Essays in Economic History and Long-Term Institutional Development

Inauguraldissertation zur Erlangung des akademischen Grades eines Doktors der Wirtschaftswissenschaften der Universität Mannheim

Angenommen durch die
Fakultät für Rechtswissenschaft und Volkswirtschaftslehre
Abteilung Volkswirtschaftslehre
Dekan: Prof. Dr. Carsten Trenkler
Referenten: Prof. Dr. Jochen Streb und Prof. Dr. Sibylle Lehmann-Hasemeyer
Tag der Prüfung: 07.04.2016

Essays in Economic History and Long-Term Institutional Development

Katharina Mühlhoff

Preface:

The papers of this dissertation deal with the nervous system of the economy - with institutions. Although all chapters focus on a specific historical case of their own, they have a common guiding theme: The question how the economy's organizational *nerves* respond to changes in social and environmental conditions and which circumstances lead to institutional dysfunction or paralysis. Anatomists know that the human body possesses a network of uncounted neurons which control the interplay of brain, skin, muscles and inner organs. Likewise, economies require institutions to establish markets and to coordinate the interaction of groups and individuals. Due to their focal role, the "rules of the game" as Douglass North[1] famously dubbed them have attracted the interest of economists from the beginnings of the discipline onwards: Classics like Smith, Mill and Ricardo joined the French Physiocrats in defending the idea of competition on free markets. Later, Weber pointed to the economic merits of rational bureaucracies, and finally the proponents of New Institutional Economics drew attention to transaction costs, asymmetric information and historical path dependency.[2] Today, scholars widely agree that economic growth depends on a society's ability establish institutions which foster health and education, prevent rent-seeking and corruption and promote mutual trust, free competition and long-term investment. Still, the general consensus notwithstanding, matters are not settled for good. For, unlike their hypothetical counterparts in economic models, real societies are not stable. They evolve, and with them the social and demographic characteristics, income distribution and preferences of their members. To promote the welfare of these changing societies, institutions must adapt to a wide range variations in the population's political and environmental conditions. In a sense, successful institutions therefore need to

[1] Douglass C. North (1990). *Institutions, Institutional Change and Economic Performance*, Cambridge University Press

[2] To give but a few, review articles of the many currents in the field include: Elinor Ostrom (2005). "Doing Institutional Analysis: Digging Deeper than Markets and Hierarchies," Handbook of New Institutional Economics, C. Ménard and M. Shirley, eds. Handbook of New Institutional Economics, pp. 819-848; Oliver E. Williamson (2000). "The New Institutional Economics: Taking Stock, Looking Ahead," Journal of Economic Literature, vol. 38 No.3, pp. 595-613; Ronald Coase (1998). "The New Institutional Economics," American Economic Review, vol. 88, Nr. 2, pp. 72-74.

mimic the neurons in our eyes, skin or ears: They have to ensure the smooth functioning of the whole system in daylight and darkness, in heat and cold, in noise and silence. This implies that we cannot assess the performance of past institutions - let alone inform current development policies - unless we understand the process of institutional change and identify the factors that promote or impede successful *institutional adaptation*.

As discussed by a voluminous body of research, obstacles to institutional change can stem from exogenous (e.g. ecological or geographical) or from man-made (political and cultural) conditions.[3] To account for these different sources of successful adaptation or institutional paralysis the following chapters scrutinize instances where particularly dramatic change of either the political, social and cultural or the scientific and ecological framework put institutions under enormous evolutionary pressure.

Starting with an example of man-made challenges of the institutional *status quo*, the first chapter analyzes how administrative modernization facilitated the transition to modern economic growth in Meiji Japan. The specific example is so intriguing because the Japanese not only experienced an industrial revolution in "fast motion", but also because they managed to enforce a course of radical institutional reform against the interests of powerful traditional elites. Uncovering the political ingredients of this staggering success thus contributes to previous research on organizational modernization and sources of institutional inertia. Besides satisfying historians' curiosity, Meiji Japan's case of institutional adaptation does – however – also have a bearing for current policies. For given that rational bureaucracies are considered indispensable foundations for sustainable economic development, understanding how a conflict-torn country with limited state capacity managed to embark on a course of almost unprecedented economic, social and scientific progress carries invaluable information for today's developing nations.

Whereas the first case focuses on political and social causes of institutional function or dysfunction, the second chapter takes a different perspective. Its vantage point is one of the most revolutionary innovations ever made in medical science: Vaccination against smallpox. With this weapon at hand, modern medicine was able to protect entire populations against the fatal ravages of a hitherto uncontrolled epidemic disease. If anything, vaccination therefore initiated a profound change in the biological environment of 19th century Europe. Yet, at the same time, the intensification of international trade and mobility during the first globalization increased the pandemic potential of the disease dramatically. As a result, health authorities in

[3] See: North, Douglass C. Institutions, institutional change and economic performance. Cambridge university press, 1990.; La Porta, Rafael, et al. "The quality of government." Journal of Law, Economics, and organization vol. 15. No.1 (1999): 222-279;Sen, Amartya. How does culture matter?. 2000.Dosi, Giovanni, and Richard R. Nelson. "An introduction to evolutionary theories in economics." Journal of evolutionary economics, vol. 4 No. 3 (1994): 153-172.;Galor, Oded, Omer Moav, and Dietrich Vollrath. "Inequality in landownership, the emergence of human-capital promoting institutions, and the great divergence." The Review of economic studies vol. 76 No.1 (2009): 143-179; Acemoglu, Daron, Simon Johnson, and James A. Robinson. "Institutions as a fundamental cause of long-run growth."Handbook of economic growth vol. 1 (2005): 385-472.; Barro, Robert J., and Jong-Wha Lee. "Sources of economic growth."Carnegie-Rochester conference series on public policy. Vol. 40. North-Holland, 1994.; Rodrik, Dani, Arvind Subramanian, and Francesco Trebbi. "Institutions rule: the primacy of institutions over geography and integration in economic development." Journal of economic growth vol. 9. No. 2 (2004): 131-165.

the 19th century faced more or less the same problems which continue to trouble epidemiologists at present: How to avoid epidemic spread if “closing the border” is no alternative? How to convince people of the merits of prevention and how to attain high immunization rates? How to ensure prompt detection and containment of potential outbreaks? And finally, how to minimize the economic costs of the chosen epidemiological policy?

The above suffices to make clear that smallpox has many lessons to teach about past and current health threats. More importantly though, some of them relate directly to our analysis of institutional adaptation. This is particularly true because, although European governments imposed largely identical policies when vaccination was first introduced, their strategies diverged markedly in the course of time. As chapter 2 attempts to make sense of this puzzling observation, it finds that the different states' institutional response depended crucially on specific geographic and political conditions. Whether these factors lead to successful institutional adaptation or whether they produced inefficiencies at the national and international levels will therefore be at the center of interest in the section.

