideas. While we discuss implications for original innovation, our theoretical analysis focuses on the second channel. Specifically, we ask to what extent the intergenerational transmission mechanism was conducive to the dissemination of best practice techniques and to which extent an apprenticeship system based on personal contacts and mostly local networks were conducive to closing gaps between best-practice and average-practice techniques.

Second, the training contract between master and apprentice (whether formal or implicit), for obvious reasons, represents a complicated transaction. For one thing, unless that transmission occurs within the nuclear family (in a father-son line), the person negotiating the transaction is not the subject of the contract himself but his parents, raising inevitable agency problems. Moreover, the contract written with the "master" by its very nature contains a great deal of incompleteness. The details of what is to be taught, how well, how fast, what tools and materials the pupil would be allowed to use, as well as other aspects such as room and board, are impossible to fully specify in advance. Equally, apart from a flat fee that many apprentices paid upfront, the other services rendered by the apprentice such as labor, were hard to fully list. It was, in a word an archetypical incomplete contract,

[1]Our emphasis on the role clans in organizing economic life for comparative development is shared with Greif and Tabellini (2010), although the mechanisms considered are entirely different.

and no imaginable improvement in contracting institutions could remedy that. As a consequence, in our theoretical analysis moral hazard in the master-apprentice relationship is the central element that creates a need for institutions to organize the transmission of knowledge.

Third, and as a result of the contractual problems in writing an apprenticeship agreement, a variety of institutional setups for supervising and arbitrating the apprentice-master relations can be found in the past. In all cases except direct parent-child relationships, some kind of enforcement mechanism was required. Basically three types of institutions can be discerned that enforced contracts and as a result ended up in one form or another regulating the industry. They were (1) informal institutions, based on reputation and trust; (2) non-state semi-formal institutions (guilds, local authorities such as the Dutch *neringen*); and (3) third party (state) enforcement usually by local authorities and courts. In many places all three of these worked simultaneously and should be regarded as complements, but their relative importance varied quite a bit and these weights may have been of crucial importance. In our theoretical analysis, we map the wide variety of historical institutions into four archetypes, namely the (nuclear) family, the clan (i.e., a trust-based institution comprising an extended family), the guild (a semi-formal institution), and the market (which comprises formal contract enforcement by a third party).

Our theoretical model builds on a recent literature in the theory of economic growth that puts the spotlight on the dissemination of knowledge through the interpersonal exchange of ideas.[2] Given our focus on pre-industrial growth, the analysis is carried out in a Malthusian setting with endogenous population growth in which the factors of production are the fixed factor land and the supply of effective labor by workers ("craftsmen") in a variety of trades.[3] Knowledge is represented as

[2]Specifically, the underlying engine of growth in our model is closely related to Lucas (2009), who in turn builds on earlier seminal contributions by Jovanovic and Rob (1989), Kortum (1997), and Eaton and Kortum (1999). Earlier explicit models of endogenous technological progress build on R&D efforts by firms, following the seminal papers of Romer (1990) and Aghion and Howitt (1992). While such models are useful for analyzing innovation in modern times, their applicability to pre-industrial growth is doubtful, partly because legal protections for intellectual property became widespread only recently.

[3]The reliance on the Malthusian mechanism to account for the lack of sustained growth in income per capita in most of the pre-industrial era has been standard in the literature on unified

the efficiency with which craftsmen can perform tasks. While there is some scope for new innovation, the main engine of technological progress is the transmission of productive knowledge from old to young workers. Young workers learn from elders through a form of apprenticeship. There is a distribution of knowledge (or productivity) across workers, and when young workers learn from multiple old workers, they can adopt the best technique that they have been exposed to. Through this process, average productivity in the economy increases over time.[4]

The central features of our analysis are that the transmission of knowledge (teaching) requires effort on the part of the master; that this leads to a moral hazard problem in the master-apprentice relationship; and that, as a consequence, institutions that mitigate or eliminate the moral hazard problem are key determinants of the dissemination of knowledge and growth.[5]

The "family" in our analysis is the polar case where no enforcement mechanism is available that reaches beyond the nuclear family, and hence children only learn from their own parents. In the family equilibrium, there is still some technological progress due to experimentation with new ideas and innovation within the family, but there is no dissemination of knowledge, and hence the rate of technological progress is low. The "clan" is an extended family where reputation and trust provide an informal enforcement mechanism. Hence, children can become apprentices of members of the clan other than their own parents (such as aunts or uncles). The clan equilibrium leads to faster technological progress compared to the family equilibrium, because productive new ideas disseminate within each clan. The "guild" in our model is a coalition of all the masters in a given trade that

growth theory since the seminal work of Galor and Weil (2000).

[4]Recent growth models that build on a process of this kind (in addition to Lucas 2009) include Alvarez, Buera, and Lucas (2008, 2013), König, Lorenz, and Zilibotti (2015), Lucas and Moll (2014), Luttmer (2007, 2015), and Perla and Tonetti (2014). Among these papers, Luttmer (2015) also considers a market environment where students are matched to teachers, although without allowing for different institutions. Particularly relevant to our work is also Fogli and Veldkamp (2012), where the structure of a network has important ramifications for the rate of productivity growth. The research is also related to models of productivity growth over the very long run such as Kremer (1993) and Jones (2001).

[5]Additional innovations relative to the recent literature on growth based on the exchange of ideas are that we examine endogenous institutional change, and that we allow for endogenous population growth. The relevance of Malthusian dynamics in the pre-industrial era has been recently emphasized by Clark (2007) and Ashraf and Galor (2011).

provides a semi-formal enforcement mechanism, but also regulates (monopolizes) apprenticeship within the trade. Finally, the "market" is a formal enforcement institution where an outside authority (such as the state) enforces contracts, and in addition rules are in place that prevent anticompetitive behavior (such limitations on the supply of apprenticeship imposed by guilds).

In terms of mapping the model into historical institutions, we regard most world regions (in particular China, India, and the Middle East) as being characterized by the clan equilibrium throughout the pre-industrial era. Here extended families were central in in organizing most aspects of economic life, including the transmission of skills between generations. The distinctive features of Western Europe are a much more central role of the nuclear family from the first centuries of the Common Era; little significance of extended families; and an increasing reliance of institutions that do not rely on family ties (such as cities and indeed guilds) starting in the Middle Ages. Hence, in the language of the model, we view Western Europe as undergoing a transition from the family to the guild equilibrium during the Middle Ages, and onwards to the market equilibrium in the centuries leading up to the Industrial Revolution.

To explain the emerging primacy of Western Europe over other world regions, the key issue is the comparative growth performance of the clan and guild institutions. Both the clan and the guild provided for apprenticeship outside the nuclear family, and a count against the guild is the anticompetitive nature of guilds, i.e., the possibility that guilds limited access to apprenticeship to raise prices. However, our analysis identifies a much more important force that explains why the Western European system of knowledge acquisition came to dominate. Namely, apprenticeship within guilds was independent of family ties, and thus allowed for dissemination of knowledge in the entire economy, whereas in a clan based system knowledge the dissemination of knowledge was impaired. A different side of the same coin is that in a clan based system, relatively little is gained by learning from multiple elders, because given that these elders belong to the same clan, they are likely to have received the same training and thus to have very similar knowledge. In contrast, in a guild (and also in the market) family ties do not limit apprenticeship, and hence the young can sample from a much wider variety of knowledge, implying that apprenticeship is more productive and knowledge disseminates more quickly. The historical evidence shows that, indeed, in Europe master and apprentice were far less likely to be related to each other than elsewhere. Moreover, the guild system sometimes included specific features, in particular journeymanship, that had the effect of providing access to a wide range of knowledge and fostering the spread of new techniques and ideas. In a narrower system based on blood relationships, such a wide exchange of ideas was not feasible.

Our framework can also be used to explore why institutional change (i.e., the adoption of guilds and, later on, the market) took place in Europe, but not elsewhere. If adopting new institutions is costly, the incentive to adopt will be lower when the initial economic system is relatively more successful, i.e., in a clan-based economy compared to a family-based economy. If the cost of adopting new institutions declines with population density, it is possible that new institutions will only be adopted if the economy starts out in the family equilibrium, but not of the clan equilibrium is the initial condition. We also discuss complementary mechanisms (going beyond the formal model) that are likely to also have contributed to faster institutional change in Europe.

The paper engages three recent literatures in economic history that have received considerable attention. One is the debate whether craft guilds were on balance a hindrance to technological progress, or whether they stimulated and supported it by supporting apprenticeship relations. (for a recent summary, see van Zanden and Prak, eds., 2013b and Ogilvie, 2004). The second new literature is the one that emphasizes the ingenuity of artisans and skilled workers in generating knowledge and minimizes the classic distinction between formal science and practical knowledge. Roberts and Schaffer stress the importance of "local technological projects" carried out by the "tacit genius of on-the-spot practitioners," and clearly they are thinking in terms of thoughtful and well-trained artisans who advanced the frontiers of useful knowledge (Roberts and Schaffer, 2007; see also Long 2011). Little in this literature, however, has emphasized the importance of the intergenerational transmission of the knowledge embedded in such "mindful hands" through the institutions of apprenticeship. The third literature is concerned with understanding economic, institutional, and cultural differences between Europe and other world regions as a source of of the relative rise of Europe and decline of other regions in the centuries leading up to the Industrial Revolution (e.g., Voigtländer and Voth 2013b, 2013a). Here we build in particular on the work by Greif and Tabellini (2015) on the role of clans in China versus "corporations" in Europe (i.e., formal organizations that exist independently of family ties) for sustaining cooperation (see also Greif 2006 and Greif, Iyigun, and Sasson 2012). However, Greif and Tabellini do not consider the implications of such institutions for the generation and dissemination of productive knowledge.

The paper is organized as follows. Section 2 describes key historical aspects of apprenticeship systems that our theory is based on. Our formal model of knowledge growth is described in Section 3. Section 4 analyzes the different apprenticeship institutions and derives their implications for economic growth. Section 5 makes use of theoretical results to analyze the rise of European technological primacy, and considers endogenous adoption of institutions. Section 6 concludes.

2 Historical Background

2.1 Learning of the Shopfloor

The acquisition of human capital through most of history took the form, in the felicitous phrase of De Munck and Soly (2007), of "learning on the shopfloor." One should not take this too literally: some skills had to be learned on board of ships or on the bottom of coal mines. Yet it remains true that learning took place through personal contact between a designated "master" and his apprentice.[6] As they point out (ibid., p. 6), before the middle of the nineteenth century there were few alternatives to acquiring useful productive skills. Some of the better schools, such as Britain's dissenting academies or the drawing schools that emerged on the European Continent around 1600, taught, in addition to the three R's, some useful skills such as draftmanship, chemistry, and geography. But on the whole, there was no substitute to the one-on-one learning process.

[6] We consider an era which is characterized by a sharp division of labor by gender and where formal apprenticeship generally was open only to boys. Hence, we will generally refer to master and apprentice as "he" throughout the paper, and our model does not distinguish two genders.

The economics of apprenticeship in the premodern world is thus based on the insight that each master artisan basically produced a two-fold set of connected outputs: he produced a commodity or service and he produced new craftsmen. In other words, he sold "human capital." The economics of such a setup explains many of the historical features of the system. The best-known, of course, is that the apprentice had to supply labor services to the master in partial payment for his training and his room and board. In some instances, this component became so large that the apprentice contract was more of a labor contract than a training arrangement.[7] It underlines the basic idea of joint production, in which the two activities—production and training—were strongly complementary.

As Humphries (2003) has pointed out, the contract between the master and the apprentice in any institutional setting is problematic in two ways. First, the flows of the services transacted for is non-synchronic (although the exact timing differed from occupation to occupation). Second, neither flows can be fully specified ex ante or observed ex post. The apprentice, by the very nature of the teaching process, is not in a position to assess adequately whether he has received what he has paid for until the contract is terminated. Even if the apprentice himself could observe the actual implementation of the contract, the details would be unverifiable for third parties and adjudicators. Because the transaction is non-repeated, the party who receives the services or payment first has an incentive to shirk. This is known ex ante, and therefore it is possible that the transaction does not take place and that the economy would suffer from the serious underproduction of training.[8] However, since that would mean that intergenerational transmission of knowledge would take place exclusively within families, some societies have come up with institutions that allowed the contracts to be enforced between unrelated people and strangers. These different institutions curbed opportunistically behavior in different ways, but they all require some kind of credible punishment. As we will see in our theoretical analysis below, the more sophisticated and effective in-

[7] Steffens (2001), pp. 124–25, observes on the basis of nineteenth century Belgian apprentices that little explicit teaching was carried out and that the learning was simply occurring through the performance of tasks.

[8] The suggestion by Epstein (2008, p. 61) that the contract could be rewritten to prevent either side from defaulting is not persuasive. For instance, he suggests that by backloading some of the payments from Master to apprentice, the latter would be deterred from defecting early—but that of course just shifts the opportunity to cheat from the apprentice to the Master.

stitutions led to better quality of training (in a precise matter we will define) and thus led to faster technological progress.

2.2 Apprenticeship in Western Europe

The evidence suggests that at least in early modern Europe the market for apprenticeship functioned reasonably well, despite the obvious dangers of market failure. Occupations that demanded more skill and promised higher lifetime earnings commanded higher premiums, and the differences in premiums that some masters could command meant that this market worked reasonably well in the sense that the apprenticeship premium seems to have varied positively with the expected profitability and prestige of the chosen occupation (Brooks 1994, p. 60). A good indicator to the working of the market for human capital, at least in Britain, was the premium that parents paid to a master. The premium paid, as noted by Minns and Wallis (2013), was not a full payment equal to the present value of the training plus room and board, which usually were much higher than the upfront premium. The rest normally was paid in kind with the labor provided by the apprentice. The premium served more than one purpose. In part, it was to insure the master against the risk of an early departure of the apprentice. But in part it reflected the quality of the training and the cost to the master, as well as its scarcity value (ibid., p. 340). More recently it has been shown that the premium worked as a market price reflecting rising and falling demand of certain occupations resulting from technological shocks (Ben Zeev, Mokyr, and Van Der Beek 2015). It is telling that not all apprentices paid the premiums: whereas 74 percent of engravers in London paid a premium in London, only 17 percent of blacksmiths did (Minns and Wallis, 2013, p. 344). It was clearly an option, though one that was resorted to quite often. If an impecunious apprentice could not pay, he had the option of committing to a longer indenture, as was the case in seventeenth century Vienna (Steidl, 2007, p. 143). In eighteenth century Augsburg a telling example is that a "big strong man was often taken on without having to pay any apprenticeship premium, whereas a small weak man would have to pay more." It is also recorded that apprentices with poor parents who could not afford the premium would end up being trained by a master who did inferior work (Reith, 2007, p.183). This market worked in sophisticated ways. What is clear is that human capital was recognized to be a valuable commodity. The formal contract signed by the apprentice in the seventeenth century included a commitment to protect the master's secrets and not to run away, as well as to not commit fornication (Smith, 1973, p. 150).

The precise operation of apprenticeship varied a great deal. The duration of the contract depended above all on the complexity of the trade to be learned, but also on the age at which youngsters started their apprenticeship. On the continent somewhere around three to four years seems to have been the norm (De Munck and Soly, 2007, p. 18). As would perhaps be expected, there is evidence that the duration of contracts grew over the centuries as techniques became more complex and the division of labor more specialized as a result of technological progress (Reith 2007, p. 183).

To what extent was the master-apprentice contract actually enforced? Historians have found that a substantial number of contracts were not completed (De Munck and Soly, 2007 p. 10). Wallis (2008, pp. 839-40) has shown that in late seventeenth century London a substantial number of apprentices were no longer with their original master before the seven mandated years on their apprenticeship were completed. The main reason was that the rigid seven year duration stipulated by the 1562 Statute of Artificers (which regulated apprenticeship) was rarely enforced, as were most other stipulations contained in that law (Dunlop, 1911, 1912).[9] As Wallis (2008, p. 854) remarks, "like many other areas of premodern regulation, the tidy hierarchy of the seven-year apprenticeship leading to mastery was more ideal than reality." Rather than an indication of contractual failure, the large number of apprentices that did not "complete" their terms indicates a greater flexibility in Britain. Moreover, many of them quit during their terms, and given the relatively small number of lawsuits filed against such apprentices (Rushton 1991, p. 94) it seems that many of the early departures were by mutual consent (see also Wallis 2008, p. 844).

The flexibility of the contracts in pre-Industrial Revolution England meant that risks of each party from suffering serious damage by the opportunistic behavior of the other were limited, and of course this meant that the institution would be

[9] For a more nuanced view, see Davies 1956, who argues that enforcement was a function of the economic circumstances, but agrees that there is little evidence of apprentices being sued or denied the right to exercise their occupation for having served fewer than seven years.

more successful in terms of transmitting existing skills between generations in an efficient manner. Once an apprentice had mastered the skill, there would be little point in staying on. Moreover, an apprentice might discover that the master or the trade that he was connected to was not the right one and in many documented cases he was "turned over" to another master—by some calculation this was true of 22 percent of all apprentices who did not complete their term (Wallis, 2008, pp. 842-43). There could be many reasons for this, of course, including the master falling sick or being otherwise indisposed. But we cannot rule out that at least some apprentices might have found that their master did not teach them best practice techniques or that the trade they were in was not as remunerative as some other.

The exact mechanics of the skill transmission process is very hard to trace by its very nature—after all the knowledge that was being taught was tacit, and mostly consisted of imitation and learning-by-doing. Moreover, our knowledge is to some extent biased by the better availability of more recent sources in which the experiences are recounted. It surely differed a great deal from occupation to occupation. All the same, in a felicitous phrase, Steffens (2001) has suggested that much of the learning occurred through apprentices "stealing with their eyes" (p. 131)—meaning that they were present in the shop but learned mostly through observation, imitation, and experimentation. The tasks to which apprentices were put at first, insofar that they can be documented at all, seem to have consisted of rather menial assignments such as making deliveries, cleaning and guarding the shop. Only at a later stage would an apprentice be trusted with more sensitive tasks involving valued customers and expensive raw materials (Lane, 1996, p. 77). Yet they spent most of their waking hours in the presence of the master and possibly more advance apprentices and journeymen, and as they aged they gradually would be trusted with more advanced tasks. Wallis (2008, p. 849) compares the process with what happens in more modern days amongst minaret builder apprentices in Yemen: "instruction is implicit and fragmented, questions are rarely posed, and reprimands rather than corrections form the majority of feedback." De Munck makes a similar point when he writes that "masters were merely expected to point out what had gone wrong and what might be improved" (cited in De Munck and Soly 2007, p. 16 and n. 79).

One of the most interesting findings of the new research on apprenticeship (which is central to the theory developed below) is that in Europe family ties were relatively less important than elsewhere in the world, such as India (Roy, 2013, pp. 71, 77). In China, guilds existed, but were organized along clan lines and it is within those boundaries that apprenticeship took place (Moll-Murata, 2013, p. 234). In Europe, family counted for a lot, but Europeans developed the capability of organizing themselves along professional lines without the dependence on kinship (Lucassen, de Moor, and van Zanden, 2008, p. 16). By the seventeenth century, the number of apprentices who were trained by people related to them was a distinct minority, estimated in London to be somewhere between 7 and 28 percent (Leunig, Minns, and Wallis 2011, p. 42). This may have been a decisive factor in the evolution of apprenticeship as a market phenomenon in Europe, but not elsewhere. Despite the fact that within the guilds the sons of masters received preferential treatment and that training with a relative resolved to a large extent the contractual problems, only a minority of craftsmen followed in the footsteps of their parents (Epstein and Prak 2008, p. 10). The examples of Johann Sebastian Bach and Leopold Mozart notwithstanding, the odd fact remains that in Europe few boys were actually trained by their fathers. Prak (2013, p. 153) has calculated that in the bricklaying industry, fewer than ten percent actually continued their fathers' trades.

2.3 Mobility and the Diffusion of Knowledge

In premodern Europe, as early as fifteenth century Flanders, artisans were mobile. In England, such mobility was particularly pronounced (Leunig, Minns, and Wallis 2011) and a considerable number of young lads from all over Britain came to be taught in London, not least of all the young James Watt and Joseph Whitworth, two of the heroes of the Industrial Revolution. But as Stabel (2007, p. 159) notes, towns and their guilds had to accept and acknowledge the skills acquired in other towns, even if they insisted that newcomers adapt to the economic standards set by the guilds.

Notwithstanding such constraints—more pronounced on the Continent than in England, where guilds were always less constricting—apprentices came to urban centers from smaller towns or rural regions, as was the case all over England and in eighteenth-century Vienna (De Munck and Soly, 2007, p. 17). Yet mobility of artisans and the skills they took and brought with them extended to all of Europe. The idea of the "journeyman" or "traveling companion" was that after the completion of their training they would travel to another city before they would qualify as masters, possibly to acquire additional skills that differed from the one they had been exposed to—much like postdoctoral students today (Lis, Soly, and Mitzman 1994, Robert 1979). As such, journeymanship was traditionally the "intermediate stage" between completing an apprenticeship and starting off as a fullfledged master. Such mobility was nothing new; journeymen and apprentices are known to have traveled extensively as early as the fourteenth century often on a seasonal basis, a practice known as "tramping." By the early modern period, this practice was fully institutionalized in Central Europe (Epstein, 2013, p. 59). Itinerant journeymen, Epstein argues, learned about different techniques practiced in different regions and were an important instrument in the spreading of best-practice techniques. The best evidence for this is that towns that believed to enjoy technological superiority forbade the practice of tramping and made apprentices swear not to practice their trades anywhere else, as happened to Nuremberg metal workers and Venetian glassmakers (ibid., pp. 60-61). Such prohibitions were ineffective at best and counterproductive at worst.

Not every apprentice had to go through journeymanship, and relatively little is known about how long it lasted and what kind of contracts it was based on. Journeymen have been regarded by much of the literature as basically employees of masters and were often organized in *compagnonnages* which frequently clashed with employers. Journeymen in many cases were highly skilled workers, but more mobile than masters. Known as "travellers" or "tramps," they often chose to bypass the formal status of master but prided themselves in their skills, considering themselves "equal partners" to masters (Lis, Soly, and Mitzman 1994, p. 19). Their mobility lent itself to the creation of networks in the same lines of work, and it stands to reason that technical information flowed fairly freely along those channels of communication. But skilled masters, too, traveled across Europe, often deliberately attracted by mercantilist states or local governments keen to promote their manufacturing industries through the recruitment of high-quality artisans. Technology diffusion occurred largely through the migration of skilled workers, or by apprentices traveling to the most renowned masters to learn the trade from them (and then returning home). Interestingly, such migration seems to have happened mostly to towns in which the industry already existed and which could ready upgrade its existing production techniques (Belfanti 2004, p. 581). Insofar that this occurs, such mobility helped disseminate best-practice techniques.

2.4 Apprenticeship and Guilds

A lively debate regarding the role of the guilds in premodern European economies has taken place in the past two decades. Traditionally relegated by an earlier literature to be a set of conservative, rent-seeking clubs, a revisionist literature has tried to rehabilitate craft guilds as agents of progress and technological innovation. Part of that story has been without any doubt that guilds were an institution that was instrumental in the smooth functioning of apprenticeships. As noted, given the potential for market failure due to incomplete contracts, incentive incompatibility, and poor information, the agreements about intergenerational transmission of skills needed enforcement, regulation, and supervision. Given the weakness of the political systems, the guilds stepped in and created a system that was functional and productive (Epstein and Prak 2008; Lucassen, de Moor, and van Zanden 2008; van Zanden and Prak 2013b). In a posthumously published essay, Epstein stated that the details of the apprenticeship contract had to be enforced through the craft guilds, which "overcame the externalities in human capital formation" and by punishing both masters and apprentices who violated their contracts (Epstein 2013, pp. 31-32). The argument has been criticized severely by Ogilvie (2014, 2016). Others, too, have found cases in which the easy nexus between guilds and apprenticeships proposed by Epstein and his followers does not quite hold up (Davids, 2003, 2007).

The reality is that some studies support Epstein's view to some extent, and others do not. The heated polemics have made the more committed advocates of both positions state their arguments in more extreme terms than they can defend. Guilds were institutions that existed through many centuries, in hundreds of towns, and for many occupations. This three-dimensional matrix had a huge number of elements, and it stands to reason that things differed over time, place, and occupation.

Most scholars find themselves somewhere in between. Guilds were at times hostile to innovation, especially in the seventeenth and eighteenth centuries, and under the pretext of protecting quality it collected exclusionary rents by longer apprenticeships and limited membership. But in some cases, such as the Venetian glass and silk industries, innovation was encouraged (Belfanti 2004, p. 576). Their attitude to training, similarly, differed a great deal over space and time. Davids (2013, p. 217), for instance, finds that in the Netherlands "guilds normally did not intervene in the conditions, registration, or supervision of [apprenticeship] contracts". Unger (2013, p. 203), after a meticulous survey, must conclude that the "precise role of guilds in the long term evolution of shipbuilding technology remains unclear." Moll-Murata (2013, p. 256), in comparing the porcelain industries in the Netherlands and China, retreats to a position that "contrasting the guild rehabilitationist" (Epstein's) and the guild-critical' positions is difficult to defend ...we find arguments supporting both propositions".

Guilds and apprenticeship overlapped, but they did not strictly need each other, especially not after 1600. Apprenticeship contracts could find alternative enforcement mechanisms to guilds. In the Netherlands local organizations named *neringen* were established by local government to regulate and supervise certain industries independently of the guilds. They set many of the terms of the apprenticeship contract, and often the length of contract and other details (Davids, 2007, p. 71). Even more strikingly, in Britain, most guilds were gradually declining after 1600 and exercised little control over the training procedures (Berlin, 2008). Moreover, informal institutions and reputation mechanism in many places helped make apprenticeship work even in the absence of guilds. As Humphries (2003) argues, apprenticeship contracts in England may have been to a large extent self enforcing in that opportunistic behavior in fairly well-integrated local societies would be severely punished. If both master and apprentice expected this in advance, in equilibrium they would not engage in opportunistic behavior and try to make their relationship as harmonious as possible. The limits to such self-enforcing contracts are obvious. Mobility of the apprentices after completion would mean that the informational support of reputations were costly and in larger communities they would be ineffective. Substantial opportunistic behavior could cause the cooperative equilibrium to unravel.

All the same, it has been stressed that despite the convincing evidence that guilds in some cases helped in the formation of human capital and supported innovation, the two economies in which technological progress was the fastest after 1600 were the Netherlands and Britain, the two countries in which guilds were relatively weak (Ogilvie, 2016). That such a correlation does not establish causation goes without saying, but it serves to warn us against embracing the revisionist view of guilds too rashly. In the theory articulated below, the growth implications of the guild system comes down to what the next best alternative is. If the state is sufficiently strong to enforce contracts and enable apprenticeship without guilds playing a role, the anticompetitive aspect of guilds dominates, and thus guilds hinder growth (consistent with faster growth in the Netherlands and Britain after 1600, when the state increasingly took over functions once served by guilds and cities). In contrast, when the state is weak, so that the choice is between apprenticeship provided via guilds or via clans, the easier dissemination of knowledge resulting from an institution not based on blood relationships dominates. We now turn to the theory that spells out these results in a formal model of knowledge transmission.

3 A Model of Pre-industrial Knowledge Growth

In this section, we develop an explicit model of knowledge creation and transmission in a pre-industrial setting. By pre-industrial, we mean that aggregate production relies on a land-based, agricultural technology that exhibits decreasing returns to the size of the population. In addition, there is a positive feedback from income per capita to population growth, implying that the economy is subject to Malthusian constraints. Compared to existing Malthusian models, the central new feature is that we develop an explicit model of the transmission of knowledge from generation to generation and the resulting technological progress. This feature allows us to analyze how institutions affect the transmission of knowledge in the economy, and how this interacts with the usual forces present in a Malthusian economy. The theory is based on the main historical aspects outlined in the previous section, namely the transmission of knowledge in a master-apprentice relationship that is subject to moral hazard; heterogeneity in knowledge that creates a need for mobility and an exchange of ideas; and a role for formal institutions such as guilds and markets in organizing apprenticeship.

3.1 Preferences, Production, and the Productivity of Craftsmen

The model economy is populated by overlapping generations of people who live through two periods, childhood and adulthood. All decisions are made by adults, whose preferences are given by the utility function:

$$u(c, I') = c + \gamma\, I', \tag{1}$$

where c is the adult's consumption and I' is the total future labor income of this adult's children. The parameter $\gamma > 0$ captures altruism towards children. The role of altruism is to motivate parents' investment in their children's knowledge.

The adults work as craftsmen in a variety of trades. At the beginning of a period, the aggregate economy is characterized by two state variables: The number of craftsmen N and the amount of knowledge k in the economy. Craftsmen are heterogeneous in productivity, and knowledge k determines the average productivity of craftsmen in a way that will be made precise below. We start by describing how aggregate output in our economy depends on the state variables N and k.

The single consumption good (which we interpret as a composite of food and manufactured goods) is produced with a Cobb-Douglas production function with constant return to scale that uses land X and effective craftsmen's labor L as inputs:

$$Y = L^{1-\alpha} X^{\alpha}, \tag{2}$$

with $\alpha \in (0, 1)$. The total amount of land is normalized to one, $X = 1$. Land is owned by craftsmen.[10]

The effective labor supply by craftsmen L is a CES aggregate of effective labor

[10] Our main results would be identical if a separate class of landowners were introduced. The model abstracts from an explicit farming sector; however, it would be straightforward to include farm labor as an additional factor of production (see Appendix A), or alternatively we can interpret some of the adults who we refer to as craftsmen as farmers.

supplied in different trades j:

$$L = \left(\int_0^1 (L_j)^{\frac{1}{\lambda}} dj \right)^{\lambda}, \tag{3}$$

with $\lambda > 1$. The elasticity of substitution between the different trades is $\lambda/(\lambda - 1)$. The distinction between different trades of craftsmen (watchmaker, wheelwright, blacksmith etc.) is important for our analysis of guilds below, which we model as coalitions of the craftsmen in a given trade. However, the equilibrium supply of effective labor will turn out to be the same in all trades, so that $L_j = L$ for all j. For most of our analysis, we can therefore suppress the distinction between trades from the notation.

We now relate the supply of effective labor by craftsmen L in efficiency units to the number of craftsmen N and the state of knowledge k. Craftsmen are heterogenous in knowledge. The productive knowledge of a craftsman i is measured by a cost parameter h_i, where a lower h_i implies that the master can produce at lower cost and hence has more productive knowledge. Intuitively, different craftsmen may apply different methods and techniques in their production, which vary in productivity. Specifically, the output q_i of a craftsman with knowledge h_i is given by:

$$q_i = h_i^{-\theta}. \tag{4}$$

The final-goods technology (2) is operated by a competitive industry. Given the Cobb-Douglas production function, this implies that craftsmen receive share $1 - \alpha$ of total output as compensation for their labor, and consequently the labor income of a craftsman supplying q_i efficiency units of craftsmen's labor is:

$$I_i = q_i \, (1 - \alpha) \, \frac{Y}{L}. \tag{5}$$

The heterogeneity in the cost parameter h_i among craftsmen takes the specific form of an exponential distribution with distribution parameter k:

$$h_i \sim \text{Exp}(k).$$

Given the exponential distribution, the expectation of h_i is given by $\mathrm{E}[h_i] = k^{-1}$. Hence, higher knowledge k corresponds to a lower average cost h_i and therefore higher productivity. We assume that the same k applies to all trades. Given the exponential distribution for h_i and (4), output q_i follows a Fréchet distribution with shape parameter $1/k$ and scale parameter $1/\theta$.

We can now express the total supply of effective labor by craftsmen as a function of state variables. The average output across craftsmen is given by:

$$q = \mathrm{E}(q_i) = \int_0^\infty h_i^{-\theta} \left(k \exp(-k h_i)\right) dh_i = k^\theta \Gamma(1-\theta), \tag{6}$$

where $\Gamma(t) = \int_0^\infty x^{t-1} \exp(-x) dx$ is the Euler gamma function. The total supply of effective craftsmen's labor L is then given by the expected output per craftsmen $\mathrm{E}(q_i)$ multiplied by the number of craftsmen N:

$$L = N\, k^\theta\, \Gamma(1-\theta). \tag{7}$$

3.2 Population Growth and the Malthusian Constraint

So far, we have described how total output (and hence output per adult) depends on the aggregate state variables N and k. Next, we specify how these state variables evolve over time. We start with population growth. Consistent with evidence from pre-industrial economies (see Ashraf and Galor 2011), the model allows for Malthusian dynamics.[11] The presence of land in the aggregate production function implies decreasing returns for the remaining factor L, which gives rise to a Malthusian tradeoff between the size of the population and income per capita. The second ingredient for generating Malthusian dynamics is a positive feedback from income per capita to population growth. While often this relationship is modeled through optimal fertility choice (e.g., Galor and Weil 2000), we opt for a simpler mechanism of an aggregate feedback from income per capita to

[11]Empirical work has found both a fertility and a mortality link to income per capita in medieval England, but gradually weakening over time (Kelly and Ó Gráda 2012 and 2014). The "fundamentalist" Malthusian assumption that all productivity gains eventually are translated into population growth so that the "iron law" holds fully is made here for simplification; our results for institutional comparisons would be similar in a framework that allows for growing income per capita even in the long term.

mortality rates. Every adult gives birth to a fixed number $\bar{n} > 1$ of children. The fraction of children that survives to adulthood depends on aggregate output per adult Y/N, namely:

$$n = \bar{n} \ \min\left[1, s\frac{Y}{N}\right]. \tag{8}$$

Here $\min[1, sY/N]$ is the fraction of surviving children, and n is the number of surviving children per adult. This function captures that low living standards (e.g., malnutrition) make people (and in particular children) more susceptible to transmitted diseases, so that low income per capita is associated with more frequent deadly epidemics. In recent times, we can also envision s to depend on medical technology (i.e., the invention of antibiotics would raise s). However, given that we analyze preindustrial growth, we will assume that s is fixed. We will also focus attention on a phase of development where the mortality tradeoff is still operative, so that survival is less than certain and $n = \bar{n}sY/N$. The law of motion for population then is:

$$N' = n \ N = \bar{n} \ s\frac{Y}{N}N = \bar{n} \ sY.$$

Consider a balanced growth path in which the stock of knowledge k grows according to a constant growth factor g:

$$g = \frac{k'}{k}.$$

In such a balanced growth path, the Malthusian features of the model economy impose a relationship between growth in knowledge g and population growth n, as shown in Proposition 1.

Proposition 1 (The Malthusian Constraint).
Along a balanced growth path, the growth factor of technology g and the growth factor of population n satisfy:

$$g^{\theta(1-\alpha)} = n^{\alpha}. \tag{9}$$

Proof. From (2) and (7) we have:

$$\frac{Y}{N} = \frac{L^{1-\alpha}}{N} = \Gamma(1-\theta)^{1-\alpha} \, k^{(1-\alpha)\theta} \, N^{-\alpha}. \tag{10}$$

Along a balanced growth path, income per capita Y/N is constant, and hence (9) has to hold in order to keep the right-hand side constant, too. ∎

The Malthusian constraint states that faster technological progress is linked to higher population growth. Given (9), a faster rate of technological progress is also associated with a higher level of income per capita. Income per capita is constant in any balanced growth path: Malthusian dynamics rule out sustained growth in living standards, because accelerating population growth would ultimately overwhelm productivity growth. Instead, economies with faster accumulation of knowledge will be characterized by faster population growth and hence, over time, increasing population density.

3.3 Apprenticeship, Innovation, and the Evolution of Knowledge

We now turn to the accumulation of knowledge in our model economy. In a given period, all productive knowledge is embodied in the adult workers. During childhood, people have to acquire the productive knowledge of the previous generation. There are two sources of increasing knowledge across generations. First, craftsmen are heterogeneous in their productive knowledge. Young craftsmen can learn from multiple adult craftsmen, and then apply the best of what they have learned. This knowledge dissemination process results in endogenous technological progress. In addition, after having acquired knowledge from the elders, young craftsmen can innovate, i.e., generate an idea that may improve on what they have learned, resulting in a second source of technological progress.

In order to model the idea that apprentices (or their parents) are subject to imperfect information on the efficiency of the different masters, we assume that the young can observe the efficiency of masters only by working with them as an apprentice. Consider an apprentice who learns from m masters indexed from 1 to m (the choice of m will be discussed below). The efficiency h_L learned during the apprenticeship process is:

$$h_L(m) = \min\{h_1, h_2, \ldots, h_m\}. \tag{11}$$

Hence, apprentices acquire the cost parameter of the most efficient (i.e., lowest cost) master they have learned from. After learning from masters, craftsmen attempt to innovate by generating a new idea characterized by cost parameter h_N. The quality of the idea is random, and it may be better or worse than what they know already. As adult craftsmen, they use the highest efficiency they have attained either through learning from elders or through innovation, so that the final cost parameter h' is given by:

$$h' = \min \{h_L, h_N\}. \tag{12}$$

As will become clear below, the model can generate sustained growth even if the rate of innovation is zero (i.e., own ideas are always worse than acquired knowledge). In that case, the dissemination process of existing ideas is solely responsible for growth. However, allowing for innovation allows for a positive rate of productivity growth even if each child learns only from a single master.