To close the discussion, chapter 3 further investigates the role of natural (e.g. biological) constraints of institutional change. To advance its argument, it considers one of the major social concomitants of the industrial revolution: the demographic transition. Building on a rich literature, the paper's approach differs from older research in that it combines elements of economic fertility theory with recent results from anthropology and evolutionary biology. The quintessence of this analysis is that social progress - like overcoming Malthusian population dynamics - can be viewed as an adaptation to altered selective pressures. Thus, long-term economic growth and substantial improvements in a society's living conditions can emerge endogenously. But whether or not the desired behaviors will be adopted or not depends on physiological and psychological processes which belong to the evolutionary blueprint of our species. In consequence, policies to promote social and economic progress might be doomed to noble failure unless they affect the entire environment such that there is no contradiction between biological necessities and economically beneficial collective behavior.

The previous glimpse into the central questions of the subsequent papers underscores the fact that thinking of institutions as the nerves of an economy is no void metaphor. Not only does a multitude of formal and informal institutions serve to coordinate economic interaction. They must also ensure that state and markets remain functional even if the prevailing conditions are thoroughly altered. In sum, the institutional nerves of past and present economies are involved in a complex web of factors, many of which escape the focus of orthodox economics.

Trying to disentangle these various relationships was an academic adventure with many detours. Keeping me on track was a strain on my advisors, colleagues, family and friends. Judging the results of this work is up to the critical reader. Yet, where pushing the boundaries of our discipline is concerned, the Meiji reformers of the next chapter would certainly have approved of the adventure. And some of them would probably have pointed to ancient Japanese folk wisdom: 井の中蛙大海を知らず. “The frog who lives in a well has no idea of the ocean.”

Cliometrica (2014) 8:201–239
DOI 10.1007/s11698-013-0099-x

ORIGINAL PAPER

The economic costs of sleaze or how replacing samurai with bureaucrats boosted regional growth in Meiji Japan

Katharina Muehlhoff

Received: 4 March 2013 / Accepted: 15 June 2013 / Published online: 6 August 2013

Abstract The notion that professional, efficient and non-corrupt bureaucracies foster economic growth is virtually uncontested. In spite of this wide consensus, central questions remain unanswered. Thus, while the harmful effects of dysfunctional administrations are extensively covered in the theoretical literature, little is known about the empirical relevance and the expected costs of insufficient administrative rationalization. And while efficient bureaucracies are considered a key ingredient to institutional performance, the existing research rarely investigates how desirable administrative structures have been implemented in history or which concrete policy measures constitute feasible reform strategies for present-day development countries. The present paper therefore aims at providing empirical evidence to dose this lacuna; to do so, it relies on the case of administrative reforms in the last three decades of the nineteenth century in Meiji Japan. Building on an exceptionally detailed set of official statistics and documentary sources, it constructs a panel of 45 Japanese prefectures and assesses the impact of heterogeneous reform implementation on canonical indicators of economic performance including measures of regional GDP, business activity and financial market development. The central results of the econometric analysis are that delayed administrative rationalization came along with a statistically significant and robust penalty on all development indicators. Moreover, this effect was remarkably persistent over time, as the data show that late-reforming prefectures performed systematically worse than the administrative forerunners until well into the twentieth century.

Keywords Meiji Japan · Administrative reform · Economic development

JEL Classification N40 · N45 · O17

K. Muehlhoff (✉)
Department of Economics, University of Mannheim, L7, 3-5, 68131 Mannheim, Germany
e-mail: kmuehlho@rumms.uni-mannheim.de

1 Introduction

Looking at early photographs of the Meiji cabinet is revealing: From the sepia-colored ranks of solemn politicians in Western-style suits, cardigans and bow-ties one figure inevitably stands out: A tall samurai in full ornate, proudly wearing his traditional "*hakama*" and the two swords. The person in question, Saigo Takamori, was one of the most powerful leaders in the coup d'état that overthrew the Tokugawa *shôgunate* and restored imperial power in 1868. Still, his prominent position did not save him from violent and premature death. Turning against his former allies, he headed a large-scale rebellion of samurai who opposed the Meiji government's attempts at abolishing ancient privileges of the warrior class. Defeated in a three-month civil war, he chose to commit *seppuku*—ritual suicide by disembowelment—on September 24, 1877.

The fate of Saigo Takamori is more than a story of triumph and tragedy in an era of cataclysmic change. It exemplifies a metamorphosis almost all developed countries underwent during their industrialization: The institutionalization of legal-rational rule by professional bureaucracies. The transition toward administrative rationality typically involved two characteristics: First, state action and the allocation of public goods became subject of universal, clearly defined rules. And second, no matter how fierce their resistance, traditional elites eventually lost their dominant influence as holders of public positions were deprived of property rights to their office and evolved into a class of government-appointed civil servants. Pre-modern forms of personal rule by privileged social groups thus gave way to rational bureaucratic governance and allowed socioeconomic development to gain momentum. In a sense, the samurai therefore had to vanish from the pictures before Japan could accomplish its leap to modernization in earnest.

The potential benefits of administrative rationality are not new to economic history. In fact, contributions like North (2009), Acemoglu et al. (2008) or Besley and Ghatak (2009) all stress the growth-promoting effects of impersonal, law-based and perpetual-lived institutions. Proceeding along similar lines, various authors including De Long and Shleifer (1993), Shleifer and Vishny (1993), (La Porta et al. 1999), Epstein (2000) and Buchheim (2006) put emphasis on limitations of the government's extractive discretion on the importance of transparent bureaucracies and on the state's capacity to assert the common public interest against the interests of established elites. Finally, theoretical ties between the establishment of legal-rational bureaucracies and economic development in Meiji Japan have been analyzed by Koh (1991), Ohshima (1994) and Mosk (2011). The literature thus provides sufficient theoretical ground to consider personal rule by traditional elite networks as detrimental for economic development. Still, the natural follow-up question "If sleaze is bad—just *how* bad is it?" is hardly addressed. And the few studies that deal explicitly with the empirical relationship between administrative rationality and economic performance largely rely on current cross-sections which prevents them from analyzing the long-term effects of retarded bureaucratic modernization.

To bridge this gap, the present paper considers the case of administrative rationalization in Japan after the Meiji Restoration. It constructs a panel of 45

Japanese prefectures[1] and employs canonical econometric techniques to assess the impact of regional heterogeneity in administrative rationality on different indicators of economic development from 1875 till 1895. This specific setup has been chosen for several reasons: First, administrative reorganization in Meiji Japan was radical and took no more than two decades starting in the 1870s. Thus, if the reforms had measurable effects, they ought to be distinct and distinguishable from secular trends. In addition, the Meiji Restoration ended a period of more than 250 years of strictly isolationist foreign policies. In consequence, Japan featured an astonishing degree of cultural and ethnical homogeneity and a regionally highly fragmented administration because she had maintained her traditional feudal system throughout the *shôgunate* period. This unique constellation facilitates both the analysis of temporal dynamics and helps disentangling the impact of institutional modernization from cultural differences. Finally, the favorable identification conditions can be fully exploited as Meiji statistics provide detailed and comprehensive data down to the local level.

Put together, the Meiji reforms not only furnish evidence on the costs of administrative encrustation to past economies. Given that state capacity in nineteenth century, Japan was just as scarce as in the developing world today, understanding what allowed the Meiji government to overcome these obstacles also carries lessons for current public sector reform.

The reminder of the paper is organized as follows: Sect. 2 gives a historical overview and discusses qualitative evidence of reform success. Section 3 investigates the relationship between administrative reform and economic development based on an econometric panel analysis. Section 4 performs a set of robustness checks and Sect. 5 concludes.

2 Historical overview

2.1 The Tokugawa legacy

The Meiji Restoration of 1868 ended over two and a half centuries of dominance by military rulers (*shôguns*) of the Tokugawa clan. Having risen to power in a period of incessant riots and civil war, the Tokugawa were anxious to craft institutions that guaranteed stability and their clan's political hegemony. Yet, skilled political craftsmen they were, they lacked the resources to assume a position of absolute sovereignty. Exerting central authority only in the areas of external affairs and the core organization of the feudal system, the *shôguns* had no means to interfere with other domains' inner affairs and neither disposed of a national army, police or bureaucracy nor a claim to the tax revenue of the *daimyo,* Japan's two-hundred-plus other feudal lords.[2] To maintain the fragile equilibrium of power, the Tokugawa

[1] Hokkaido and Okinawa were excluded because even before the Restoration, their structural and cultural background differed substantially from the other Japanese regions.

[2] For a summary of Tokugawa institutions and political thought see: Macklin (1992), Harootunian (1970) and Beasley (1972).

imposed policies of almost complete isolation from international relations, social stratification and regional sectionalism. This settlement proved successful in preventing internal turmoil and bloodshed, yet it was bought at the cost of impeding joint action at the national level let alone targeted economic or social development policies.

Under the stress of the forced opening of Japan to the Western powers in 1854, the structural weaknesses of the feudal arrangements became evident and helped forging an alliance of court nobles and *daimyo*—notably the lords of Satsuma, Chôshû and Tosa provinces—who overthrew the *shôgunate* and declared the Meiji Emperor's restoration to power in 1868. Well aware of their precarious situation vis-à-vis the West, the leaders of the imperial government embarked on an ambitious modernization course under the notion of *Fukoku Kyôhei* ("a rich country and a strong army"). Yet, this strategy was unfeasible unless they secured stable sources of public revenue and assumed full control of society and economy.[3]

Consequently, the following two decades were marked by the central government's struggle to replace the domains' semi-autonomous administration with a centralized, loyal and professional civil service—in short: to support their *Fukoku Kyôhei* program by almost a textbook example of Weberian bureaucratic rationalization. Still, due to historical and current political conditions, the timing and success of these measures differed greatly across regions. This very characteristic makes the Japanese case a valuable natural experiment. However, to fully exploit it, three preliminary questions need to be settled: First, which deficiencies rendered the traditional system inimical to economic growth? Second, how did the Meiji government address these obstacles and did successful administrative reform *really* foster economic development? Finally, given the reforms' record varied across regions, how can this variation be used to construct quantitative measures of reform success for analytical purposes?

2.2 Weaknesses of the traditional administration

2.2.1 Lack of suitable incentives at the national and local levels

The traditional system exhibited weaknesses both at the structural and the organizational level. The former were due to the maintained territorial and administrative fragmentation after the civil war period: Given the *shôguns'* inability to secure the monopoly of power, domain independence was fiercely protected by the feudal lords and fostered by the Tokugawa in a *divide et impera* fashion. Administrative procedures therefore continued to follow regional customary law

[3] Among the Meiji leaders, there were divisions of opinion whether to put more emphasis on economic or on military development goals. Still, the need to improve administrative performance was less contested as it was beyond doubt that military and economic strength relied on the government's ability to ensure stable sources of revenue and to assert its monopoly of power. Moreover, there is evidence that transferring competencies to a strong central government headed by representatives from the domains which had supported the anti-*shôgunate* coalition was conceived as a means to mitigate cleavages between the powerful Restoration domains, which in turn reinforced the case for administrative restructuring. For further sources and references see: Ramseyer and Rosenbluth (1998), McLaren (1916), DeVere Brown/Hirota (ed.), The Diary of KidoTakayoshi (1983), Jansen and Rozman (1986).

which arguably increased transaction costs in inter-dominial economic activities.[4] Potential benefits from over-regional trade and division of labor were accordingly hard to realize, and although Tokugawa accounts probably exaggerate when they claim that perpetrators who set foot on another clan's land risked their lives, fragmentation shaped economic exchange and travelers sometimes spent hours having their belongings searched or waiting for permission to cross-dominial borders (Matsushita 2006: 105 Steiner (1965): 50).

In addition to the problems at the national level, the feudal administrations also suffered from deficiencies at the local level, especially from organizational inertia and resistance against occupational upward (or downward) mobility. At best, these conditions gave officeholders few incentives to perform; in more adverse cases, they provided the ground for impenetrable local patronage networks, administrative discretion and corruption. These problems too reflected the Tokugawas' attempts at preserving the existing social order: To start with, *samurai* status being a necessary condition for assuming civil service positions, *shôgunate* legislation which restricted the number of each lords' armed vassals had virtually eliminated occupational mobility in and out the administrative apparatus (Yamamura 1974: 70–85).[5] The absence of merit-based staffing and remuneration policies was further compounded by the fact that traditional Japanese law prevented the *non-samurai* population from bypassing the magistrate's office and appealing directly to a higher authority. Although exceptions were made in, e.g., the case of a local magistrate abusing his position, villages daring to complain could be punished for *esso*—illegal bypassing—or simply for "causing unnecessary trouble to the authorities" (Befu 1966: 29,43). On the other hand, bureaucrats themselves had nothing to gain from reporting illicit behavior of colleagues or subordinates because the traditional law maintained the Confucian principle of *renza*, the common responsibility of groups (Steenstrub 1996: 153). As a result, both potential whistle-blowers and superiors risked being punished for crimes they merely reported. Reprimands were accordingly rare and it was far less exceptional for public service positions to remain in one family for generations than being the object of frequent promotions or dismissals.

Incentive problems were reinforced by the fact that the low- and middle-ranking posts of local administrators were unattractive from a monetary perspective: Both according to contemporary sources and work from present-day scholars, the salaries of petty officials like clerks were hardly sufficient to cover the necessary daily living expenses of a family (Murakami 1997: 10–20; Nishizawa 2004: 195–198). Higher-ranking administrators for their part were vassals of the Tokugawa or the domain

[4] Specialization and national market integration where even more directly impeded by *shôgunate* strategies to prevent alliances of opposing lords (e.g., by deliberately neglecting the maintenance of certain transit and trading routes) and by numerous *daimyo* who took recourse to measures like the implementation of domain monopolies in the supply of regional products to increase the economic leverage of their fiefs.