Recall that the distribution of the h_i among adult craftsmen is exponential with distribution parameter k. The distribution of new ideas is also exponential, and the quality of new ideas depends on existing average knowledge:

$$h_N \sim \mathrm{Exp}(\nu k).$$

That is, the more craftsmen already know, the better the quality of the new ideas that are generated. The parameter ν measures relative importance of transmitted knowledge and new ideas. If ν is close to zero, most craftsmen rely on existing knowledge, and if ν is large, innovation rather than the dissemination of existing ideas through apprenticeship is the key driver of knowledge.

The exponential distributions for ideas imply that, given the knowledge accumulation process described by (11) and (12), the knowledge distribution preserves its shape over time (as in Lucas 2009). Specifically, if each young craftsman learns from m masters that are drawn at random we have:[12]

$$h_L = \min\{h_1, h_2, \ldots, h_m\} \sim \text{Exp}(mk),$$
$$h' = \min\{h_L, h_N\} \sim \text{Exp}(mk + \nu k).$$

Hence, with m randomly chosen masters per apprentice, aggregate knowledge k evolves according to:

$$k' = (m + \nu)k. \tag{13}$$

The market for apprenticeship interacts with population growth. In particular, if each master takes on a apprentices, and each apprentice learns from m masters, the condition for matching demand and supply of apprenticeships is:

$$N'm = Na. \tag{14}$$

We ignore integer constraints and treat m and a as continuous variables. Below, we will focus on equilibria where each apprentice chooses the same number of masters m, and each master has the same number of apprentices a.

We now arrive at the core of our analysis, namely the question of how the number and identity of masters for each apprentice is determined. Apprenticeship is associated with costs and benefits. While working as an apprentice with a master, each apprentice produces $\kappa > 0$ units of the consumption good (this is in addition to the output generated by the aggregate production function). This output is controlled by the master. In turn, a master who teaches a apprentices incurs a utility cost $\delta(a)$, where $\delta(0) = 0$, $\delta'(a) > 0$, and $\delta''(a) > 0$ (i.e., the cost is increasing and convex in a). Incurring this cost is necessary for transmitting knowledge to the apprentices. We assume for simplicity that the function $\delta(\cdot)$ is quadratic, i.e. $\delta(a) = \frac{\bar{\delta}}{2}a^2$.

If a master takes on a apprentices but then puts in no effort in teaching, the apprentices still generate output κa by assisting the master in production. Thus, there is a moral hazard problem: Masters may be tempted to take on apprentices, appropriate production κa, but not actually teach, saving the cost $\delta(a)$.

[12]This result follows from the min stability property of the exponential distribution. In particular, if h_a and h_b are independent exponentially distributed random variables with rates k_a and k_b, then $\min[h_a, h_b]$ is exponentially distributed with rate $k_a + k_b$.

Dealing with this moral hazard problem is a key challenge for an effective system of knowledge transmission. The danger of moral hazard is especially severe here because the very nature of apprenticeship defines it as the mother of all incomplete contracts (see Section 2). In a modern market economy, we envision that such problems are dealt with by a centralized system of contract enforcement. In such a system, a parent would write contracts with masters to take on the children as apprentices. A price would be agreed on that is mutually agreeable given the cost of training apprentices and the parent's desire, given altruistic preferences (1), to provide the children with future income. Courts would ensure that both parties hold up their end of the bargain.

In pre-industrial societies with a lack of an effective system of contract enforcement, other institutions have to step in to make possible an effective transmission of knowledge from the elders to the young. Our view is that variation in such institutions of knowledge transmission across countries and world regions plays a central role in shaping economic success and failure in the pre-industrial era. After a brief discussion of model assumptions, we analyze specific, historically relevant institutions in the context of our model of knowledge-driven growth.

3.4 Discussion of Model Assumptions

Our model of growth in the pre-industrial economy is stylized and relies on a set of specific assumptions that yield a tractable analysis. We conclude our description of the model with a discussion of the role and plausibility of the assumptions that are most central to our overall argument. Most importantly, apprenticeship institutions matter in our economy because the knowledge of masters is not publicly observable. This creates the incentive for apprentices to sample the knowledge of multiple masters to gain productive knowledge, and implies that institutions that determine how apprentices are matched to masters matter for growth. To maintain tractability, in the model the lack of information on productivity is extreme: Nothing at all is known about the productivity of different masters, even though there is wide variation in their actual productivity. Taken at face value, this assumption is clearly implausible. However, possible concerns about the role of this assumption can be addressed in two ways.

First, in our model all productivity differences between masters are actual productivity differences, i.e., masters who know more produce more. A realistic alternative possibility is that at least some variation in knowledge is in terms of "latent" productivity, i.e., some masters may know techniques and methods that will turn out to be highly productive and important at a later time when combined with other knowledge, but do not give a productivity advantage in the present. A well known example are the various inventions of Leonardo da Vinci which could not be implemented given the knowledge of his age, but which turned into productive knowledge centuries later. Similarly, the success of the steam engine was in large part based on a set of gradual improvements in craftmen's ability in working metal to precise specifications; steam engines work only if the piston can move easily in the cylinder while also providing a tight fit. Many of these improvements in techniques would have been of comparatively little value when first invented, but turned out to be hugely important later on. Along these lines, in Appendix B we describe an extension of our model where a craftsmen output can be constrained by the state of aggregate knowledge. This version leads to exactly the same implications as the simpler setup described here, but actual variation in productivity is much smaller than variation in latent productivity, so that lack imperfect information on underlying productivity appears more plausible.

Second, it would be possible to relax the assumption of total lack of information about productivity, and instead assume that an informative, but imperfect, signal of each master's productivity was available. In such a setting, more productive masters could command higher prices for apprenticeships, they would employ a larger number of apprentices, and the spread of productive knowledge would be faster. As we document in Section 2, the historical evidence for Europe suggests that, indeed, more productive and knowledgeable masters were able to command higher prices and attract more apprentices. However, as long as information on productivity is less than perfect, the basic tradeoffs articulated by our analysis and the comparative growth implications of the institutions analyzed below would be the same. Less-than-perfect information about productivity is highly plausible. For example, even in today's world of instant communication and online discussion boards graduate students do not have perfect information about which advisor would be the best match for them, and in pre-industrial times information frictions were comparatively much more severe. We adopt the extreme case of complete lack of observability for tractability; without this assumption the distribution of knowledge would not preserve its shape over time, so that we would have to rely on numerical simulation for all results.[13]

Finally, the master-apprentice relationship in the model is simplified compared to reality. We use a setting with one-sided moral hazard, i.e., masters can cheat apprentices, but not vice versa. In reality, moral hazard was a major concern on both sides of the master-apprentice relationship. Once again, the assumption is introduced merely to simplify the analysis. It would be straightforward to introduce two-sided moral hazard in our setting, and the role of institutions for mitigating moral hazard would be unchanged. In addition, in the model apprentices interact with all masters they learn from in the same way, and they make a one-time choice of the number of masters to learn from. As ever, reality is substantially more complicated; choices of whom to learn from unfolded sequentially over time, and apprentices generally did only a single full apprenticeship, followed by other shorter interactions such as spells spent with a master during journeymanship. Once again, these assumptions are for simplicity and tractability, but are not central to our main results regarding the role of institutions for knowledge transmission. Having said that, when matching the model to data care should be taken to account for the fact that "apprenticeships" in the model correspond to a wider range of interactions (in particular journeymanship in addition to apprenticeship) in reality.

4 Comparing Institutions for Knowledge Transmission

The crucial issue in our theory is how the moral hazard problem inherent to apprenticeship is resolved. If masters do not make an effort to teach their apprentices, a parent will have no incentive to send a child to be apprenticed with a master outside the family. Without effort on the part of the master, the child would not learn anything, whereas the master would gain the apprentice's production κ. The par-

[13] Luttmer (2015) provides an alternative approach for modeling the assignment of students to teachers. Luttmer's model has the advantage of allowing for observability and not relying on an unbounded support of existing knowledge, albeit at the cost of a considerably more complex analysis.

ent would be better off keeping the child at home, thereby keeping output κ in the family. Thus, for apprenticeship outside the immediate family to be possible (and thus for knowledge to disseminate), an enforcement mechanism is needed that provides incentives for masters to put in effort.

4.1 Centralized versus Decentralized Institutions

We consider two types of institutions, characterized by centralized versus decentralized enforcement. Under centralized enforcement, people can write contracts that specify that the master has to put in effort (and also the price of apprenticeship), and there is a centralized system (such as courts) that punishes any person that breaks a contract. In contrast, in a decentralized system no such central authority exists, and instead people have to form coalitions to maintain a sufficient threat of punishment to resolve the moral hazard problem.

We should note that in reality, the distinction between centralized and decentralized institutions is less sharp than in our theory. Even where centralized enforcement institutions existed, they were often complemented by a self-enforcing mechanism based on trust and reputation (Humphries 2003; Mokyr 2008). In many societies, but most of all in England, market relationships were linked up to social relationships. Such linkages are a strong incentive toward cooperative behavior that could make even complex contracts self-reinforcing (Spagnolo 1999). Opportunistic behavior in any kind of relationship would lead to the erosion of reputation in all dimensions, and thus could be costly. For example, a master found out to treat apprentices badly might lose not only future apprentices, but also damage relations with his customers and suppliers. The same was true for the apprentices, who would be damaged in their future careers if they were known to have reneged on contracts.

To allow for the possibility of decentralized enforcement, we assume that each adult has the possibility of inflicting a utility cost (damage) on any other adult.[14] However, the punishment that a single adult can mete out is not sufficient to induce a master to put in effort, i.e., the punishment is lower than the cost of train-

[14] The cost can be interpreted as physical punishment, destruction of property, or as spreading rumors that induce others not to buy from the individual in question.

ing a single apprentice. In contrast, coalitions of multiple people can always make threats that are sufficient to guarantee compliance. An effective threat of punishment therefore requires coordination among parents. Coordination, in turn, requires communication: For shirking of a master to have consequences, the fact of the shirking has to be communicated to all would-be punishers. Thus, the extent to which people are able to communicate with each other partly determines how much knowledge transmission is possible.

Over time, societies have differed in the extent and manner in which individuals where connected in communication networks. We consider two different scenarios for decentralized enforcement, the "family" and the "clan," which we consider particularly relevant for contrasting Europe during the Early Middle Ages to China, India, and the Middle East during the same period and beyond.

The decentralized systems correspond to a period where centralized enforcement was not yet sufficiently effective. Even when courts existed, contract enforcement was often costly, slow, and uncertain. More importantly, for centuries the reach of the state and hence its courts was severely curtailed. Europe, for example, used to consist of hundreds of independent sovereign entities, and the reach of the law outside one's immediate surroundings (say, the city of residence) was limited. With this in mind, the first centralized enforcement institution that we consider is organized not by the state but by a coalition of all the masters in a given trade: a "guild." The guild monitors the behavior of its members and enforces the apprenticeship contracts between parents and masters. However, the guild also has anti-competitive features: the guild can set the price for apprenticeship, thereby exploiting its monopoly in a given trade. Guilds played a central role in organizing economic life in Europe during the Middle Ages, and our theory will allow us to assess their implications for knowledge creation and dissemination.

The final institution that we consider is the "market," where there is a centralized enforcement system for all trades as in a modern market economy. Importantly, under this institution the government not only enforces contracts, but also prevents collusion: trades are no longer allowed to form guilds that limit entry and lower competition, and both parents and masters act as price takers. The market institution corresponds to the final stages of the pre-industrial economy, when in

Europe nation states became powerful and increasingly abolished the traditional privileges of guilds.

4.2 The Family

Decentralized institutions enforce apprenticeship agreements through the formation of coalitions of parents that coordinate on a sufficient threat of punishment for shirking masters. Different decentralized institutions are distinguished by the size of these coalitions and the identity of their members. For the formation of a coalition to be feasible, the members have to be able to communicate with each other regarding the behavior of masters. Hence, one polar case is where members of different families are not able to communicate with each other, so that no coalitions can be formed. The lack of communication rules out coordinating on punishing shirking masters. As a consequence, apprenticeship outside the immediate family is impossible, i.e., each child only learns from the parent. In principle, the moral hazard problem is present even within the family. However, in utility (1) parents care about their own children, and we assume that the degree of altruism γ is sufficiently high for parents to never shirk when teaching their own children. The result is a "family equilibrium," i.e., an equilibrium where knowledge is transmitted only within dynasties, but there is no dissemination of knowledge across dynasties.

Formally, under decentralized institutions we model the knowledge accumulation decisions as a game between the craftsmen of a given generation. The strategy of a given craftsman has three elements:

1. Decide whether to send own children to others as apprentices for training, and if so, which compensation to pay to the masters of one's children.
2. Decide whether to exploit one's own apprentices (if any).
3. Decide whom to punish (if anyone).

We focus on Nash equilibria.[15] The strategy profile for the family equilibrium is as follows:

[15]Given that there are subsequent generations, one could also define a dynamic game involving all generations. However, given that preferences are of the warm-glow type, decisions of future generations do not affect the payoffs of the current generation, so that dynamic considerations do not change the strategic tradeoffs faced by the players.

- All craftsmen train their children on their own.
- If (off the equilibrium path) a master gets someone else's child as an apprentice, the master exploits the apprentice.
- No one ever punishes anyone.

If communication outside the immediate family is impossible, the family equilibrium is the only equilibrium. The family equilibrium can also occur as a "bad" equilibrium in an economy where more communication links are available, but people fail to coordinate on a more efficient punishment equilibrium.[16]

Now consider the balanced growth path under the family equilibrium. We assume that the Malthusian feedback, parameterized by the maximum number of children $\bar{n}$ and the survival function $s(\cdot)$, is sufficiently strong for dynamics to lead to a balanced growth path in which income per capita is constant.[17] The following proposition summarizes the properties of the balanced growth path.

Proposition 2 (Balanced Growth Path in Family Equilibrium).
If altruism is sufficiently strong (i.e., γ is sufficiently large), there exists a unique balanced growth path under the family equilibrium with the following properties:

(a) Each child trains only with his own parent: $m^{\mathrm{F}} = 1$, and $a^{\mathrm{F}} = n^{\mathrm{F}}$.

(b) The growth factor g^{F} of knowledge k is:

$$g^{\mathrm{F}} = 1 + \nu.$$

(c) The growth factor n^{F} of population N is:

$$n^{\mathrm{F}} = (1+\nu)^{\frac{(1-\alpha)\theta}{\alpha}}.$$

(d) Income per capita y^{F} is constant and satisfies:

$$y^{\mathrm{F}} = \frac{Y}{N} = \frac{(1+\nu)^{\frac{(1-\alpha)\theta}{\alpha}}}{\bar{n}s}.$$

[16] For any communication structure, the family equilibrium always exists, because in the expectation that no one else will punish shirking masters it is optimal to (i) never punish shirking masters either and (ii) not send one's own children to be apprenticed outside the family.

[17] The required assumptions can be made precise; what is key is that maximum population growth is larger than the maximum rate of productivity growth.

Proof. See Appendix C. ■

The condition for sufficient altruism reflects that parental altruism should be strong enough to overcome the disutility of teaching one's children. The rate of technological progress is positive in the family equilibrium, but small. This is because the only source of progress is the new ideas of craftsmen (recall that ν measures the quality of new ideas). New ideas are passed on to children, which makes children, on average, more productive than the parents. However, knowledge does not disseminate across dynasties. Given the growth rate of knowledge $g^{\mathrm{F}} = 1 + \nu$, Malthusian dynamics ensure that population grows just fast enough to offset productivity growth and yield constant income per capita.

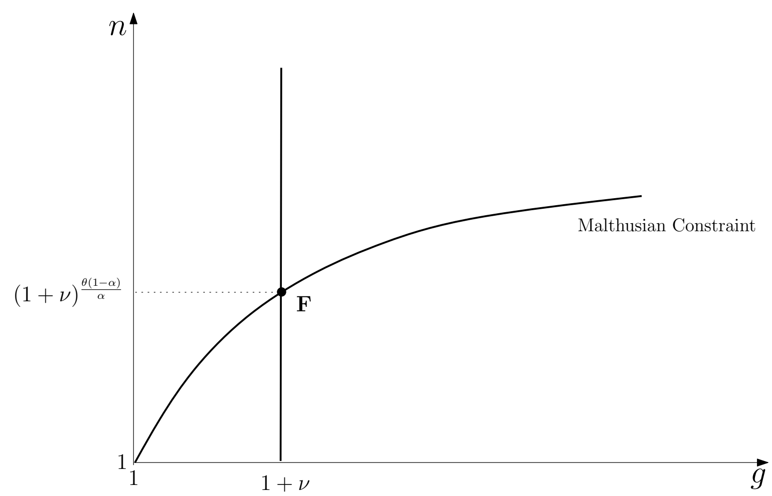

Figure 1: Productivity and Population Growth in the Family (**F**) Equilibrium

Figure 1 represents the determination of the balanced growth path in the family equilibrium. The concave curve represents the Malthusian constraint given by (9).[18] The intersection between this constraint and the line $g = 1 + \nu$ gives the balanced growth equilibrium **F**.

4.3 The Clan

Next, we consider economies where there is communication within an extended family or clan. While many other structures could be considered, the clan has

[18]It is concave if $\theta(1-\alpha) < \alpha$ but does not need to be.

a particular historical significance because of its importance for organizing economic exchange in the major world regions outside Europe. Formally, we consider a setting where all members of a dynasty who share an ancestor o generations back can communicate (here $o = 0$ corresponds to the family equilibrium, $o = 1$ means siblings are connected, and so on). Now consider a potential "clan equilibrium" with the following equilibrium strategy profiles:

- All craftsmen send their children do be trained by all current members of the clan, and parents compensate masters for the apprenticeship by paying each $\delta'(a) - \kappa$ (the marginal cost), where a is the number of apprentices per master.
- All masters put effort into teaching.
- If (off the equilibrium path) a master cheats an apprentice, all current members of the clan punish the master. In addition, if (off the equilibrium path) a craftsman does not participate in required punishment, that craftsman gets punished also.

For example, if $o = 1$, children get trained not just by their parents, but also by their aunts and uncles. For $o = 2$, second-degree relatives serve as masters, and so on.[19]

Along a balanced growth path, the total number of adults belonging to the clan is $(n^{\text{c}})^o$, where n^{c} is the rate of population growth in the balanced growth path. The following proposition summarizes the properties of the balanced growth path in the clan equilibrium.