[5] This pattern was admittedly thwarted toward the end of the Tokugawa period when the need to counter the increasingly palpable threat of Western aggression created demand for specialist services in the *shôgunate's* and the *daimyos'* bureaucracies. Yet, advancement in public service careers continued to be a matter of the putative office holder's birth rank and personal ties to the feudal lord rather than his individual ability or achievement (Izumi (2001), p. 15).

lords and had *ipso facto* a legal claim to fixed rice stipends the size of which depended on birth rank but not on the office occupied. If they held official functions, they received a salary, *yakudaka*, which varied according to the post (*yaku*) assumed. However, *yakudaka* were paid only insofar, as they exceeded the hereditary claims[6] and the marginal revenue of assuming office was as low as the losses of a bad magistrate who was removed from his functions and continued to receive his full stipend (Nishizawa 2004: 191). That both the lack of social permeability and the failure to impose meritocratic principles perceptively reduced institutional quality was evident to contemporary observers. A popular handbook for prospective civil servants in the mid-Meiji period for instance summarized the shortcomings of the feudal bureaucracies as follows:

> Due to their despotism, the policies of the Tokugawa had become deeply flawed: Usually the offspring of families who had acquired particular deeds in the distant past were appointed in a hereditary fashion. As a result, pedigree was of central importance while ability and achievement did not matter at all […] Hence, it is now time to establish recruitment rules which address this ancient predicament and prevent incompetent, degenerated and corrupt individuals from assuming public offices so easily (Ôtsu 1881: 8,22).

This observation is further corroborated by quantitative research of present-day historians who found that among the local magistrates who had been officially reprimanded or punished at some point of their professional lives, the fraction of administrators who held hereditary positions was disproportionally large (Izumi 2001: 17–20; Nishizawa 2004: 55).

2.2.2 Problems of asymmetric information

In many parts of Japan, the unfortunate incentive structure was aggravated by difficulties to supervise and control local officials. Again this was due to specific feudal regulations. In particular, the *daimyo* had been subjected to the so-called *sankin kotai* system of alternate attendance which required them to spend every other year in Edo, the *shôgunate's* capital. With few exceptions, local administration reflected this national pattern in the sense that samurai were obliged to settle in the capital of their home provinces. Since no high-ranking samurai was to permanently take office in a rural area, the district magistrates, commonly called *daikan*, typically left everyday village administration to the village headmen and other local elites unless the need to secure the district's tax revenue or to enforce a minimum of public order required direct intervention (Befu 1968: 303–312). That this settlement created undesired economic side effects was especially apparent in the area of taxation: As the total staff at a *daikan's* office rarely exceeded 20–30 persons, local magistrates were unable to arrange for frequent controls of all villages under their

[6] Thus, a *samurai* official with a stipend of 100 *hyo* and a *yakudaka* of 150 *hyo* would be compensated only for the difference between his stipend and the *yakudaka* and therefore receive an annual remuneration of 50 *hyo*.

jurisdiction.[7] Not surprisingly so, land and tax registers were far from even approximately capturing the extent to which new lands were cultivated and existing ones improved (Ravina 1998: 53). Accordingly, the distribution of individual tax loads was unequal and tax assessments or undefined property rights a frequent source of dispute. The economic harm of administrative fragmentation was—however—not confined to the increased transaction and contract costs that resulted from the erosion of the higher authorities' capacity to act as informed conflict-solving entities. It also stemmed from downright illicit activities of local magistrates and village elites: As discussed before, administrative tasks were carried out by a number of semi-autonomous actors, all of whom had little incentive to report illicit activities. Hence, in cases that required the interaction of different entities (e.g., the taxation of newly cultivated land) village elders, clerks and local magistrates formed a hierarchy of successive monopolists with every actor having the power to hush up mistakes or to collect bribes without worrying about intervention from higher-ranking officials. In economic theory, this constellation increases both the prevalence of institutional failure and the equilibrium bribe level of corrupt officials.[8] Although the number of surviving (and sufficiently detailed) records is limited, this pattern seems to fit *shôgunate*-era reality: In plain numbers, historians have estimated that embezzlement and underreporting of newly cultivated land left the Tokugawa and the *daimyo* with annual tax losses of no less than 15 % (Tacke 1975: 52). Not less common was the practice of "Nawa-gokuro," i.e., the undermeasurement of arable land, a custom the higher authorities would even tolerate to a certain extent (usually up to 20–30 %) be it as a means of tax alleviation be it because of their incapability to supervise local officials in remote areas (Nakamura 1966: 53–56). This quantitative result is supported by numerous, often vivid qualitative descriptions: The particularly elaborate accounts of unrests in Omi province in the late 1830s reveal for instance that villagers had been expected to pay bribes first to the surveyor's advance guard and then—moving up the hierarchy of successive monopolists—to his clerks and finally himself. Having paid the survey team a total of 1000 *ryo*—i.e., the monetary equivalent of 1000 *koku* of rice or 1 % of the average agricultural yield of a small feudal domain—at least one village succeeded in preventing the taxation of all its newly cultivated agricultural areas (Ono 1932: 275–289). Moreover, in 1866 toward the end of the Tokugawa era, a *daikan* and his subordinates are reported to have spent about 30 % of the taxes from their jurisdiction in Shinshû province on feasts, alcohol and personnel luxuries (Nishizawa 2004: 203).[9]

Summing up, two adverse effects of the traditional government system stick out: At the national level, administrative and territorial fragmentation precluded large-scale investment in growth-promoting public goods, impeded market integration

[7] Even if a *daikan* devoted the entire harvest season (i.e., about 3 months) to the inspection of arable lands and actual revenue, he was left with no more than one to three days per village—the time required to travel not included.

[8] For an overview of these models see: Machlup and Taber (1960), Shleifer and Vishny (1993).

[9] Even in less extreme cases, the fact that peasants routinely paid 2.5–3 % in excesses of the official tax rate to win the favor of local officials caused not even the slightest surprise among contemporary observers (Mizu no Tomenaga 1980/1981: 8–9).

and reduced the potential for *Smithian* growth. At the local level, asymmetries of information and unfavorable incentives stripped the higher authorities off the ability to universally enforce proper administrative conduct which arguably increased the prevalence and the degree of institutional failure. Finally, it should be noted that the historical sources suggest that harmful network structures emerged only if influential traditional cliques within the bureaucracy managed to resist external supervision and control. How the Meiji leaders attempted to break these ties and why their success varied across regions will be discussed in the next two sections.

2.3 The Meiji administrative reforms

After the Meiji Restoration, structural reform toward national unification under a strong central government proceeded rapidly: In summer 1869, the *daimyo* had—in most cases voluntarily after having been granted compensations—returned their lands to the Emperor. Two years later, in 1871, the domains were completely abolished and replaced by prefectures. Purely administrative units, the latter were stripped off the domains' traditional autonomy and ruled by centrally appointed governors—from 1878 assisted by elected assemblies with advisory competences. Change on the town and village level proceeded along the same lines. From 1872 to 1888, a series of laws was issued according to which towns and villages lost their role as natural administrative units to newly defined districts (*ku*). Institutions like the village assemblies survived the administrative transition but lost most of their self-governing functions.[10] *De iure*, the laws of 1869 through 1888, thus put an end to the traditional system's administrative fragmentation. This structural improvement notwithstanding the *de facto* modernization of administrative procedures necessitated improvement of civil servants' incentives, supervision, education and training.

To reach this goal, the Meiji leaders relied on a policy of strategic replacement and staffed key positions in the administration of the newly created prefectures with samurai from the domains which had led the anti-Tokugawa alliance in the Restoration War. Providing a prompt, relatively inexpensive means to break informal networks and to secure the central government's monopoly of power this strategy gradually turned into the general rule that civil service positions should be assigned to candidates who were not natives of the prefecture where they held office. Later—not the least due to the growing popular demand for political participation and government transparency—this ad hoc step toward administrative rationalization was supplemented by efforts to commit officeholders to universal codes of conduct and to institutionalize the merit principle in the staffing of public positions.[11] Once the reforms of 1869 till 1888 had provided Tokyo with the means to impose uniform standards and to reward compliance, strategic staffing thus allowed the central to dissolve local cliques and to redirect officials' loyalties and career perspectives from the traditional *daimyo*-domain to the *Tennô*-Japan level.

[10] For details on the return of the domains and the reforms of the prefectures, towns and villages see Jansen and Rozman (1986: 103), Steiner (1965: 24–52).

[11] For details see: Spaulding (1967) and Motoyama (1997).

Important to note, personnel restructuring moreover came along with far-reaching legal reforms that centered on the individual and abandoned the traditional principle of group responsibility. The new regulations required, for example, both high-ranking officials and petty bureaucrats to submit their records to internal audits and revisions of penal law significantly increased precision in the application of malpractice laws. Both in Tokugawa and early Meiji penal law, corruption was merely defined as "officials accepting bribes." By contrast, the Penal Code of 1882 and the "Regulations on the Personal Conduct of Officials" of 1887 bound superiors and judges to clear instructions by stipulating that:

> Officials may not receive in connection with their official functions any present from others, be it as an acknowledgment of service rendered, as a fee, or in any other name or under whatever pretext.

For high-ranking bureaucrats, the structural reforms of the 1870s created an additional element of control when the government installed elected assemblies at the prefectural and local levels. In fact, the power of these assemblies was limited to advisory functions with no direct influence on the legislative process. They nevertheless constituted a supervision mechanism in the sense that even if a governor could veto any of the assemblies' decisions, he had to file detailed reports in order to obtain the Home Ministry's permission to do so.[12]

Not only Civil servants' incentives were modified by the deterrent effect of closer supervision and higher risk of punishment, but the government also sought to increase commitment by installing a system of rewards and internal competition. For one, salaries in the Meiji civil service compared favorably both with the *yakudaka* of *shôgunate* times and the wages of comparable private sector employments.[13] In the early years of Meiji, the incentive effect of high salaries was reinforced by the fact that individuals whose education qualified them for public offices were mainly drawn from the former samurai class. As samurai privileges like the claim to hereditary, stipends were revoked at the same time and since educational reform and the blurring of status barriers allowed more and more commoners to enter the administration, officials' salaries—and the risk of losing them in case of incompetency or misconduct—became much stronger a stimulus than the traditional *yakudaka.*[14] Furthermore, the government rewarded outstanding performance on the job. Governors who took office in prefectures renown as rebellious received special allowances and preferential

[12] Imperial Rescript on the Prefectural Assemblies, chapter 1, article 5 and chapter 4, article 33–34., in: McLaren, Walter Wallace, Japanese Government Documents, Yokohama (1914, p. 273).

[13] Data from the 1880s for instance show that the median rank and median annual income of civil servants who held an official rank were 14th and 300 Yen, respectively, which was more than twice the wage of a manager at the private Mitsui bank (Kasuya 2005: 224).

[14] Incentives in the civil service were increasingly shaped by inner-administrative competition. This process was initiated by the revoking of samurai privileges which caused birth rank and local family ties to lose their role as criteria of selection to merit-based indicators: Immediately after 1868, mostly activists of the Restoration movement were appointed. From the 1870s with the establishment of the first law schools till the 1887 general examination ordinance, appointment was made a matter of educational attainment and performance in standardized national examinations. As a result, roughly one-third of government positions was occupied by non-samurai newcomers after the first two decades of the Meiji period.

promotion to high government posts (DeVere 1966: 220; Baxter 1994: 199). Lower-ranking officials could earn yearly bonuses up to the amount of their monthly salary.[15] Finally, capitalizing on caste mentality and symbols deep-rooted in collective consciousness, civil service was idealized as the continuation of *samurai* values and lifestyle. In certain cases, exceptional achievement was therefore rewarded by elevation to the prestigious *chokunin* rank (personally granted by the Tennô) and uniforms of soldiers or police officers included swords, the ancient emblem of *samurai* status (Westney 1987: 53, 76, 95).

As the Meiji policies come close to a textbook example of Weberian bureaucratic modernization, it is straightforward to ask whether they did indeed translate into sizeable improvements of institutional quality. The answer is an unequivocal "Yes" and is supported both by testimonies of Western and Japanese observers and by quantitative evidence. Thus, while sharply criticizing the central government on various occasions, prominent Meiji intellectual and politician Nishimura Shigeki contributed an essay to the prestigious political journal of the "Society of the 6th Year of Meiji (*Meirokusha*)" in which he claimed that

> They [i.e. the reforms] have destroyed such despotic measures as forced loans, compulsory assistance to neighboring villages and property confiscations that were practiced during feudal days.[16]

With particular regard to corruption, twentieth-century historian McLaren further notes that

> Corruption of this kind [i.e. local government officials accepting bribes] was, however, easily enough dealt with when the central Government chose to suppress it, and after a few years, it had practically disappeared. (McLaren 1916: 129)

From an empirical point of view, the benevolent assessment of contemporaries and historians is corroborated in that the number of popular protests and uprisings prompted by corruption or malpractice of local government officials steadily declined after the restoration. Figure 1 depicts the annual number of incidents registered by the authorities in the period of 1835–1900. While increasing surveillance after the creation of rural police forces and the implementation of particularly unpopular policies (e.g., the radical deflationary policy of Minister of Finance Matsukata Masayoshi in 1884) caused a rise in the incidence of protests recorded, corruption-induced incidents exhibit a strongly negative time trend in both absolute and relative terms.[17]

[15] Law on the Duties of prefectural officials, issued July 27, 1876, article II,2, in: McLaren, Walter Wallace, Japanese Government Documents, Yokohama (1914, p. 268).

[16] The corresponding article can be found in: Nishimura, Shigeki, 'On Change', Meiroku-Zasshi vol. 43 (Nov. 1875).

[17] Corresponding records were kept on the prefectural and local levels and listed the dates, location, the reason and the number of persons involved in a protest. Thanks to the work of Aoki these data have been cataloged both for the Tokugawa and Meiji periods. In the present graph, I classified an uprising as caused by malfeasance if "officials' corruption," "despotic administration" and the like were reported as the causes of the disorder. For my data see: Aoki, Kouji, Hyakusho Ikki no Nenjiteki Kenkyu, Tokyo (1974), statistical appendix; Aoki, Kouji, Meiji Noumin Soujo no Nenteki Kenkyu, Tokyo (1967).