Proposition 3 (Balanced Growth Path in Clan Equilibrium).
There is threshold $o_{\max} > 0$ such that if $o < o_{\max}$ and if altruism is sufficiently strong (i.e., γ is sufficiently large), there exists a balanced growth path in the clan equilibrium with the following properties:

(a) The number of masters per child m is given by the number of adults in the clan, $m^{\text{c}} = (n^{\text{c}})^o$, and the number of apprentices per master is $a^{\text{c}} = (n^{\text{c}})^{o+1}$.

[19] In reality, if there is sufficient information sharing within the clan, it may not be necessary to receive full training from all clan members. Instead, one could assume that apprentices initially search over the entire group, sample the masters' knowledge, but then spend most of their time learning from the clan member identified to have the lowest h. We adopt the simpler notion of learning equally from all masters to preserve the symmetry that makes the problem tractable.

(b) *The growth factor g_k^{c} of knowledge k is the solution to:*

$$g^{\text{c}} = 1 + \frac{\nu \left(n^{\text{c}}\right)^{o}}{g^{\text{c}} - \nu}. \tag{15}$$

(c) *The growth factor n_k^{c} of population N is given by:*

$$n^{\text{c}} = \left(g^{\text{c}}\right)^{\frac{(1-\alpha)\theta}{\alpha}}.$$

(d) *Income per capita is constant and satisfies:*

$$y^{\text{c}} = \frac{\left(g^{\text{c}}\right)^{\frac{(1-\alpha)\theta}{\alpha}}}{\bar{n}s}.$$

For $o = 0$, the balanced growth path coincides with the family equilibrium, whereas for $o > 0$ knowledge growth, population growth, and income per capita are higher compared to the family equilibrium. The growth g^{c} of knowledge k is increasing in the size of the clan o.

Proof. See Appendix D. ■

Parallel to the family equilibrium, the condition on sufficiently high altruism ensures that parents find it worthwhile to pay for the training of their children.[20] The upper bound $o_{\max}$ on the size of the clan limits productivity growth to a level where the Malthusian feedback is sufficiently strong to generate a balanced growth path with constant income per capita.

The clan equilibrium leads to a higher growth rate compared to the family equilibrium because children learn from more masters. In particular, they benefit not just from the new ideas of their own parent, but also from the new ideas of their aunts, uncles, and other current members of the clan. Thus, new knowledge disseminates more widely compared to the family equilibrium. However, there is still

[20] Another possibility is that altruism is at a level sufficient for parents to want to send their children to some, but not all, available masters. Characterizing the balanced growth path in this case is more complicated, because the selection of which masters to train with is non-trivial. Nevertheless, the basic shortcoming of the clan-based institution, namely that different masters have similar knowledge and so less new can be learned by visiting more of them, would still apply in this type of equilibrium.

no dissemination of knowledge across clans. Equation (15) implies that as long as $\nu > 0$ (there is some innovation), a higher o (larger clans) leads to faster growth. However, if there are no new ideas, $\nu = 0$, the growth rate in the clan equilibrium is zero. Intuitively, in a clan the masters of a given apprentice all trained with the same masters when they were apprentices, which implies that they all started out with the same knowledge. If the masters don't have new ideas of their own, studying with multiple masters does not provide any benefit over studying with only one of them. Hence, knowledge does not accumulate across generations.

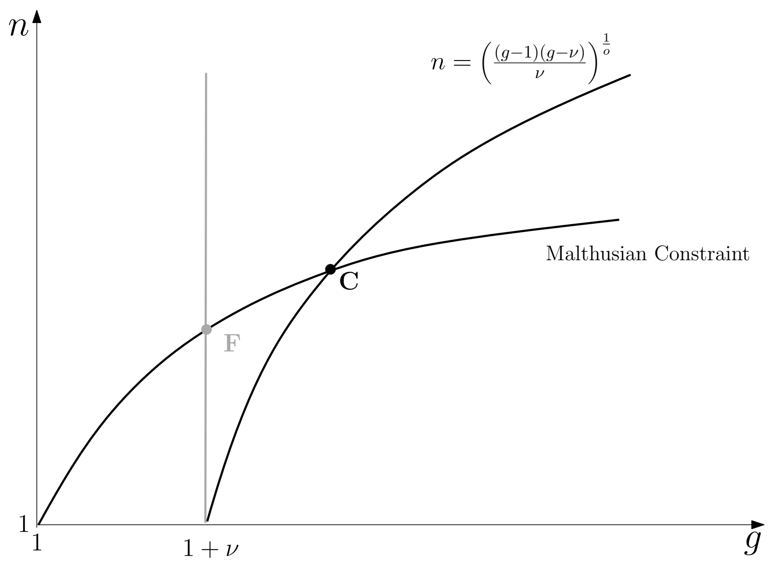

Figure 2: Productivity and Population Growth in the Clan (**C**) Equilibrium

Figure 2 represents the determination of the balanced growth path in the clan equilibrium. In addition to the Malthusian constraint (9), we have drawn the function

$$n = \left(\frac{(g-1)(g-\nu)}{\nu}\right)^{\frac{1}{o}}, \tag{16}$$

which is derived from (15). This function is decreasing from 1 to 0 as g goes from 0 to ν, and increasing from 0 to $+\infty$ as g goes from 1 to ∞. Is it not defined for $g \in (\nu, 1)$. For $g = 1 + \nu$ it is equal to 1. If $o < \bar{o}$, it crosses the Malthusian constraint at some point **C**, which is the clan equilibrium. The positively sloped relationship between population growth and technical progress (16) can be interpreted as follows: When fertility is higher, the clan is bigger, there are more masters and more new ideas. Faster technological progress follows.

4.4 The Market

At the opposite extreme (compared to the family) of enforcement institutions, we now consider outcomes in an economy with formal contract enforcement (as in the usual complete-markets model). All contracts are perfectly and costlessly enforced, so that masters who promise to train apprentices do not shirk. There is a competitive market for apprenticeship. Given market price p for training apprentices, masters decide how many apprentices to train, and parents decide how many masters to pay to train their children. In equilibrium, p adjusts to clear the apprenticeship market.

A craftsman's decision to take on apprentices is a straightforward profit maximization problem. In particular, given price p a master will choose the number of apprentices a to solve:

$$\max_a \{pa + \kappa a - \delta(a)\}.$$

The benefit of taking on apprentices derives from the price p as well as the apprentices' production κ, and the cost is given by $\delta(a)$. Optimization implies that in equilibrium:

$$p = \delta'(a) - \kappa.$$

Now consider parents' choice of the number of masters m that their children should learn from. Given p, parents will choose m to maximize their utility (1):

$$\max_m \mathrm{E}\{-p\ m\ n + \gamma\ I'\},$$

where n is the number of children. Each child's income depends on m, because learning from a larger number of masters increases the expected productivity (and hence income) of the child. The objective function is concave, because as m rises, the probability that an additional master will have higher productivity than all masters already visited declines.

Lemma 1. *The first-order condition for the parent's problem implies:*

$$\delta'(a) - \kappa = \gamma\theta(1-\alpha)\frac{1}{m+\nu}\frac{Y'}{N'}. \tag{17}$$

Proof. See Appendix E. ■

Notice that the decision problem implicitly assumes that the young apprentice gets m independent draws from the distribution of knowledge among the elders, as though the masters were drawn at random. The possibility of independent draws from the knowledge distribution is a key advantage of the market system over the clan system. In a clan, the potential masters have similar knowledge (because they learned from the same "grand" master), and hence the gain from studying with more of them is limited (there is still some gain because of the new ideas generated by masters). Of course, it would be even better to study only with masters known to have superior knowledge. We assume, however, that a master's knowledge can only be assessed by studying with them; hence, choosing masters at random is the best one can do. The market equilibrium gives rise to a unique balanced growth path, which is characterized in the following proposition.

Proposition 4 (Balanced Growth Path in Market Equilibrium).
The unique balanced growth path in the market equilibrium has the following properties:

(a) The number of apprentices per master a^{M} solves (17):

$$\delta'(a^{\mathrm{M}}) - \kappa = \gamma\theta(1-\alpha)\frac{1}{a^{\mathrm{M}}/n^{\mathrm{M}}+\nu}y^{\mathrm{M}}, \tag{18}$$

and the number of masters per child m^{M} is the solution to $m^{\mathrm{M}} = a^{\mathrm{M}}/n^{\mathrm{M}}$.

(b) The growth factor g^{M} of knowledge k is given by:

$$g^{\mathrm{M}} = m^{\mathrm{M}} + \nu.$$

(c) The growth factor n^{M} of population N is given by:

$$n^{\mathrm{M}} = (g^{\mathrm{M}})^{\frac{(1-\alpha)\theta}{\alpha}}.$$

(d) *Income per capita is constant and satisfies:*

$$y^{\mathrm{M}} = \frac{Y}{N} = \frac{(g^{\mathrm{M}})^{\frac{(1-\alpha)\theta}{\alpha}}}{\bar{n}s}.$$

The market equilibrium yields higher growth in productivity and population and higher income per capita than the clan equilibrium and the family equilibrium.

Proof. See Appendix F. ■

To analyze the equilibrium, we can plug the expressions for a^{M}, m^{M}, and y^{M} into (17) to get:

$$\delta'((g^{\mathrm{M}} - \nu)n^{\mathrm{M}}) - \kappa = \gamma\theta(1-\alpha)\frac{1}{g^{\mathrm{M}}}\ \frac{n^{\mathrm{M}}}{\bar{n}s}. \tag{19}$$

This equation describes a relationship between g^{M} and n^{M} which we call the "apprenticeship market," as it is derived from demand for apprenticeship and the equilibrium condition on the apprenticeship market. Equation (19) can be rewritten as:

$$n^{\mathrm{M}} = \frac{\kappa}{\bar{\delta}(g^{\mathrm{M}} - \nu) - \dfrac{\gamma\theta(1-\alpha)}{\bar{s}\bar{n}\ g^{\mathrm{M}}}} \tag{20}$$

This function of g^{M} is plotted in Figure 3. The negative relationship between population growth and the rate of technical progress in (19) can be interpreted as follows. When fertility is higher, the market for apprenticeships is tighter, the equilibrium price of apprenticeship is higher, and parents demand fewer masters. Hence faster population growth is associated with lower productivity growth.

The market equilibrium leads to faster growth than the clan equilibrium because knowledge is disseminated across ancestral boundaries throughout the entire economy. The masters teaching apprentices represent a wider range of knowledge, implying that more can be learned from them.[21] All of this is made possible by having a different enforcement technology of the apprenticeship contracts, relying on courts rather than on punishment by clan members.

[21] The contrast between Clan and market equilibrium is an example of social structure being important for economic outcomes; a similar application to technology diffusion is provided in the recent work by Fogli and Veldkamp (2012).

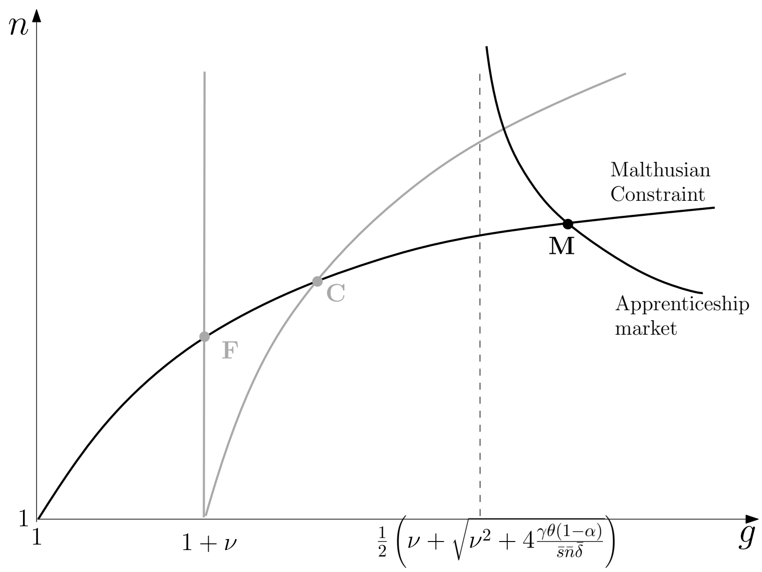

Figure 3: Productivity and Population Growth in the Market (**M**) Equilibrium

4.5 The Guild

Historically, economies did not transition directly from the Family or clan equilibrium to the market equilibrium; rather, there are intermediate stages of semi-formal enforcement through institutions other than the state. In Europe, the key intermediate institution is the guild system, which for centuries regulated apprenticeship and knowledge transmission, at a time when state power was still weak. Guilds had other, anti-competitive features, which motivate a long historical debate on the merits, or lack thereof, of guilds in terms of promoting economic performance. We now provide a formal characterization of a "guild equilibrium" as an intermediate step between the family equilibrium and the market equilibrium.

We envision a guild as an association of all masters involved in the same trade. In the production function (3), the effective labor supply from many different trades is combined with limited substitutability across trades, so that market power can arise. Allowing for heterogeneous labor supply by different trades, the labor income of a craftsman i in trade j is:

$$I_{ij} = q_{ij}\left(1-\alpha\right)\frac{Y}{L}\left(\frac{L_j}{L}\right)^{\frac{1}{\lambda}-1}.$$

Apprentices choose the most attractive trade. In equilibrium, the net benefit of joining as an apprentice is equalized across trades, so that for all j we have:

$$EI'_{ij} - p_j m_j = Eq'_{ij} (1 - \alpha) \frac{Y'}{L'} - pm. \tag{21}$$

Collusion among masters in a given guild leads to social costs and benefits compared to the clan equilibrium. The costs are the usual downsides from limited competition; the guild has an incentive to raise prices and limit entry. Guilds enforced labor market monopsonies, and as a result often limited the number of apprentices that each master was allowed to take on at one time, specifying the number of years each apprentice had to spend with his master, or even specifying long gaps that had to elapse between taking on one apprentice and the next one (Trivellato 2008, p. 212; Kaplan 1981, p. 283). The purpose of these limitations was to limit supply and increase exclusionary rents, which for our purposes means that technological progress is slowed down compared to a market equilibrium. However, the upside is that the guild cuts across different dynasties and thus represents the full range of knowledge in the given trade. If the guild also enforces apprenticeship contracts (in the same fashion as in the clan equilibrium above), there is more scope for knowledge accumulation. Thus, in the absence of strong centralized contract enforcement institutions (i.e., if the Clan and not the Market is the relevant alternative), the guild has a genuinely positive role to play.[22]

Consider the choice of a guild j of setting the price of apprenticeship p_j within the trade, or equivalently, of choosing the number a_j of apprentices per master. The guild maximizes the utility of the masters in the trade. If the guild lowers a_j, the effective supply of craftsmen's labor in trade j in the next generation goes down. Due to limited substitutability across trades, this increases future craftsmen's income in the trade, and thus the price p_j that today's apprentices are willing to pay. Thus, as in a standard monopolistic problem, the guild will raise p_j to a level above the marginal cost of training apprentices. The maximization problem of the guild

[22]This feature provides a contrast between our work and other recent research on the economic role of guilds, such as Desmet and Parente (2014).

can be expressed as:[23]

$$\max_{a_j} \{p_j a_j - \delta(a_j) + \kappa a_j\} \tag{22}$$

subject to:

$$S_j N' m_j = N a_j,$$
$$p_j = \gamma \frac{\partial E I'_{ij}}{\partial m_j},$$
$$E I'_{ij} - p_j m_j = (1-\alpha) \frac{Y'}{N'} - pm.$$

Here S_j is the endogenous fraction of apprentices choosing to join trade j. We have $S_j = 1$ in equilibrium; however, the guild solves its maximization problem taking the behavior of all other trades as given, so that S_j varies with p_j and a_j in the maximization problem of the guild. The second constraint represents the optimal behavior of parents sending their children to trade j (equalizing p_j to the marginal benefit of training with an additional master). The third constraint stems from the mobility of apprentices across trades (from (21)). These two equations represent the two market forces limiting the power of the guild. Notice that Y'/N' is exogenous for the guild j, because each trade is of infinitesimal size.

Lemma 2. *At the symmetric equilibrium, the solution to the maximization problem (22) satisfies:*

$$\delta'(a) - \kappa = \Omega(m)\ \gamma\theta(1-\alpha) \frac{1}{m+\nu} \frac{Y'}{N'} \tag{23}$$

with $\Omega(m) < 1$.

Proof. See Appendix G. ■

Thus, the condition determining equilibrium in the apprenticeship market is of the same form as in the market equilibrium (see Lemma 1), but with the benefit from apprenticeship scaled down by a factor strictly smaller than one. Hence, the extent of apprenticeship (and productivity growth) will be lower compared to the market equilibrium. In the limit where trades become perfect substitutes, $\lambda \to 1$,

[23]For simplicity, and realistically for the European case, we assume that the children of masters in trade j will look for apprenticeship in other trades. This can be rationalized by a small role for "talent" in choosing trades.

we have that $\Omega(m) \to 1$, i.e., guilds have no market power and the problem of the guild leads to the same solution as the market (Lemma 1).

We can now characterize the balanced growth path in the guild equilibrium.

Proposition 5 (Balanced Growth Path in Guild Equilibrium).
The unique balanced growth path in the guild equilibrium has the following properties:

(a) The number of apprentices per master a^{G} *solves (23):*

$$\delta'(a^{\mathsf{G}}) - \kappa = \Omega\left(\frac{a^{\mathsf{G}}}{n^{\mathsf{G}}}\right) \gamma\theta(1-\alpha)\frac{1}{(a^{\mathsf{G}}/n^{\mathsf{G}}+\nu)}y^{\mathsf{G}}, \tag{24}$$

and the number of masters per child m^{G} *is the solution to* $m^{\mathsf{G}} = a^{\mathsf{G}}/n^{\mathsf{G}}$.

(b) The growth factor g^{G} *of knowledge* k *is given by:*

$$g^{\mathsf{G}} = m^{\mathsf{G}} + \nu.$$

(c) The growth factor n^{G} *of population* N *is given by:*

$$n^{\mathsf{G}} = (g^{\mathsf{G}})^{\frac{(1-\alpha)\theta}{\alpha}}.$$

(d) Income per capita is constant and satisfies:

$$y^{\mathsf{G}} = \frac{Y}{N} = \frac{(g^{\mathsf{G}})^{\frac{(1-\alpha)\theta}{\alpha}}}{\bar{n}s}.$$

The guild equilibrium yields lower growth in productivity and population and lower income per capita than the market equilibrium.

Proof. See Appendix H. ■

The guild equilibrium is represented in Figure 4, where the apprenticeship market is described by Equation (24). This relationship is similar apprenticeship market condition in the market equilibrium, but with a shift to the left because of the market power of the guild, represented by the term $\Omega(\cdot)$.

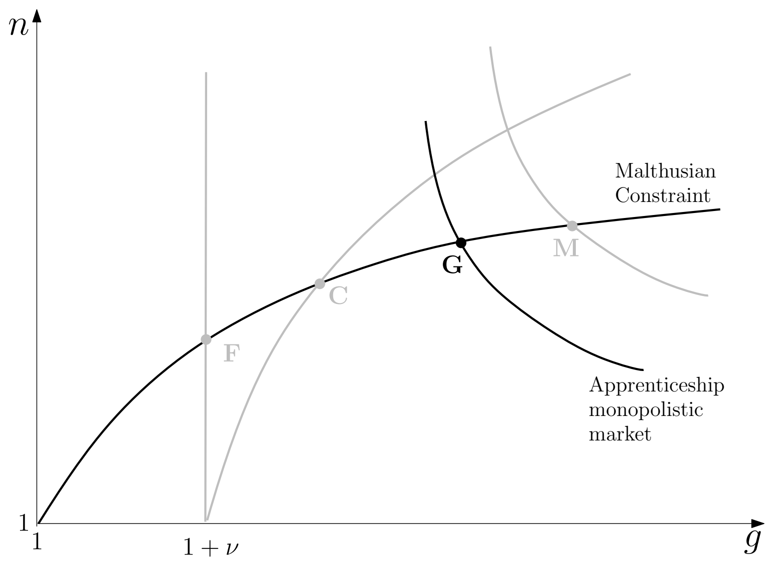

Figure 4: Productivity and Population Growth in the Guild (**G**) Equilibrium

For explaining the rise of European technological supremacy, the key comparison is between the growth performance of the guild equilibrium (which we view as representing Europe for much of the period from the Middle Ages to the Industrial Revolution) and the clan equilibrium (which characterizes other regions such as China, India, and the Middle East). There are forces in both directions; guilds foster growth compared to clans because knowledge can disseminate across ancestral lines, but at the same time the anticompetitive behavior of guilds may limit access to apprenticeship. For this reason, the ranking of growth rates depends on parameters. The guild will lead to faster growth if λ is sufficiently small, because a low λ (close to 1) implies that guilds have little market power, so that the guild equilibrium is close to the market equilibrium. Moreover, the guild also generates faster growth if the rate of innovation ν (i.e., the relative efficiency of new versus existing ideas) is close to zero. In this case, most of growth is due to the dissemination of existing ideas rather than by the generation of new knowledge, and guilds dominate clans in terms of dissemination (recall that the growth rate in the clan equilibrium is zero if $\nu = 0$).

Perhaps the most important comparison is that the guild would always lead to more growth than the clan if the number of masters m were the same in both systems. In the guild, conditional on m, masters are selected in the best possible way (namely as independent draws from the distribution of knowledge, which maximizes the probability that something new can be learned from an additional master).[24] In contrast, in the clan the knowledge of the multiple masters that a given apprentice learns from is necessarily correlated, given that all masters started out with the same initial knowledge available in the clan. Thus, for a given m, in a clan apprentices are exposed to a smaller variety of ideas, and (on average) they learn less. Hence, the only scenario where the clan could generate more growth than the guild is where the market power of guilds is so strong that they would reduce m to well below the level prevalent in the clan. If anything, the historical evidence points in the opposite direction: through the multiple interactions that apprenticeship and journeymanship provided, the European guild system is likely to have been characterized by at least as many learning opportunities as contemporary clan based systems. From the perspective of our model, faster technological progress in Western Europe compared to other world regions would then be the necessary consequence.

4.6 Apprenticeship Institutions and Growth in a Parameterized Economy

We now illustrate our results with a parameterized example of the model economy. We do not formally calibrate the model, but choose parameter values that yield broadly plausible growth rates for the historical period considered. One period (generation) is interpreted as 25 years. We first set $\alpha = 0.8$, $\theta = 0.25$, $\gamma = 0.1$, $\bar{n} = 2$, $s = 7.5$, $o = 3$, $\kappa = 0.02$, and $\lambda = 4$. We then set ν (the relative efficiency of new ideas) so as to reproduce a growth rate of population of 0.86 percent per generation in the family equilibrium, which matches the estimated growth of population between 10000 BCE and 1000 CE in Clark (2007), Table 7.1. This yields $\nu \approx 0.15$. We set the cost of training apprentices such that the number of masters per apprentice m is identical in the clan and guild equilibria, which yields $\bar{\delta} \approx 0.019$. It would be more realistic to allow for a higher m in the guild equilibrium, but by equalizing m across the institutional regimes allows us to isolate the

[24]Note that while the guild may limit access to apprenticeship, it does benefit from allowing for an efficient choice of masters, because this raises the expected benefit of learning from masters and hence the price apprentices are willing to pay.

additional growth in the guild equilibrium compared to the clan equilibrium that arises solely as a consequence of the increased variety in masters' knowledge. Any growth effects due to a higher m in the guild would be in addition to this effect.

	$g-1$	$n-1$	m	Y/N
Family Equilibrium (**F**)	14.7%	0.86%	1	6.724
Clan Equilibrium (**C**)	14.9%	0.87%	1.018	6.725
Guild Equilibrium (**G**)	16.4%	0.96%	1.018	6.730
Market Equilibrium (**M**)	17.7%	1.02%	1.030	6.735

Table 1: Balanced Growth Paths for Different Apprenticeship Institutions in Parameterized Economy ($g-1$ is productivity growth; $n-1$ is population growth, and also growth in total output; m is the number of masters each apprentice learns from; Y/N is income per capita)

Table 1 displays the balanced growth rates for knowledge k and population N under each apprenticeship institution, together with the number of masters per apprentice m and income per adult Y/N. Notice that since Y/N is constant in the balanced growth path, the growth rate of total output Y is equal to the growth rate of population N. We find that productivity growth, population growth, output growth, and income per capita are all increasing as we proceed from family to clan, guild, and ultimately the market. Unlike in Malthusian theories with fixed productivity growth, in our model the long-run level of income per capita is endogenous, albeit constant over time within a given regime. An economy undergoing a gradual transition from one regime to the next would experience changes income per capita.

In terms of the growth performance of the different institutional regimes, we notice that the growth advantage of the clan compared to the family is small. In contrast, guild and market yield substantially higher growth rates than either family or clan. The market yields the highest growth rate. However, moving from clan to guild already generates 60 percent of the growth rate differential between clan and market, despite the fact that (by our choice of parameterization) the number of apprentices is identical in the clan and guild equilibria.[25] Hence, the easier dissemination of knowledge for a given m provided by a system that is not constrained by bloodlines accounts for the majority of the overall growth effect of better institutions. Moving from guild to market yields an additional growth effect through a higher m (because in the market system, guilds are not able to restrict access to apprenticeship).

5 The Rise of Europe's Technological Primacy

One of the central questions about pre-industrial economic growth is why Western Europe pulled ahead of other world regions in the centuries leading up to the industrial revolution. Over time, Western Europe achieved technological primacy over the previous leaders, and on the eve of the industrial revolution living standards where substantially higher than elsewhere. The more rapid growth in Western Europe during this period has been identified by a number of researchers as a key precondition for successful industrialization.[26]

As the preceding analysis makes clear, in our view apprenticeship institutions that promoted the dissemination of knowledge lie at the heart of Western Europe's success. However, this still leaves open the question of why Europe adopted superior institutions such as the guild and ultimately the market, whereas other regions failed to do so. To deal with this issue, we now consider mechanisms that explain transitions over time between different apprenticeship institutions.

As already hinted at earlier, we view most of the pre-industrial world as characterized by the clan equilibrium. The contrast between Asia and Europe in systems of knowledge transmission is emphasized by van Zanden (2009): "We can distinguish two different ways to organize such training: in large parts of the world the family or the clan played a central role, and skills were transferred from fathers to sons or other members of the (extended) family. In fact, in parts of Asia, being

[25]Notice that in all cases, the number of masters per apprentice m is either equal to one or close to one. While we treat m in the model as a continuous variable, an alternative interpretation is that even in the more efficient instititional regimes, relatively few workers benefit from learning from multiple masters, generating a low average m in the aggregate.

[26]The rise of Europe in the centuries leading up to the industrial revolution is also the subject of two recent papers by Voigtländer and Voth (2013a, 2013b). However, they emphasize demographic changes, whereas our analysis is about differences in productivity growth.

a craftsman was largely hereditary. ... In contrast to the relatively closed systems in which the family played a central role, Western Europe had a formal system of apprenticeship—organized by guilds or similar institutions—and in principle open to all." Similarly, van Zanden and Prak (2013a) write: "In China, training was provided by relatives, and hence a narrow group of experts, instead of the much wider training opportunities provided by many European guilds." The importance of clans in organizing economic life in China has recently been emphasized in the work of Avner Greif with different coauthors (e.g., Greif and Tabellini 2010, Greif, Iyigun, and Sasson 2012, and Greif and Iyigun 2013). Regarding India, Roy (2013) writes: "Two stylized models of training-cum-regulation, however, can be distinguished. One of these involved relatively open recruitment of apprentices, and regulation by means of an indenture contract recognized by the state or guilds. The other system involved recruitment restricted by ethnicity and kinship, and regulation of younger workers by their intimate relations. ... The latter exercised a wider scope in India." Finally, Kumar and Matsusaka (2009) report an array of historical evidence documenting the pre-industrial importance of kinship networks in China, India, and the Islamic world compared to Europe.

In contrast, we view Europe as an economy that started out closer to the family equilibrium, but then adopted more efficient, market-based institutions in the centuries leading up to the industrial revolution, thereby overtaking the rest of the world. Following Greif and Tabellini (2010), the underlying reasons why family, not clan, dominated in Europe are related to religion. In Europe the Christian church actively discouraged practices that sustain kinship groups, and by the ninth century the nuclear family predominated. Mitterauer (2010) describes two signs of the emergence of the nuclear family: the distinction between paternal and maternal relatives disappeared from the Romance languages by 600 CE (and from other European languages soon after); and spiritual kinships analogous to blood kinships were established (such as godmother and godfather).

Once the initial equilibrium has been selected, the future path each of economy will vary, depending on which one will adopt new institutions first (path dependency). The central question becomes therefore is why family-based Europe adopted superior institutions over time, whereas other clan-based regions did not.

One potential explanation is that it is precisely the fact that Europe started out in the low-growth family equilibrium that fostered the more rapid adoption of superior institutions. If adopting the guild or market systems is costly, the incentive for adoption depends on the the performance of the existing institutions. In our view, other world regions had less to gain from adopting new institutions, given that the clan-based system performed well for most purposes.

To formalize this possibility, consider the option of adopting the guild system at fixed per-family cost of $\mu(N)$. We let depend this cost on the density of population, with $\mu'(N) < 0$ reflecting the idea that the adoption of guilds is cheaper when density is high (in conformity with the fact that the incidence of guilds increases with population density, see de Munck, Lourens, and Lucassen 2006). This cost $\mu(N)$ can either be seen as an aggregate cost to set up guilds or courts, or, linked to an individual decision, i.e., the cost of moving from a small town to a larger city where contract enforcement institutions are in place.

Formally, we need to compare the utility of the parents from keeping the current system, $u^{\text{F}\to\text{F}}$ and $u^{\text{C}\to\text{C}}$ with the one of adopting the guild based apprenticeship, $u^{\text{F}\to\text{G}}$ and $u^{\text{C}\to\text{G}}$. Let us consider two economies having the same population N_0 and knowledge k_0, one in the family system, the other in the clan system. The distribution of income is thus the same in these two economies, as is mean income y_0 and the number of children n_0. Adults have to decide whether to pay the cost μ to adopt the guild. If the guild is adopted, the equilibrium price of apprenticeship will be p_0, the number of apprentices per master will be a_0, and the number of masters teaching one apprentice will be m_0. The income of the children under the guild system will be given by: $y^{\text{F}\to\text{G}} = y^{\text{C}\to\text{G}} \equiv y_1^{\text{G}}$. Let us now write the utility in the four cases:

$$\begin{aligned}
u^{\text{F}\to\text{F}} &= y_0 + \gamma n_0 y_1^{\text{F}} + \kappa n_0 - \delta(n_0) \\
u^{\text{F}\to\text{G}} &= y_0 + \gamma n_0 y_1^{\text{G}} + \kappa a_0 - \delta(a_0) - \mu(N_0) \\
u^{\text{C}\to\text{C}} &= y_0 + \gamma n_0 y_1^{\text{C}} + \kappa (n_0)^{o+1} - \delta((n_0)^{o+1}) \\
u^{\text{C}\to\text{G}} &= y_0 + \gamma n_0 y_1^{\text{G}} + \kappa a_0 - \delta(a_0) - \mu(N_0)
\end{aligned}$$

The following propositions gives the main result.

Proposition 6 (Transition from Family and Clan to Guild).
Consider two economies with the same initial knowledge and population, one in the family equilibrium, and one in the clan equilibrium. If $\lim_{N\to 0}\mu(N) = \infty$ *and* $\lim_{N\to\infty}\mu(N) = 0$*, there exist population thresholds* $\underline{N}$ *and* $\overline{N}$*, with* $\underline{N} < \overline{N}$*, such that:*

- ▷ *if* $N_0 < \underline{N}$*, none of the economies adopt the guild institution.*
- ▷ *if* $\underline{N} \leq N_0 \leq \overline{N}$*, only the economy in the family equilibrium adopts the guild institution.*
- ▷ *if* $\overline{N} < N_0$*, both economies adopt the guild institution.*

Proof. See Appendix I. ■

Given equal populations, the incentive to pay the fixed cost will be lower when the initial economic system is more successful, i.e., a clan-based economy will be less likely to adopt than a family-based economy.[27]

Take now two hypothetical economies starting with the same low level of population. Suppose that one of them starts in the family equilibrium, whereas in the other the clan equilibrium prevails. In both economies, population is low and the guild equilibrium is not adopted, but there is still some technical progress and population grows. Given Proposition 3, population growth is higher in the clan economy. The question is which economy will first reach the population threshold that makes adopting the market optimal. The family economy has a lower threshold value, but, as it grows more slowly, it is not clear that it will adopt the guild first. A possible trajectory is the one shown in Figure 5. Here, the family economy adopts the guild earlier, at date t_1, which allows it to catchup and overtake the c economy. Later on, the clan economy may or may not reach its own threshold above which it is optimal to adopt the guild, depending on the properties of the cost function $\mu(\cdot)$, and in particular how it behaves when population becomes large. In the picture, the threshold is reached at some date t_2.

Proposition 7 (Clan as an Absorbing State).

[27] This explanation applies earlier work by Avner Greif (see in particular Greif 1993 and Greif 1994) on institutional change to the issue of human capital and knowledge transmission.

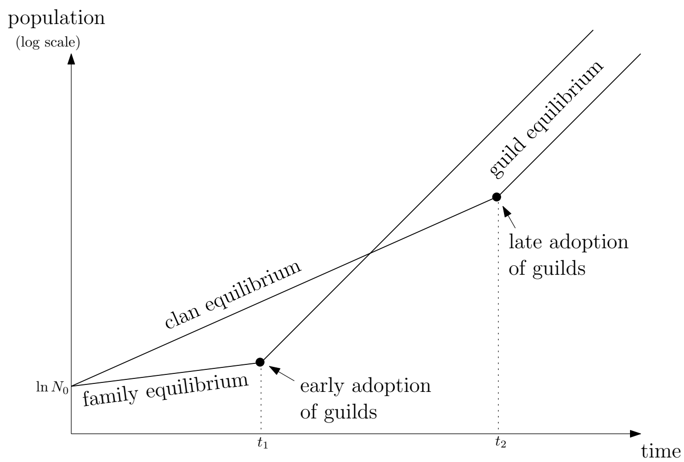

Figure 5: Possible Dynamics of Two Economies Starting from the Same N_0

If

$$\lim_{N\to\infty}\mu(N) > \gamma n^{\mathtt{C}}(y_1^{\mathtt{G}} - y^{\mathtt{C}}) + \kappa(a_1^{\mathtt{G}} - (n^{\mathtt{C}})^{o+1}) - \delta(a_1^{\mathtt{G}}) + \delta((n^{\mathtt{C}})^{o+1}),$$

the economy in the clan equilibrium never adopts the guild.

Here the values of $n^{\mathtt{C}}$ and $y^{\mathtt{C}}$ are defined in Proposition 3, and $y_1^{\mathtt{G}}$, $a_1^{\mathtt{G}}$ are values from the guild equilibrium after one period, starting from the clan balanced growth path as the initial condition. The proposition holds because, in a Malthusian context, income per person $y^{\mathtt{C}}$ and $y_1^{\mathtt{G}}$ remains bounded.

In reality, complementary mechanisms are likely to have also contributed to the failure of clan-based economies to adopt more efficient apprenticeship institutions. For example, the clan-based organization had many advantages over the family other than faster knowledge growth; indeed, most of economic and social life was organized around the clan. Hence, for a successful clan-based economy, the cost of giving up the existing system of social organization (in favor of the guild system) is likely to have been much higher that what our stylized model suggests. Another dimension is that the dominance of the nuclear family in Europe created a need early on for organizations that cut across family lines. Guilds are such an organization that is independent of families, but there are many antecedents that had a similar legal status (such as monasteries or independent cities). Hence, earlier institutional developments may have made the adoption of guilds in Europe much cheaper compared to clan-based societies.

6 Conclusion

The aim of this paper was to examine the sources of productivity growth in pre-industrial societies, with the specific objective to explain the relative ascendency of Western Europe in the centuries leading up to industrialization. We developed a model of person-to-person exchanges of ideas, and argued that apprenticeship institutions that regulate the transmission of tacit knowledge between generations are central for understanding the performance of pre-industrial economies.

In our analysis, we have put the spotlight on differences across institutions in the dissemination of knowledge. Of course, a second channel of productivity growth is innovation, i.e., the creation of entirely new knowledge. Our analysis does allow for innovation in the form of new ideas, but we have held this aspect of productivity growth constant across institutions. A natural extension of our work would be to examine how the institutional differences that we have identified as driving differences in the dissemination of knowledge may affect incentives for original innovation. We plan to address this issue in future research.

References

Aghion, Philippe, and Peter Howitt. 1992. "A Model of Growth through Creative Destruction." *Econometrica* 60 (2): 323–51.