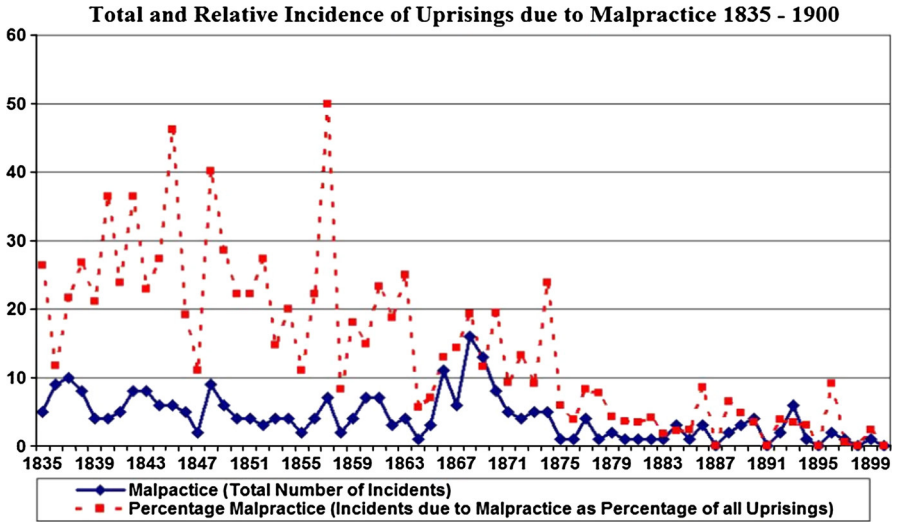


Fig. 1 Total and relative incidence of uprisings

2.4 Regional variation

The general trend notwithstanding reform success varied grossly across regions. The following section prepares for the statistical analysis by sketching the reasons behind this mixed record. Considering these factors is imperative because the qualitative historical evidence shows to what extent the survival of traditional networks and administrative sleaze jeopardized economic development after the restoration. Moreover, understanding *why* the bureaucracies of certain prefectures were more impenetrable than others provides the main ingredient for the econometric analysis namely a quantitative measure of "reform success."[18]

The reasons of retarded administrative rationalization stemmed largely from path dependency and the territorial arrangements set in place by the Tokugawa in the seventeenth century. After their victory in the civil wars, they had increased their fief to about 25 % of the country and controlled economic and strategic centers like Osaka, Kyoto and Nagasaki, large ports and the main producing regions of raw materials like silver, iron or copper. To consolidate their power, the *shôguns* encircled their territories by a *cordon sanitaire* of *daimyo* (the *fudai daimyo*) who had supported them during the wars. Members of the anti-Tokugawa coalition, the *tozama daimyo*, were ousted from central Japan and pushed to the periphery where—albeit after having lost territories to the *fuda i*—they continued to rule their ancestral fiefs.

[18] One particularly valuable source to investigate the experiences of the different prefectures is the *Meiji 15nen/16nen—Chihô-Junsatsu-shi Fukumeishô* (*CJF*), a collection of reports submitted to the Meiji government in 1882/1883 which gives detailed accounts of the economic, political and administrative state of affairs in every of the then existing Japanese prefectures.

Since entry to the *samurai* class was strictly regulated, this constellation gave rise to a distinct pattern with the reduced *tozama* territories featuring very high densities of warrior population, whereas the Tokugawa and the *fudai* commanded relatively large areas in relation to the number of vassals whose rank qualified them for administrative posts (Izumi 2001: 20–21). As a result, *shôgunal* ordinances that *de iure* obliged *samurai* to reside in the castle town of their home province were *de facto* unenforceable in the jurisdiction of many *tozama* lords. In these areas, a class of so-called *goshi* ("*samurai farmers*") emerged who combined social characteristics of warrior and peasant status and frequently assumed low- to middle-ranking posts in the local administration.[19] Yet, while the *goshis'* in-depth knowledge of regional particularities and close personal ties to their native communities made them apt administrators, the *tozama* bureaucracies became much more vulnerable to the influence of patronage and personal networking than the *fudai* and Tokugawa administrations.[20] In the latter, the number of potential officeholders was so restricted that the district of a single magistrate easily rivaled the size of a small feudal domain which made close ties to single localities hard to establish. Although hereditary offices and limits to occupational mobility existed also in the *shôgunate's* sphere of influence, magistrates were therefore more often appointed from without and changed positions more frequently than in *tozama* territories. In other words: the transition from feudal retainers who ruled based on traditional authority to a class of salaried, professional civil servants was more advanced in the former Tokugawa and *fudai* regions. By contrast, the landed samurai of the large *tozama* domains was better able to shield their stakes from the central government's attempts at control and merit-based recruitment.

Adding to that, reform activities reflected political currents within the Meiji oligarchy and long-standing personal relationships made it hard for the leaders to confront former associates in the local bureaucracies of their native (*tozama*-)domains. What is more, mistrust and internal cleavages drove them to jealously guard the stakes of their home provinces from external intervention be it from the central government or other prefectures.

Mixed reform success was thus largely predetermined by the different prefectures' proneness to develop strong personal networks within their administrations. That the survival of these traditionally status- and loyalty-based structures proved fatal in socioeconomic terms is illustrated by a voluminous body of historical sources. The following testimonies of contemporary observers are intended to provide a flavor of the extent and the actual mechanisms behind retarded administrative rationalization. As a starting point, they take the case of Kagoshima prefecture (formerly Satsuma domain) where substantial reforms occurred only after the Rebellion of Saigo Takamori in 1877 and continued to be a cumbersome business for a long time afterward. As late as 1898, a Meiji-era scholar noted:

[19] For instance, *gôshi* were granted the right to wear arms and could be appointed to a limited number of public offices. Yet, their hereditary stipends were typically so small that they had to rely on additional sources of income—mostly agricultural activities.

[20] Regional expertise allowed them to better assess potential tax yields and to react early to popular unrest. Full blown peasant uprisings were accordingly far rarer events in *gôshi* controlled areas than in the jurisdiction of the *shôgunate* and its allies. Matsushita (2006: 200), Paik et al. (2012).

> The influence of former samurai is admittedly high in all the South-Western parts of the country. Yet this is true to the extreme in the former Satsuma domain. Here, public affairs no matter how big or small they may be lie entirely in the hands of samurai cliques and among the deputies to the national parliament and the prefectural and municipal assemblies, the prefectural, district and village officials, the policemen, judges, clerks (…) teachers and even among the students of middle and high schools 99 % are of samurai descent (Honmi 1898: 37).