Alvarez, Fernando E., Francisco J. Buera, and Robert E. Lucas, Jr. 2008. "Models of Idea Flows." NBER Working Paper 14135.

———. 2013. "Idea Flows, Economic Growth, and Trade." NBER Working Paper 19667.

Ashraf, Quamrul, and Oded Galor. 2011. "Dynamics and Stagnation in the Malthusian Epoch." *American Economic Review* 101 (5): 2003–41.

Bar, Michael, and Oksana Leukhina. 2010. "The role of mortality in the transmission of knowledge." *Journal of Economic Growth* 15 (4): 291–321 (December).

Belfanti, Carlo Marco. 2004. "Guilds, Patents, and the Circulation of Technical Knowledge." *Technology and Culture* 45 (3): 569–89 (July).

Ben Zeev, Nadav, Joel Mokyr, and Karine Van Der Beek. 2015. "Flexible Supply of Apprenticeship in the British Industrial Revolution." unpublished ms.

Brooks, Christopher. 1994. "Apprenticeship, Social Mobility and the Middling Sort, 1550–1800." In *The Middling Sort of People: Culture, Society and Politics in England, 1550-1800*, edited by Jonathan Barry and Christopher Brooks. London: Palgrave Macmillan.

Clark, Gregory. 2007. *A farewell to alms: a brief economic history of the world*. Princeton: Princeton University Press.

Cowan, Robin, and Dominique Foray. 1997. "The Economics of Codification and the Diffusion of Knowledge." *Industrial and Corporate Change* 6 (3): 595–622.

Davids, Karel. 2003. "Guilds, Guildmen and Technological Innovation in Early Modern Europe: The Case of the Dutch Republic." Economy and Society in the Low Countries Working Papers, No. 2.

———. 2007. "Apprenticeship and Guild Control in the Netherlands, c. 1450-1800." In *Learning on the Shop Floor: Historical Perspectives on Apprenticeship*, edited by Steven L. Kaplan Bert De Munck and Hugo Soly, 65–84. New York: Berghahn Books.

———. 2013. "Moving Machine-makers: Circulation of Knowledge on Machine-building in China and Europe between c. 1400 and the Early Nineteenth century." Chapter 6 of *Technology, Skills and the Pre-Modern Economy in the East and the West*, edited by Jan L. van Zanden and Marten Prak. Boston: Brill.

Davies, Margaret Gay. 1956. *The Enforcement of English Apprenticeship: a Study in Applied Mercantilism, 1563-1642*. Cambridge, MA: Harvard University Press.

de la Croix, David, and Omar Licandro. 2015. "The Longevity of Famous People from Hammurabi to Einstein." *Journal of Economic Growth* 20:263–303.

de Munck, Bert, Piet Lourens, and Jan Lucassen. 2006. "The establishement and distribution of crafts guilds in the Low Countries 1000-1800." Chapter 2 of *Crafts Guilds in the Early Modern Low Countries - Work, Power and Representation*, edited by Maarten Prak, Catharina Lis, Jan Lucassen, and Hugo Soly. Aldershot: Ashgate.

De Munck, Bert, and Hugo Soly. 2007. "Learning on the Shop Floor in Historical Perspective." In *Learning on the Shop Floor: Historical Perspectives on Apprenticeship*, edited by Steven L. Kaplan Bert De Munck and Hugo Soly, 3–32. New York: Berghahn Books.

Desmet, Klaus, and Stephen L. Parente. 2014. "Resistance to Technology Adoption: The Rise and Decline of Guilds." *Review of Economic Dynamics* 17 (3): 437–58.

Dunlop, O. Jocelyn. 1911. "Some aspects of early English apprenticeship." *Transactions of the Royal Historical Society, Third Series* 5:193–208.

———. 1912. *English Apprenticeship and Child Labor*. New York: MacMillan.

Eaton, Jonathan, and Samuel Kortum. 1999. "International Technology Diffusion: Theory and Measurement." *International Economic Review* 40 (3): 537–70.

Epstein, Stephan R. 2008. "Craft Guilds, the Theory of the Firm, and the European Economy, 1400-1800." In *Guilds, Innovation and the European Economy, 1400-1800*, edited by Stephan R. Epstein and Maarten Prak, 52–80. Cambridge: Cambridge University Press.

Epstein, Stephan R. 2013. "Transferring Technical Knowledge and Innovating in Europe, c. 1200-c. 1800." Chapter 1 of *Technology, Skills and the Pre-Modern Economy in the East and the West*, edited by Jan L. van Zanden and Marten Prak. Boston: Brill.

Epstein, Stephan R., and Maarten Prak. 2008. "Introduction: Guilds, Innovation and the European Economy, 1400-1800." In *Guilds, Innovation and the European Economy, 1400-1800*, edited by Stephan R. Epstein and Maarten Prak, 1–24. Cambridge: Cambridge University Press.

Fogli, Alessandra, and Laura Veldkamp. 2012. "Germs, Social Networks and Growth." NBER Working Paper 18470.

Foray, Dominique. 2004. *The Economics of Knowledge*. Cambridge, MA: MIT Press.

Galor, Oded. 2011. *Unified Growth Theory*. Princeton: Princeton University Press.

Galor, Oded, and David N. Weil. 2000. "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond." *American Economic Review* 90 (4): 806–28.

Greif, Avner. 1993. "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition." *American Economic Review* 83 (3): 525–48.

———. 1994. "Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies." *Journal of Political Economy* 102 (5): 912–50.

———. 2006. "Family Structure, Institutions, and Growth: The Origins and Implications of Western Corporations." *The American Economic Review* 96 (2): 308–312.

Greif, Avner, and Murat Iyigun. 2013. "Social Organizations, Violence, and Modern Growth." *American Economic Review* 103 (3): 534–38.

Greif, Avner, Murat Iyigun, and Diego Sasson. 2012. "Social Institutions and Economic Growth: Why England and Not China Became the First Modern Economy." Unpublished Working Paper, University of Colorado.

Greif, Avner, and Guido Tabellini. 2010. "Cultural and Institutional Bifurcation: China and Europe Compared." *American Economic Review* 100 (2): 135–40.

———. 2015. "The Clan and the Corporation: Sustaining Cooperation in China and Europe." Unpublished manuscript, Bocconi University.

Humphries, Jane. 2003. "English Apprenticeship: A Neglected Factor in the First Industrial Revolution." Chapter 3 of *The Economic Future in Historical Perspective*, edited by Paul A. David and Mark Thomas. Oxford: Oxford University Press.

———. 2010. *Childhood and Child Labour in the British Industrial Revolution*. Cambridge Studies in Economic History - Second Series. Cambridge: Cambridge University Press.

Jones, Charles I. 2001. "Was an Industrial Revolution Inevitable? Economic Growth Over the Very Long Run." *The B.E. Journal of Macroeconomics* 1 (2): 1–45 (August).

Jovanovic, Boyan, and Rafael Rob. 1989. "The Growth and Diffusion of Knowledge." *The Review of Economic Studies* 56 (4): 569–82.

Kaplan, Steven L. 1981. "The luxury guilds in Paris in the eighteenth century." *Francia* 9:257–98.

Kelly, Morgan, Joel Mokyr, and Cormac Ó Gráda. 2014. "Precocious Albion: A New Interpretation of the British Industrial Revolution." *Annual Review of Economics* 6 (1): 363–389 (08).

Kelly, Morgan, and Cormac Ó Gráda. 2012. "The Preventive Check in Medieval and Preindustrial England." *The Journal of Economic History* 72 (04): 1015–1035 (December).

———. 2014. "Living standards and mortality since the middle ages." *Economic History Review* 67 (2): 358–381 (05).

König, Michael, Jan Lorenz, and Fabrizio Zilibotti. 2015. "Innovation vs. Imitation and the Evolution of Productivity Distributions." Forthcoming, Theoretical Economics.

Kortum, Samuel S. 1997. "Research, Patenting, and Technological Change." *Econometrica* 65 (6): 1389–419.

Kremer, Michael. 1993. "Population Growth and Technological Change: One Million B.C. to 1900." *Quarterly Journal of Economics* 108 (3): 681–716.

Kumar, Krishna B., and John G. Matsusaka. 2009. "From families to formal contracts: An approach to development." *Journal of Development Economics* 90 (1): 106–119.

Lane, Joan. 1996. *Apprenticeship in England 1600-1914*. Boulder: Westview Press.

Leunig, Tim, Chris Minns, and Patrick Wallis. 2011. "Networks in the Premodern Economy: The Market for London Apprenticeships, 1600-1749." *The Journal of Economic History* 71 (6): 413–443.

Lis, Catharina, Hugo Soly, and Lee Mitzman. 1994. ""An Irresistible Phalanx": Journeymen Associations in Western Europe, 1300-1800." *International Review of Social History* 39 (8): 11–52.

Long, Pamela O. 2011. *Artisan/Practitioners and the Rise of the New Sciences, 1400-160*. Cornvallis: Oregon State University Press.

Lucas, Jr., Robert E. 2009. "Ideas and Growth." *Economica* 76 (301): 1–19.

Lucas, Jr., Robert E., and Benjamin Moll. 2014. "Knowledge Growth and the Allocation of Time." *Journal of Political Economy* 122 (1): 1–51.

Lucassen, Jan, Tine de Moor, and Jan Luiten van Zanden. 2008. "The Return of the Guilds: Towards a Global History of the Guilds in Pre-Industrial Times." In *The Return of the Guilds*, edited by Jan Lucassen, Tine de Moor, and Jan Luiten van Zanden, 5–18. Cambridge: Cambridge University Press.

Luttmer, Erzo G. J. 2007. "Selection, Growth, and the Size Distribution of Firms." *The Quarterly Journal of Economics* 122 (3): 1103–1144.

———. 2015. "An Assignment Model of Knowledge Diffusion and Income Inequality." Unpublished Manuscript, University of Minnesota.

Minns, Chris, and Patrick Wallis. 2013. "The price of human capital in a pre-industrial economy: Premiums and apprenticeship contracts in 18th century England." *Explorations in Economic History* 50 (3): 335–350.

Mitterauer, Michael. 2010. *Why Europe?: The Medieval Origins of Its Special Path*. Chicago: University Of Chicago Press.

Mokyr, Joel. 2002. *The gift of Athena. Historical Origins of the Knowledge Economy*. Princeton: Princeton University press.

———. 2008. "The Institutional Origins of the Industrial Revolution." In *Institutions and Economic Performance*, edited by Elhanan Helpman, 64–119. Cambridge MA: Harvard University Press.

Moll-Murata, Christine. 2013. "Guilds and Apprenticeship in China and Europe: The Jingdezhen and European Ceramics Industries." Chapter 7 of *Technology, Skills and the Pre-Modern Economy in the East and the West*, edited by Jan Luiten van Zanden and Marten Prak. Boston: Brill.

Ogilvie, Sheilagh. 2004. "Guilds, Efficiency and Social Capital: Evidence from German Proto-industry." *Economic History Review* 57 (2): 286–333.

———. 2014. "The Economics of Guilds." *Journal of Economic Perspectives* 28 (4): 169–92 (Fall).

———. 2016. "The European Guilds: an Economic Assessment." Unpublished book ms.

Perla, Jesse, and Christopher Tonetti. 2014. "Equilibrium Imitation and Growth." *Journal of Political Economy* 122 (1): 52–76.

Prak, Maarten. 2013. "Mega-structures of the Middle Ages: The Construction of Religious Buildings in Europe and Asia, c.1000-1500." Chapter 4 of *Technology, Skills and the Pre-Modern Economy in the East and the West*, edited by Jan Luiten van Zanden and Marten Prak. Boston: Brill.

Reith, Reinhold. 2007. "Apprentices in the German and Austrian Crafts in Early Modern Times: Apprentices as Wage Earners?" In *Learning on the Shop Floor: Historical Perspectives on Apprenticeship*, edited by Steven L. Kaplan Bert De Munck and Hugo Soly, 179–99. New York: Berghahn Books.

Robert, Leeson. 1979. *Travelling Brothers*. London: George Allen & Unwin.

Roberts, Lissa, and Simon Schaffer. 2007. "Preface." In *The Mindful Hand: Inquiry and Invention from the late Renaissance to early Industrialization*, edited by Simon Schaffer Lissa Roberts and Peter Dear, xiii–xxvii. Amsterdam: Royal Netherlands Academy of Arts and Sciences.

Romer, Paul M. 1990. "Endogenous Technological Change." *Journal of Political Economy* 98 (5): S71–S102.

Roy, Tirthankar. 2013. "Apprenticeship and Industrialization in India, 1600-1930." Chapter 2 of *Technology, Skills and the Pre-Modern Economy in the East and the West*, edited by Jan Luiten van Zanden and Marten Prak. Boston: Brill.

Rushton, Peter. 1991. "The Matter in Variance: Adolescents and Domestic Conflict in the Pre-Industrial Economy of Northeast England, 1600-1800." *Journal of Social History* 25 (1): 89–107.

Smith, Steven R. 1973. "The London Apprentices as Seventeenth-Century Adolescents." *Past & Present* 61:149–161.

Spagnolo, Giancarlo. 1999. "Social relations and cooperation in organizations." *Journal of Economic Behavior & Organization* 38 (1): 1–25 (January).

Stabel, Peter. 2007. "Social Mobility and Apprenticeship in Late Medieval Flanders." In *Learning on the Shop Floor: Historical Perspectives on Apprenticeship*, edited by Steven L. Kaplan Bert De Munck and Hugo Soly, 158–78. New York: Berghahn Books.

Steffens, Sven. 2001. "Le métier volé: transmission des savoir-faire et socialization dans les métiers qualifiés ay XIXe siècle." *Revue du Nord* 15:121–135.

Steidl, Annemarie. 2007. "Silk Weaver and Purse Maker Apprentices in Eighteenth and Nineteenth Century Vienna." In *Learning on the Shop Floor: Historical Perspectives on Apprenticeship*, edited by Bert De Munck, Steven L. Kaplan, and Hugo Soly, 133–57. New York: Berghahn Books.

Trivellato, Francesca. 2008. "Guilds, Technology, and Economic Change in early modern Venice." In *Guilds, Innovation and the European Economy, 1400-1800*, edited by Stephan R. Epstein and Maarten Prak, 199–231. Cambridge: Cambridge University Press.

van Zanden, Jan Luiten. 2009. *The Long Road to the Industrial Revolution: the European Economy in a Global Perspective, 1000-1800*. Boston: Brill.

van Zanden, Jan Luiten, and Marten Prak. 2013a. "Technology and Human Capital Formation in the East and West Before the Industrial Revolution." Chapter 1 of *Technology, Skills and the Pre-Modern Economy in the East and the West*, edited by Jan Luiten van Zanden and Marten Prak. Boston: Brill.

———. 2013b. *Technology, Skills and the Pre-Modern Economy in the East and the West*. Boston MA: Brill.

Voigtländer, Nico, and Hans-Joachim Voth. 2013a. "How the West "Invented" Fertility Restriction." *American Economic Review* 103 (6): 2227–64.

———. 2013b. "The Three Horsemen of Riches: Plague, War, and Urbanization in Early Modern Europe." *The Review of Economic Studies* 80 (2): 774–811.

Vollrath, Dietrich. 2011. "The agricultural basis of comparative development." *Journal of Economic Growth* 16 (4): 343–370.

Wallis, Patrick. 2008. "Apprenticeship and Training in Premodern England." *The Journal of Economic History* 68 (3): 832–861.

A Extension with Farmers

In this section, we sketch how the model can be extended by including farm labor as a separate input in production. This extension addresses the concern that in the pre-industrial era, most people were engaged in food production, whereas craftspeople made up a smaller fraction of the population.

The single consumption good (which we interpret as a composite of food and manufactured goods) is produced with a Cobb-Douglas production function with constant return to scale that uses land X, farm labor N_f, and effective craftsmen's labor L as inputs:

$$Y = (N_f)^{1-\alpha-\beta} L^{\beta} X^{\alpha},$$

with α, $\beta \in (0, 1)$. The total amount of land is normalized to one, $X = 1$, and land is owned by farmers.

The aggregate state variables are now three: N_f (population of farmers), N_m (population of craftsmen), and k. Let $N = N_m + N_f$ be the total number of adults. Farmers and craftsmen have the same survival rate and that there is no intergenerational mobility across occupations. Hence, the laws of motion for population are:

$$N'_m = n\, N_m, \quad N'_f = n\, N_f, \quad N' = N'_m + N'_f = n\, N.$$

As a consequence, the share of both groups in the total population is constant. We define ρ as

$$\rho = \frac{N_m}{N}.$$

The assumptions of equal population growth and no occupational mobility are made for simplicity. In reality, it is well known that in pre-industrial times cities (where craftsmen were concentrated) experienced much higher mortality than the countryside, so that there was net migration into cities. Allowing for such rural-urban migration could be accommodated in our framework and would leave the main results intact, but it would come at the cost of complicating the analysis. Given that our focus is on knowledge transmission rather than urbanization, we abstract from such features here.

Given those two changes to the specification, the rest of analysis carries on. Two new parameters are involved. The market equilibrium condition (18) becomes:

$$\delta'(a^{\mathrm{M}}) - \kappa = \frac{\gamma\theta\beta}{\rho} \frac{1}{a^{\mathrm{M}}/n^{\mathrm{M}} + \nu} y^{\mathrm{M}}$$

It becomes easier to analyze the role of parameter α, as now, this one only plays a role in the Malthusian constraint. A low α corresponds to labor-intensive agriculture. In this case returns to population size decrease at a lower rate, and hence an increase in productivity growth leads to a larger shift in population growth and income per capita compared to the case of a large land share.[28] It modifies the incentives to move to another mode of organizing apprenticeship, as detailed in Proposition 8.

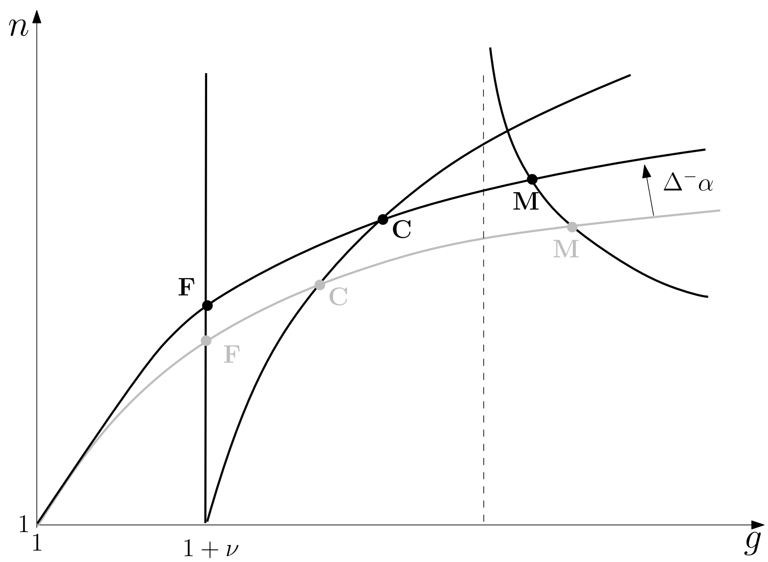

Figure 6: Effect of labor intensive agriculture

Proposition 8 (Effect of Labor-Intensive Agriculture).
If Agriculture is labor intensive (low α), the long-term gains in growth from moving from family to market, $g^{\mathrm{M}} - g^{\mathrm{F}}$, and from clan to market, $g^{\mathrm{M}} - g^{\mathrm{C}}$, are reduced. But the gain from moving from family to clan, $g^{\mathrm{C}} - g^{\mathrm{F}}$, is increased.

Figure 6 shows the result graphically. An implication of this proposition is that a country with labor intensive agriculture has more incentives to adopt the clan than the family.

[28] Vollrath (2011) argues that agriculture in pre-industrial China was more labor intensive compared to Europe (due to the possibility of multiple crops per year, wet-field rice production etc.), and that this accounts for part of observed differences in living standards.

B Extension with the "Leonardo da Vinci" Assumption

In this section, we allow for the possibility that advanced techniques may be "ahead of their time," in the sense that their full productive value can only be realized at a higher state of aggregate knowledge. For example, Leonardo da Vinci invented a number of machines and devices that could be successfully built only centuries later. To capture this feature in a simple way, we assume that each craftsman has a potential output which is linked to own knowledge h_i, but that the craftsman may be constrained by the average state of knowledge in the economy. Specifically, the potential output $\bar{q}_i$ of a craftsman with knowledge h_i is given by:

$$\bar{q}_i = h_i^{-\theta}. \tag{25}$$

The actual output q_i of a craftsman cannot exceed the average potential output $\bar{q} = \mathrm{E}(\bar{q}_i)$ in the economy, so that:

$$q_i = \min\{\bar{q}_i, \bar{q}\}. \tag{26}$$

The assumption that individual productivity is constrained by aggregate knowledge is not essential to any of our results. Still, without this assumption in any period there would be some masters with arbitrarily high output, which is implausible. With the constraint, a good part of the knowledge in the economy is latent knowledge that will unfold its full potential only in later generations. This closely relates to Mokyr (2002)'s argument that the growth of productivity is constrained by the epistemic base on which a technique rests. The more people who use a technique understand the science behind it they understand, the broader the base. According to Mokyr (2002), this basis was very narrow in pre-modern Europe and gradually became wider.

Figure 7 illustrates the distribution of actual output q_i among craftsmen for two values of knowledge k, where the dashed line corresponds to a higher state of knowledge. The kinks in the distribution functions represent the points above which potential productivity is constrained by the average knowledge in the society. Because of the specific shape of the exponential distribution, the share of craftsmen who are constrained by the average knowledge is constant (and given by $1/e$).

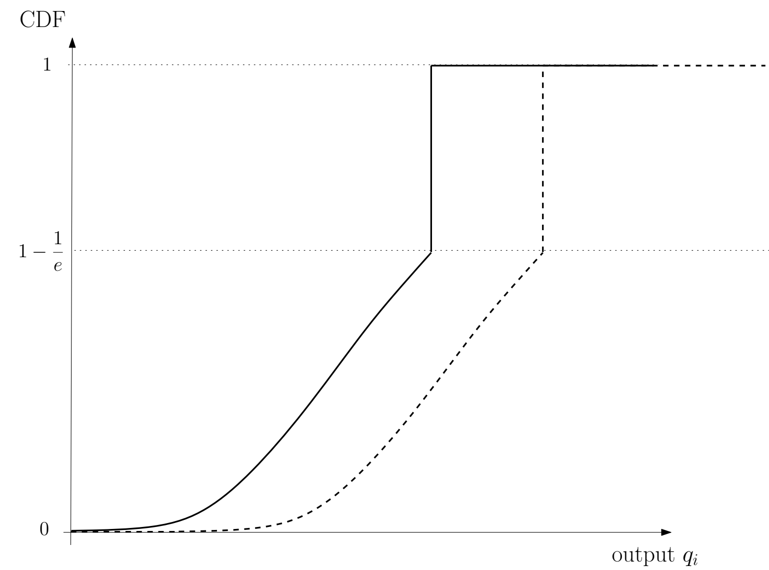

Figure 7: Distribution Function of Productivity at time t (solid line) and $t' > t$ (dashed line)

Given the exponential distribution for h_i and (25), potential output $\bar{q}_i$ follows a Fréchet distribution with shape parameter $1/k$ and scale parameter $1/\theta$.

We can now express the total supply of effective labor by craftsmen as a function of state variables. The average potential output across craftsmen is given by:

$$\bar{q} = \int_0^\infty h_i^{-\theta} \left(k \exp(-kh_i)\right) dh_i = \int_0^\infty k^\theta (kh_i)^{-\theta} \exp(-kh_i)\; kdh_i = k^\theta \Gamma(1-\theta),$$

where $\Gamma(t) = \int_0^\infty x^{t-1} \exp(-x) dx$ is the Euler gamma function. Given (26), the actual output q_i of a craftsman is:

$$q_i = \min\{\bar{q}_i, \bar{q}\} = \min\left[h_i^{-\theta}, k^\theta \Gamma(1-\theta)\right].$$

The threshold for h_i below which craftsmen are constrained by average knowledge is given by:

$$\hat{h} = k^{-1} \Gamma(1-\theta)^{-1/\theta}.$$

The expected supply of output of a given craftsman is:

$$\begin{aligned}\mathrm{E}(q_i) &= \int_0^{\hat{h}} k^\theta \Gamma(1-\theta) k \exp[-kh_i] dh_i + \int_{\hat{h}}^{\infty} h_i^{-\theta} k \exp[-kh_i] dh_i \\ &= k^\theta \Lambda.\end{aligned}$$

Here Λ is a constant given by:

$$\Lambda = \Gamma(1-\theta) + \exp[-\Gamma(1-\theta)^{-1/\theta}]\theta\Gamma(-\theta) + \Gamma(1-\theta, \Gamma(1-\theta)^{-1/\theta}),$$

and $\Gamma(t,s) = \int_s^\infty x^{t-1} \exp(-x) dx$ is the incomplete gamma function. The total supply of craftsmen's labor L in efficiency units is then given by the expected output per craftsmen $\mathrm{E}(q_i)$ multiplied by the number of craftsmen N:

$$L = N\, k^\theta\, \Lambda.$$

In sum, all the results in the benchmark continue to hold with the "Leonardo" assumption, up to some constant terms: $\Gamma(1-\theta)$ should be replaced by Λ.

In the proof of Proposition 3, the "Leonardo" assumption requires to compute $\partial \mathrm{E}(q')/\partial k'$ differently. In particular, in deriving $\mathrm{E}(q')$ with respect to k' we should be careful in taking as given the future average society knowledge $\bar{q}' = (k')^\theta \Gamma(1-\theta)$, i.e. the externality in (26). More precisely, (26) should be written as:

$$\mathrm{E}(q') = \int_0^{\hat{h}'} \underbrace{(k')^\theta \Gamma(1-\theta)}_{\bar{q}\text{ (exogenous)}} k' \exp[-k'h_i] dh_i + \int_{\hat{h}'}^{\infty} h_i^{-\theta} k' \exp[-k'h_i] dh_i$$

where

$$\hat{h}' = \Gamma(1-\theta)^{-1/\theta}/k'.$$

Integrating we obtain:

$$\mathrm{E}(q') = (1 - \exp[\Gamma(1-\theta)/\theta])\, \bar{q} + (k')^\theta\, \Gamma(1-\theta, \Gamma(1-\theta)^{-1/\theta})$$

and the derivative is

$$\frac{\partial \mathrm{E}(q')}{\partial k'} = \theta\Gamma(1-\theta, \Gamma(1-\theta)^{-1/\theta})(k')^{\theta-1} = \Theta(k')^{\theta-1}.$$

with $\Theta = \theta\Gamma(1-\theta, \Gamma(1-\theta)^{-1/\theta})$. Compared to the benchmark, there is a factor $\Gamma(1-\theta, \Gamma(1-\theta)^{-1/\theta})$ instead of a factor $\Gamma(1-\theta)$.

C Proof of Proposition 2

The threshold for sufficient altruism is given by

$$\hat{\gamma} = \frac{(\delta(n^{\mathrm{F}}) - \kappa n^{\mathrm{F}})\,\bar{n}s}{(1-\alpha)(n^{\mathrm{F}})^2}\Gamma(1-\theta), \quad \text{with } n^{\mathrm{F}} = (1+\nu)^{\frac{(1-\alpha)\theta}{\alpha}}. \tag{27}$$

We claim that the balanced growth path exists if $\gamma > \max\{0, \hat{\gamma}\}$.

Let us first compute the balanced growth path, supposing it is incentive compatible. The growth rate of knowledge in (b) comes from (13) where we have imposed $m = m^{\mathrm{F}} = 1$. Population growth is obtained using (9). Income per capita in (c) derives from (8). Utility in (d) is derived as follows: Income of a given craftsman is $q_i\,(1-\alpha)\,\frac{Y}{L} + \kappa a^{\mathrm{F}}$ (from (5)). Expected income, using (6), is $k^{\theta}\Gamma(1-\theta)(1-\alpha)\frac{Y}{L} + \kappa a^{\mathrm{F}} + \alpha y^{\mathrm{F}}$, where αy^{F} is income from owning land. Given the value of L from (7), this simplifies into $(1-\alpha)\frac{Y}{N} + \kappa a^{\mathrm{F}} + \alpha y^{\mathrm{F}}$. The future labor income of the child is $(1-\alpha)y^{\mathrm{F}}$.

For the balanced growth path defined above to be incentive compatible, i.e. fathers are indeed willing to provide their kids with teaching, the cost of teaching should be less than the gain in the income, as evaluated by the altruistic parents. Normalizing the labor income of a craftsman without training to zero, this condition is:

$$\delta(n^{\mathrm{F}}) - \kappa n^{\mathrm{F}} < \gamma\, n^{\mathrm{F}} E\left[(1-\alpha)q\frac{Y}{L}\right] = \gamma\, n^{\mathrm{F}}(1-\alpha)y^{\mathrm{F}}$$

Using (6) and (7), this condition determines a lower bound on the altruism factor γ:

$$\hat{\gamma} > \frac{(\delta(n^{\mathrm{F}}) - \kappa n^{\mathrm{F}})\,\bar{n}s}{(1-\alpha)(n^{\mathrm{F}})^2}\Gamma(1-\theta).$$

D Proof of Proposition 3

The required threshold for altruism is given by:

$$\gamma > \frac{\delta'\left((1+\nu)^{\frac{(1-\alpha)\theta}{\alpha}}\right) - \kappa}{\frac{(1-\alpha)\nu\theta}{\bar{n}s}(1+\nu)^{\frac{(1-\alpha)\theta}{\alpha}-2}}.$$

Let us first compute the balanced growth path, supposing it is incentive compatible. To determine the number of apprentices per master, we use (14). Equation (15) is derived as follows. Each apprentice learns from $m^{\mathrm{c}} = (n^{\mathrm{c}})^{o}$ masters. As the draws for the initial knowledge of masters are not independent (all these masters were educated by the same persons), their acquired knowledge k^i is the same, but they had different ideas on their own drawn from $\mathrm{Exp}(\nu k_{-1})$. The acquired knowledge of the apprentices thus follows:

$$(k^i)' = k^i + m^{\mathrm{c}}\nu k_{-1}$$

The final knowledge of the apprentices is given by

$$k' = (k^i)' + \nu k$$

Using $(k^i)' = k' - \nu k$ and $(k^i) = k - \nu k_{-1}$ in the first equation, we get

$$k' - \nu k = k - \nu k_{-1} + m^{\mathrm{c}}\nu k_{-1} \tag{28}$$

which leads to (15) where $g = k'/k$.

The average utility expression takes into account the flows between master and apprentices. Adults are paid as masters by a^{c} parents the amount of $\delta'(a^{\mathrm{c}}) - \kappa$. Their disutility net of income from apprentices is $\delta(a^{\mathrm{c}}) - \kappa(a^{\mathrm{c}})$. They also pay as parents the same amount $\delta'(a^{\mathrm{c}}) - \kappa$ for each of their children n^{c} to each of their master m^{c} . The balance is:

$$a^{\mathrm{c}}(\delta'(a^{\mathrm{c}}) - \kappa) - (\delta'(a^{\mathrm{c}}) - \kappa) - m^{\mathrm{c}}n^{\mathrm{c}}(\delta'(a^{\mathrm{c}}) - \kappa) = -(\delta(a^{\mathrm{c}}) - \kappa a^{\mathrm{c}}).$$

As population is growing, they have to pay more for their children than what they get as teacher.

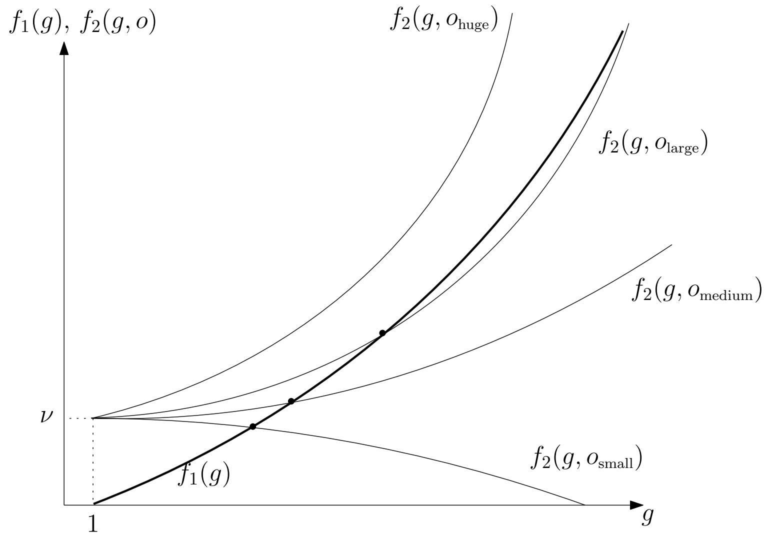

Figure 8: Equation (29)

From Equation (15) and the value of n^{c}, the growth rate $g = g^{\text{c}}$ should satisfy:

$$g^2 - (1+\nu)g + \nu = \nu g^{\frac{(1-\alpha)\theta o}{\alpha}}. \tag{29}$$

The left hand side is a convex function $f_1(g)$, while the right hand side is a function of g and o, $f_2(g,o)$. Figure 8 represents these two functions for different sizes of the clan. At the minimum possible value of g, the left hand side is smaller than the right hand side:

$$f_1(1) = 0 < f_2(1,o) = \nu \forall o > 0.$$

Several cases may occur:

- For $o \leq \frac{\alpha}{(1-\alpha)\theta}$, $f_2(g,o)$ is concave, crosses $f_1(g)$ once, and there exist a solution to the equality (29).
- For $\frac{\alpha}{(1-\alpha)\theta} < o \leq \frac{2\alpha}{(1-\alpha)\theta}$, one can show applied l'Hospital twice that

$$\lim_{g\to\infty} \frac{f_1(g)}{f_2(g,o)} > 1,$$

  implying that $f_2(g,o)$ is below $f_1(g)$ for large g and crosses it once. Hence there exist a solution to the equality (29).

- If $o > \frac{2\alpha}{(1-\alpha)\theta}$ but not "too large", $f_2(g, o)$ cuts $f_1(g)$ twice. There are two balanced growth paths.
- For o very large, the function $f_2(g, o)$ which is above $f_1(g)$ for $g = 1$, stays above it as g increases. In that case, there is no solution to Equation (29), and no balanced growth path. The interpretation is that the clan is so big that technological level and population grow at an accelerating rate.

We can summarize these findings by defining $\bar{o}$, such that if $o \leq \bar{o}$, a balanced growth path exists.

Differentiating (29), the effect of o on g is given by:

$$\frac{dg}{do} = \frac{\partial f_2/\partial o}{\partial f_1/\partial g - \partial f_2/\partial g}.$$

The term $\partial f_2/\partial o$ is positive as $g > 1$. We conclude that increasing o increases g for equilibria where $\partial f_1/\partial g > \partial f_2/\partial g$, that is where $f_2(g, o)$ cuts $f_1(g)$ from above as g increases.

Let us now consider whether such an equilibrium is incentive compatible. If o is too low, the threat of punishment is insufficient to prevent shirking, and only the family equilibrium exists. If o is large, parents may no longer be willing to apprentice their children with all current adult members of the clan. In other words, the clan should be large enough for the threat of punishment to ensure compliance, but small enough for parents to be willing to pay the apprenticeship fee. Hence, there exists thresholds $o_{\min}$ and $o_{\max}$ such that if $o_{\max} > o > o_{\min}$ the clan equilibrium is incentive compatible.

Assuming that the punishment technology is such that one person can only do negligible harm to another, but more than one person can exert a much more damaging action, sufficient in any case to deter shirking, we get $o_{\min} = 0$.[29] Let us now consider $o_{\max}$.

The clan equilibrium is sustained if, for each child, the marginal cost paid to the

[29] Considering o as a continuous variable

master is less than the expected marginal benefit, as priced by the parents:

$$\delta'(a) - \kappa \leq \gamma \frac{\partial \mathrm{E}(q')}{\partial k'} \frac{\partial k'}{\partial m} (1-\alpha) \frac{Y'}{L'}. \tag{30}$$

The right hand side represents the expected effect on individual productivity of meeting an additional master. $(1-\alpha)\frac{Y'}{L'}$ is exogenous for the individual, and the altruism parameter γ reflects that the marginal benefit is seen from the point of view of the parent.