The detrimental influence of local samurai elites is equally stressed in a report from the *Chiho-Junsatsushi Fukumeisho* (*CJF*):

> Only after the Satsuma Rebellion has the prefecture been integrated in the central government's sphere of influence. (…) In a few years, many projects were initiated and he [i.e. the new governor] managed to reestablish the public order. However, once he struggles to handle things in *a way that wins him the favor of the powerful clans in the prefecture, he is unable to undertake decisive action*. CJF I: 759)

The struggle against local networks and institutional encrustation was not confined to Kagoshima. An internal government report about Ishikawa prefecture—like Kagoshima a former *tozama* domain with high ratios of native officeholders—noted the following:

> They (i.e. the local samurai) have not yet woken from their feudalistic dreams and do not know any other way of making their livelihood than turning to the prefectural government or their former *daimyo* to request aid and financial relief. (…) The middle and lower ranking officials of the prefecture are mainly natives who (…) might know Ishikawa and in particular Kanazawa [i.e. the prefectural capital] but are largely ignorant of anything else. Due to their excessive provincialism they have no idea of the currents of the time and visiting the prefectural offices *today* is like inspecting the *daimiate* administration of the *feudal past*. Furthermore, one hears that Ishikawa officials loathe collaborating with appointees from other prefectures and that they had the bad habit of forming cliques. (…). (CJF I: 517)

How serious a threat conservative elite attitudes presented to long-term growth perspectives is stressed by observers in Kagoshima and Ishikawa alike. For instance, a Kagoshima teacher attributed the prefecture's failure to take part in Japan's rapid industrialization to social inertia and explicitly stressed the fact that "The former samurai of Satsuma show not the slightest interest for applied sciences and their sons are particularly weak in the field of mathematics" (Matsushita 2006: 189).[21] And the Ishikawa report in the *Chiho Junsatsu-shi Fukumeisho* denounces the harmful impact of successful lobbying activities by influential samurai networks:

[21] Indeed his observation seems to the state of affairs in Kagoshima well enough as both Meiji era statistics and studies of present-day historians show that the prefecture ranked among the last where school attendance and educational outcomes were concerned (Ohkawa 1957: 24, Honmi 1898: 42).

> In order to promote samurai enterprises, large funds have been distributed and numerous businesses have been set up subsequently. Yet, most of these start-up firms are caught in a state of decline and stagnation with no hope of future recovery. This situation is due to the samurais' lack of perseverance and their strong spirit of dependence.[…] It is not sufficient to acknowledge that these factors (i.e. the lack of suitable incentives) determine the current state of samurai enterprises and that the easy availability of public funds and aid remains ineffective. Besides that, there is no denying that samurai in this region *use employment creation as a pretext to acquire government subventions and protection*. (CJF I: 518–520)

Things were hardly different in Kôchi prefecture, formerly Tosa domain, where according to the *Chiho Junsatsu-shi Fukumeisho* "the prefecture's veteran officials 'sternly opposed' all appointees from other prefectures and managed to render the implementation of central government ordinances utterly impossible" (CJF I: 741). Important to note, the report goes on and points out that these unfortunate circumstances started to change noticeably as the newly appointed governor successfully pushed for rapid personnel restructuring.

Other regions fared far better. In the Kantô plain for instance, the booming prefectures in the vicinity of Tokyo mostly consisted of former *shôgunate* territories and had experienced comprehensive, hardly opposed personal restructuring and merit-based appointments from the first years of Meiji onwards. In the 1880s, the fruits of these efforts became apparent as government reports found no fault with the functioning of the prefectural administrations but in the case of Kanagawa where the following—excusable—flaw was found:

> Compared to other regions, administrative procedures are somewhat slow, but upon noting this fact it has to be taken into consideration that the prefecture's engagement in trade and contacts with foreigners raise particular challenges for the administration. (CJF II: 1379)

The development of many former Tokugawa and *fudai* territories not only illustrates that early administrative rationalization fostered economic activity by facilitating specialization and trade, but also shows how the reforms built on the higher degree of occupational mobility in the *shôgunate* territories where the need for specialist services had paved the way to administrative positions for low-ranking samurai and commoners even before the restoration. Given that social permeability in these regions tended to be higher, the successive abolition of samurai privileges and occupational barriers for commoners was met with relatively little resistance as former vassals who lost their stipends and public offices frequently turned to business activities without the administration. Capitalizing on this historical precondition, administrative modernization could then provide the final impetus to overcome long-established institutional deficiencies as can be seen, for example, in the case of Tochigi prefecture: Here, municipal bureaucracies were ridden by endemic corruption and collusion between popularly elected local officials and village elites. Since patronage networks constituted a permanent cause of conflict and reduced both the prefectural and central government's capacity to enforce their

monopoly of power, the governor of Tochigi prefecture temporarily suspended all elections for administrative offices and replaced the officials in question by appointees from other prefectures. The corresponding reports from the *Chiho Junsatsu-shi Fukumeisho* reveal that this radical step did indeed perceptively reduce the incidence of officials' misconduct and illicit activities. Yet they touch on two equally noteworthy details, namely first that the governor invested time and effort to obtain the approval of the prefectural assembly and second that the central government *nevertheless* required him to officially *justify* his infringing upon the general election procedures (CJF II: 928–929).

The bottom line of these testimonies is clear: The replacement of feudal bureaucrats with centrally appointed civil servants was no sign of central government discretion and despotism. It was an indicator of ongoing administrative modernization. In addition, all accounts highlight the fact that while administrative rationalization proved conductive for economic growth, the replacement of traditional structures, let alone the implementation of development plans like the *Fukoku Kyohei* policies, was possible only after local personal networks were stripped off their influence.

In terms of analytic procedure, this observation creates exciting avenues for quantitative research. Given that the Meiji leaders' instrument of choice to suppress regional particularism consisted in strategic staffing policies *and* seeing that this strategy apparently served their ends, it is possible to quantify early reform success by the extent to which the government enforced personal restructuring in the different prefectural administrations. In order to assess the economic benefits from administrative rationalization, the statistical analysis then only needs to construct variables which capture the Meiji leaders' recruitment policies and to relate these measures to indicators of economic development.

3 Econometric analysis

3.1 Data and estimation strategy

The first indicator of strategic staffing is the variable, Native_{tj} which denotes the share of native employees (i.e., the number of persons who held office in the prefecture where they had been born) in prefecture j in year t. Native_{jt} can therefore be interpreted as a measure of the administration's openness to over-regional influences. To refine the picture drawn by this general indicator, I consider the distribution of native and non-native officeholders across the administrative hierarchy. This extension is necessitated by the fact that—as shown by, e.g., the experiences of Ishikawa and Kôchi prefectures—high-ranking native officials could and *did* use their position to influence regional policies in favor of traditional interests.

To account for these factors, a second variable, $\text{RatioNativeHigh}_{jt}$ (RNH_{jt}) was introduced to the model. It gives the ratio of the fraction of native staff in the top layers of the administration to the fraction of natives among all employees in prefecture j and year t. RNH_{jt} thus reflects the hierarchical distribution of foreigners and natives. Values of $\text{RNH}_{jt} > 1$ thereby indicate that native officeholders

clustered at the top positions, whereas $\text{RNH}_{jt} < 1$ implies that high-ranking posts were hold by foreigners and subordinate positions by natives.