Let us first consider the term $\partial \mathrm{E}(q')/\partial k'$. From (6), the derivative is

$$\frac{\partial \mathrm{E}(q')}{\partial k'} = \theta \Gamma(1-\theta)(k')^{\theta-1}.$$

The term $\partial k'/\partial m$ can be directly obtained using (28) and is equal to νk_{-1}. Finally, the term Y'/L' can be transformed into:

$$\frac{Y'}{L'} = \frac{Y'}{N'} \frac{N'}{L'} = \frac{Y'}{N'} \frac{1}{\Gamma(1-\theta)(k')^{\theta}} \tag{31}$$

using (7). Condition (30) can now be rewritten as:

$$\delta'(a) - \kappa \leq \gamma \theta (k')^{\theta-1} \ \nu k_{-1} \ (1-\alpha) \frac{Y'}{N'} \frac{1}{(k')^{\theta}}. \tag{32}$$

which, along a balanced growth path, reduces to

$$\delta' \left((g^{\mathrm{c}})^{\frac{(1-\alpha)\theta(o+1)}{\alpha}} \right) - \kappa \leq \frac{\gamma(1-\alpha)\nu\theta}{\bar{n}s} \ (g^{\mathrm{c}})^{\frac{(1-\alpha)\theta}{\alpha} - 2}.$$

The left hand side is increasing in o as $g^{\mathrm{c}} > 1$, g^{c} is increasing in o, and $\delta(a)$ is convex. If it is smaller than the right hand side for the minimum value of o ($o = 0$, $g^{\mathrm{c}} = 1 + \nu$), i.e. if

$$\delta' \left((1+\nu)^{\frac{(1-\alpha)\theta}{\alpha}} \right) - \kappa < \frac{\gamma(1-\alpha)\nu\theta}{\bar{n}s} (1+\nu)^{\frac{(1-\alpha)\theta}{\alpha} - 2}.$$

then either the right hand side becomes larger than the left hand side for some value of $o = \hat{o} > 0$, or they never intersect, in which case $\hat{o}$ is infinite. Notice that the right hand side is decreasing in g^{c}, and hence in o, provided that

$$(1-\alpha)\theta < 2\alpha,$$

in which case, $o_{\max}$, the maximum size of the clan which is incentive compatible, is necessarily finite.

In the analysis of the incentive compatibility, we have assumed that g^{c} was defined, we show above that it is not the case $o > \bar{o}$. Hence, the threshold above which a balanced growth path exists and is incentive compatible is $o_{\max} = \min\{\bar{o}, \hat{o}\}$

E Proof of Lemma 1

The first-order condition for the parent's problem can be written as:

$$\underbrace{p}_{\text{marginal cost}} \quad n = \gamma\, n \underbrace{\frac{\partial \mathrm{E}(q')}{\partial k'} \frac{\partial \mathrm{E}(k')}{\partial m} (1-\alpha) \frac{Y'}{L'}}_{\text{marginal benefit } (\partial I'/\partial m)}.$$

Remembering from (6) that

$$\frac{\partial \mathrm{E}(q')}{\partial k'} = \theta \Gamma (1-\theta)(k')^{\theta-1},$$

and using $k' = k(m+\nu)$ (from (13)) as masters are now drawn randomly, we obtain, in equilibrium:

$$\delta'(a) - \kappa = p = \gamma \theta \Gamma (1-\theta) \frac{(k')^{\theta}}{m+\nu} (1-\alpha) \frac{Y'}{L'}.$$

Using (7), this equation simplifies into

$$\delta'(a) - \kappa = \gamma \theta (1-\alpha) \frac{1}{m+\nu} \frac{Y'}{N'}.$$

F Proof of Proposition 4

For a fixed number of masters m, the market equilibrium yields higher growth in productivity because those masters are drawn randomly. Indeed, from the proof of Proposition 3, productivity in **C** follows

$$k' = (1+\nu)k + (m-1)\nu k_{-1}$$

which implies a growth rate equal to:

$$g^{\mathrm{C}} = \frac{1+\nu+\sqrt{(1+\nu)^2+4(m-1)\nu}}{2}$$

which is always less than the growth rate with the market, $m+\nu$, for $\nu > 0$ and $m > 1$.

Moreover, the equilibrium number of masters m^{M} is higher in the market equilibrium compared to the clan equilibrium. Indeed, m^{M} balances marginal cost and benefit. If the clan equilibrium had a higher number of masters, it would not be incentive compatible (parents would not like to pay all those masters).

Notice that, in the computation of the utility, the payments from apprentices, pa^{M}, and for children, $pn^{\mathrm{M}}m^{\mathrm{M}}$, balance.

G Proof of Lemma 2

The maximization problem of the guild is:

$$\max_{a_j} \left\{ p_j a_j - \delta(a_j) + \kappa a_j \right\}$$

subject to:

$$\begin{aligned} S_j N' m_j &= N a_j, \\ p_j &= \gamma \frac{\partial EI'_{ij}}{\partial m_j}, \\ \gamma EI'_{ij} - p_j m_j &= \gamma(1-\alpha)\frac{Y'}{N'} - pm. \end{aligned} \tag{33}$$

Replacing p_j and p by their value from the second constraint into the third leads to:

$$\gamma (S'_j)^{\frac{1}{\lambda}-1}\left(\frac{k'_j}{k'}\right)^{\frac{\theta}{\lambda}} - \frac{\theta}{m_j+\nu}\,(S'_j)^{\frac{1}{\lambda}-1}\left(\frac{k'_j}{k'}\right)^{\frac{\theta}{\lambda}} m_j = \gamma - \frac{\theta}{m+\nu}\, m.$$

which can be solved for S_j':

$$S_j' = \left[\frac{(\gamma - \theta)\, m + \gamma\nu}{(\gamma - \theta)\, m_j + \gamma\nu} \left(\frac{m_j + \nu}{m + \nu} \right)^{1 - \frac{\theta}{\lambda}} \right]^{\frac{\lambda}{1-\lambda}} \tag{34}$$

The second constraint can be rewritten as:

$$p_j = \gamma\theta(1-\alpha) \frac{1}{m+\nu} \frac{Y'}{N'} \frac{(\gamma - \theta)\, m + \gamma\nu}{(\gamma - \theta)\, m_j + \gamma\nu}$$

and the equilibrium on the apprenticeship market (33) is:

$$a_j = n\ S_j'\ m_j$$

These constraints implies that lowering the supply of apprenticeships a_j increases the price p_j. It is now easier to express the maximization program in terms of m_j:

$$\max_{m_j} \left\{ \left(\frac{\gamma\theta(1-\alpha)}{m+\nu} \frac{Y'}{N'} \frac{(\gamma - \theta)\, m + \gamma\nu}{(\gamma - \theta)\, m_j + \gamma\nu} + \kappa \right) n\ S_j'\ m_j - \delta \left(n\ S_j'\ m_j \right) \right\}$$

The first order condition is:

$$\begin{aligned} &- (\gamma - \theta) \left(\frac{\gamma\theta(1-\alpha)}{m+\nu} \frac{Y'}{N'} \frac{(\gamma - \theta)\, m + \gamma\nu}{((\gamma - \theta)\, m_j + \gamma\nu)^2} \right) S_j'\ m_j + \\ &\left(\left(\frac{\gamma\theta(1-\alpha)}{m+\nu} \frac{Y'}{N'} \frac{(\gamma - \theta)\, m + \gamma\nu}{(\gamma - \theta)\, m_j + \gamma\nu} + \kappa \right) - \delta' \left(n\ S_j'\ m_j \right) \right) \left(\frac{\partial S_j'}{\partial m_j}\, m_j + S_j' \right) = 0 \end{aligned}$$

with

$$\begin{aligned} \frac{\partial S_j'}{\partial m_j} = \frac{\lambda}{1-\lambda} & \left[\frac{(\gamma - \theta)\, m + \gamma\nu}{(\gamma - \theta)\, m_j + \gamma\nu} \left(\frac{m_j + \nu}{m + \nu} \right)^{1 - \frac{\theta}{\lambda}} \right]^{\frac{\lambda}{1-\lambda} - 1} \\ & \times \left(-(\gamma - \theta) \frac{(\gamma - \theta)\, m + \gamma\nu}{((\gamma - \theta)\, m_j + \gamma\nu)^2} \left(\frac{m_j + \nu}{m + \nu} \right)^{1 - \frac{\theta}{\lambda}} \right. \\ & \qquad \left. + \left(1 - \frac{\theta}{\lambda} \right) \frac{(\gamma - \theta)\, m + \gamma\nu}{(\gamma - \theta)\, m_j + \gamma\nu} \left(\frac{m_j + \nu}{m + \nu} \right)^{-\frac{\theta}{\lambda}} \right). \end{aligned}$$

At the symmetric Nash equilibrium between guilds, this becomes:

$$-\left(\gamma-\theta\right)\left(\frac{\gamma\theta(1-\alpha)}{m+\nu}\,\frac{Y'}{N'}\,\frac{1}{\left(\gamma-\theta\right)m+\gamma\nu}\right)\,m+\\\left(\left(\frac{\gamma\theta(1-\alpha)}{m+\nu}\,\frac{Y'}{N'}+\kappa\right)-\delta'\left(n\,m\right)\right)\left(\frac{\partial S_j'}{\partial m_j}\,m+1\right)=0$$

with:

$$\frac{\partial S_j'}{\partial m_j}=\frac{\lambda}{1-\lambda}\left(\frac{-(\gamma-\theta)}{\left(\gamma-\theta\right)m+\gamma\nu}+\left(1-\frac{\theta}{\lambda}\right)\right),$$

which we can rearrange into:

$$\frac{\gamma\theta(1-\alpha)}{m+\nu}\,\frac{Y'}{N'}\left(\frac{-\left(\gamma-\theta\right)m}{\left(\gamma-\theta\right)m+\gamma\nu}+\frac{\partial S_j'}{\partial m_j}\,m+1\right)\\+\left(\kappa-\delta'\left(n\,m\right)\right)\left(\frac{\partial S_j'}{\partial m_j}\,m+1\right)=0$$

and finally:

$$\delta'(nm)-\kappa=\left(\frac{\gamma\theta(1-\alpha)}{m+\nu}\,\frac{Y'}{N'}\right)\Omega(m)$$

with:

$$\Omega(m)=\frac{\left(\frac{-(\gamma-\theta)m}{(\gamma-\theta)m+\gamma\nu}+1+\frac{\lambda m}{1-\lambda}\left(\frac{-(\gamma-\theta)}{(\gamma-\theta)m+\gamma\nu}+\left(1-\frac{\theta}{\lambda}\right)\right)\right)}{1+\frac{\lambda m}{1-\lambda}\left(\frac{-(\gamma-\theta)}{(\gamma-\theta)m+\gamma\nu}+\left(1-\frac{\theta}{\lambda}\right)\right)}.$$

$\Omega(m)$ can be further simplified into:

$$\Omega(m)=\frac{1-\lambda+(\lambda-\theta)m-\frac{(\gamma-\theta)m}{(\gamma-\theta)m+\gamma\nu}}{1-\lambda+(\lambda-\theta)m-\lambda\,\frac{(\gamma-\theta)m}{(\gamma-\theta)m+\gamma\nu}}$$

When $\lambda\to1$, $\frac{\partial S_j'}{\partial m_j}\to\infty$ and $\Omega(m)\to1$.

H Proof of Proposition 5

Before considering the optimization problem, let us compute the effect of changing m_j on income I_j. Using Equation (7), we can compute the relative quantity of efficient labor in sector j as:

$$\frac{L'_j}{L'} = S'_j \left(\frac{k'_j}{k'}\right)^{\theta}.$$

With this expression, and with Equation (6), which adapts to trade j as $Eq_j = \Gamma(1-\theta)k_j^{\theta}$, it is convenient to rewrite expected income as:

$$EI'_{ij} = (1-\alpha)\,\frac{Y'}{N'}(S'_j)^{\frac{1}{\lambda}-1}\left(\frac{k'_j}{k'}\right)^{\frac{\theta}{\lambda}}.$$

Let us now compute the effect of changing m_j on individual income. Using the result in (31), and $k'_j = (m_j+\nu)k_j$, we get:

$$\frac{\partial EI'_{ij}}{\partial m_j} = \theta\Gamma(1-\theta)\frac{(k'_j)^{\theta}}{m_j+\nu}\,(1-\alpha)\,\frac{Y'}{L'}\left(\frac{L'_j}{L'}\right)^{\frac{1}{\lambda}-1}$$

which can moreover be simplified into, using (31):

$$\frac{\partial EI'_{ij}}{\partial m_j} = \theta(1-\alpha)\frac{(k'_j)^{\theta}}{(k')^{\theta}}\frac{1}{m_j+\nu}\,(1-\alpha),$$

leading to:

$$\frac{\partial EI'_{ij}}{\partial m_j} = \theta(1-\alpha)\frac{1}{m_j+\nu}\,\frac{Y'}{N'}(S'_j)^{\frac{1}{\lambda}-1}\left(\frac{k'_j}{k'}\right)^{\frac{\theta}{\lambda}}. \tag{35}$$

Marginal income is therefore equal to expected income multiplied by: $\frac{\theta}{m_j+\nu}$.

To study the equilibrium, one should consider Equation (24) replacing a^{G}, m^{G} and y^{G} by their value:

$$\delta'((g^{\mathrm{G}}-\nu)n^{\mathrm{G}}) - \kappa = \gamma\theta(1-\alpha)\frac{1}{g^{\mathrm{G}}}\;\frac{n^{\mathrm{G}}}{\bar{n}s}\;\Omega(g^{\mathrm{G}}-\nu) \tag{36}$$

This equation describes a relationship between g^{G} and n^{G} which we call the "apprenticeship monopolistic market", as it is derived from the demand for apprenticeship and the monopolistic behavior of the guild. Equation (36) can be rewritten as:

$$n^{\mathtt{G}} = \frac{\kappa}{\bar{\delta}(g^{\mathtt{G}} - \nu) - \dfrac{\gamma\theta(1-\alpha)}{s\bar{n}\ g^{\mathtt{G}}}\Omega(g^{\mathtt{G}} - \nu)}.$$

If we compare this expression with the equivalent in the market equilibrium, Equation (20), we see that the denominator is necessarily larger. Hence, for any given g, $n^{\mathtt{G}} < n^{\mathtt{M}}$. It follows that $g^{\mathtt{G}} < g^{\mathtt{M}}$.

Notice finally that, as in the market equilibrium, the payments from apprentices, $pa^{\mathtt{G}}$, and for children, $pn^{\mathtt{G}}m^{\mathtt{G}}$, balance in the computation of the utility.

I Proof of Proposition 6

We can compute the gains of adopting the guild institution as:

$$\begin{aligned} u^{\mathtt{F}\to\mathtt{G}} - u^{\mathtt{F}\to\mathtt{F}} &= \gamma n_0(y_1^{\mathtt{G}} - y_1^{\mathtt{F}}) + \kappa(a_0 - n_0) - \delta(a_0) + \delta(n_0) - \mu(N_0) \\ u^{\mathtt{C}\to\mathtt{G}} - u^{\mathtt{C}\to\mathtt{C}} &= \gamma n_0(y_1^{\mathtt{G}} - y_1^{\mathtt{C}}) + \kappa(a_0 - (n_0)^{o+1}) - \delta(a_0) + \delta((n_0)^{o+1}) - \mu(N_0) \end{aligned}$$

$\underline{N}$ makes people in the family equilibrium indifferent between adopting the guild or not, i.e. it solves:

$$\gamma n_0(y_1^{\mathtt{G}} - y_1^{\mathtt{F}}) + \kappa(a_0 - n_0) - \delta(a_0) + \delta(n_0) - \mu(\underline{N}) = 0.$$

One should show that for $N_0 = \underline{N}$, people in the clan equilibrium do not want to adopt the guild,i.e.

$$\gamma n_0(y_1^{\mathtt{G}} - y_1^{\mathtt{C}}) + \kappa(a_0 - (n_0)^{o+1}) - \delta(a_0) + \delta((n_0)^{o+1}) - \mu(\underline{N}) < 0$$

This is true if:

$$\gamma n_0 y_1^{\mathtt{C}} + (\kappa(n_0)^{o+1} - \delta((n_0)^{o+1})) > \gamma y_1^{\mathtt{F}} + \kappa n_0 - \delta(n_0). \tag{37}$$

Let us define the following function:

$$\psi(m) = \gamma n_0 u(m) + \kappa m n_0 - \delta(m n_0).$$

$u(m)$ is the function that relates future income to number of masters learning from, in the context of the clan equilibrium. From (10) and (28), we get:

$$u(m) = \Gamma(1-\theta)^{1-\alpha} (\underbrace{(1+\nu)k - \nu k_{-1} + m\nu k_{-1}}_{k'})^{(1-\alpha)\theta} (N')^{-\alpha}.$$

Hence, $u(m)$ is increasing and concave in m. As a consequence, $\psi(m)$ is also concave in m ($\delta(\cdot)$ is convex). To see whether it is increasing, we can compute:

$$\psi'(m) = \gamma\ n_0(1-\alpha)\Gamma(1-\theta)^{1-\alpha}\theta(k')^{(1-\alpha)\theta-1}\nu k_{-1} (N')^{-\alpha} + \kappa n_0 - \delta'(mn_0)n_0.$$

We also know from Appendix D that the clan equilibrium is sustained if the marginal cost paid to the master is less than the expected marginal benefit, i.e. $\delta'((n_0)^{o+1}) - \kappa \leq \partial y_1^{\mathsf{C}}/\partial m_0$, which implies, from (32) and (10):

$$\gamma\theta\ (1-\alpha)\Gamma(1-\theta)^{1-\alpha} (k')^{(1-\alpha)\theta-1}\nu k_{-1} (N')^{-\alpha} + \kappa - \delta'(mn_0) \geq 0.$$

This individual level condition implies that, at the aggregate equilibrium, we have $\psi'(m) > 0$. Using the mean value theorem for derivatives, we know there exists $\tilde{m} \in [1, m^{\mathsf{C}}]$ such that $(\psi(m^{\mathsf{C}}) - \psi(1))/(m^{\mathsf{C}} - 1) = \psi'(\tilde{m})$. As $\psi(\cdot)$ is concave, $\psi'(\tilde{m}) > \psi'(m^{\mathsf{C}}) > 0$ which proves $\psi(m^{\mathsf{C}}) > \psi(1)$ and inequality (37) holds.

$\overline{N}$ makes people in the clan equilibrium indifferent between adopting the guild or not, i.e. it solves

$$\gamma n_0(y_1^{\mathsf{G}} - y_1^{\mathsf{C}}) + \kappa(a_0 - (n_0)^{o+1}) - \delta(a_0) + \delta((n_0)^{o+1}) - \mu(\overline{N}) = 0.$$

One should show that for $N_0 = \overline{N}$, people in the family equilibrium also want to adopt the guild, i.e.:

$$\gamma n_0(y_1^{\mathsf{G}} - y_1^{\mathsf{F}}) + \kappa(a_0 - n_0) - \delta(a_0) + \delta(n_0) - \mu(\overline{N}) > 0.$$

This is true as $\mu(\overline{N}) < \mu(\underline{N})$.