This general rationale set out, I proceeded to use the two explanatory variables Native_{jt} and RNH_{jt} to account for inter-prefectural differences in three fundamental indicators of economic development. The indicators considered are measures of prefectural per capita income, business activity and financial development. Formally, the corresponding reduced-form model is captured by the following equation:

$$z_{jt} = \beta_1 \cdot \text{Native}_{jt} + \beta_2 \cdot \text{RatioNativeHigh}_{jt} + \gamma_{jt} \cdot X_{jt} + \text{cons} + \varepsilon_{jt}$$

z thereby denotes the value of development indicator i in prefecture j and year t. β_1 and β_2 capture the partial effects of changes either in the fraction of native officeholders or in their distribution within the hierarchy. X_{jt} is a vector of prefectural-specific control variables with their coefficients γ_{jt} and ε_{jt} is a white noise error term.

As the empirical implementation of the model is concerned, the estimation relies entirely on data obtained from primary sources and official statistics. The necessary information to construct the two central explanatory variables, Native_{jt} and RNH_{jt}, was from the so-called *Meiji Kan-in Roku* (=Lists of public employees during the Meiji Period).[22] Based on this evidence, I obtained Native_{jt} by counting the number of native administrators and dividing the result by the total number of prefectural officials.[23] RNH_{jt} was constructed accordingly by assessing the fraction of natives

[22] The "*Meiji Kan-in Roku*" were published from the early years of Meiji onwards and contain comprehensive monthly lists of all administrative posts in the different prefectures together with the names and native provinces of the corresponding officeholders.

[23] An important advantage of "*Meiji Kan-in Roku*" is the fact that it proves relatively robust against changes of prefectural boundaries in the early Meiji period. In fact, a central government act known as the *Haihan-Chiken* first transformed the feudal territories of the Tokugawa period into 72 prefectures, a number which was subsequently reduced to 47. Consequently, the variable "Native" might misstate the fraction of native to non-native officials if officeholders remained in their ancestral fief but were no longer recorded as natives when the territory in question was amalgamated into another prefecture. Fortunately, though, the historical circumstances and certain characteristics of the *Kan-In Roku* mitigate potential bias from this source. For one, there is no evidence that changes in prefectural boundaries were more frequent in regions which had performed particularly well or badly before the Meiji Restoration. The direct effect of territorial modifications was therefore probably small. As the indirect effect of misstating the true number of natives in newly formed prefectures is concerned, it suffices to note the following characteristics of the *Kan-In Roku*: Officeholders were listed with their names and family names. Hence, by following backwards the career path of an administrator, I could verify his native province even if prefectural boundaries changed in the course of his professional life. The problem then came down to dealing with new entrants to the civil service. *Kan-In Roku* treats this staff always in a way that—to the extent old prefectures ceased to exist—natives of these territories were recorded as born in the prefecture to which the old province was annexed. Changes in prefectural boundaries therefore tended to slightly overstate the true fraction of native to non-native officials. Bias from this source is, however, likely to distort the expected statistical estimates downwards rather than upwards. The reason is as follows: Since changes in prefectural boundaries were frequently undertaken for efficiency reasons, they ought to have improved administrative performance after a certain time. Distortions from the misclassification of newly appointed officials also becoming sizeable with a time lag, prefectures with overstated values of "Native" were arguably already reaping the benefits from the efficiency gains after the geographical restructuring. Hence, if high realizations of "Native" indicated institutional backwardness, the positive impact of improved administrative performance after boundary changes will at least partially offset this effect in the econometric estimation. This implies that *if* the statistical analysis confirms the anecdotical historical evidence—i.e., if the coefficients of "Native" are negative—these estimates would constitute a lower bound for the true relationship between administrative backwardness and economic development.

among the high-ranking officials (i.e., officials who held the so-called *chokunin* rank)[24] and dividing the resulting fraction by Native_{jt}.

Turning to the set of development indicators, a preliminary problem to overcome consisted in the fact that statistical data from the early Meiji era do not allow us to construct annual GDP estimates at the prefectural level. To introduce a measure of regional per capita income, I used the following approximation: From the first post-restoration years onwards, the central government levied national taxes on consumption goods and services like alcohol, sugar, soy sauce, tobacco, postal fees and transport. Data on each prefecture's per capita tax yield have been collected and published in the Statistical Yearbook of the Japanese Empire (*Nihon Teikoku Toukei Nenkan*). These comprehensive time series show that neither the set of the targeted goods and services nor the tax rate varied across prefectures. What is more, the items which were subject to national taxes were consumed across the entire country and can reasonably be assumed normal or superior in private consumption (Ohkawa 1957, statistical annex). Thus, the corresponding demand should have been sufficiently elastic to ensure that heterogeneity in regional per capita income translated into differences in the tax yield. Following this rationale, the econometric model approximates interprefectural income differences with the variable RelativeTax_{tj} which denotes the per capita yield of consumption taxes in prefecture j and year t divided by the average national tax yield. By construction, realizations of $\text{RelativeTax}_{tj} > 1$ therefore indicate above-average tax revenue—and presumably higher per capita income—in prefecture j, whereas $\text{RelativeTax}_{tj} < 1$ suggests low tax yields and per capita income.[25]

In addition to the main indicator RelativeTax, I consider two auxiliary measures of development. The first builds on the notion that besides per capita income, the institutional heterogeneity captured by Native_{jt} and RNH_{tj} probably also influenced regional business activity. To investigate this possibility, I relied on the variable Startup_{jt} which can be derived from the *Shuyo kigyo no keifuzu*, a compilation of corporate genealogies edited by the business historians Yagura and Ikushima. This source contains the histories of more than 1,000 firms that were established during the Meiji era and listed on the Tokyo stock exchange in 1984.[26] Based on this

[24] As mentioned before, the chôkunin rank was exclusively awarded by the Emperor and was restricted to high-ranking officials.

[25] The reason for relying on relative rather than their absolute tax yields stems from the fact that—while taxation regulations were centrally determined—the set of taxed items and the corresponding rates changed over time. As a result, absolute tax revenue is not directly comparable for different years which rendered its use unsuitable for the present analysis.

[26] *Shuyo kigyo no keifuzu* arguably constitutes the most comprehensive record of business activity in the early Meiji period and is the subject of a recent in-depth analysis by (Tang 2011). His work features a prefectural-level database with the annual number of company start-ups from the first year of Meiji onwards.

unique information, I was able to introduce $Startup_{jt}$, i.e., the number of newly established businesses in prefecture j and year t as an indicator of private entrepreneurial activity.[27]

The set of dependent variables is completed by a measure of financial market development: The variable $Lendingrate_{jt}$ gives the average annual lending rate on standardized loans collected by banks in prefecture j and year t. The corresponding data have been collected by the prefectural statistical offices and provide the percentage lending rates on secured private loans of 100 Yen. From the early 1880s onwards, these lists are published on a yearly basis in the *Fuken Tôkeisho*, the "Statistical Yearbooks of the Japanese Prefectures." Starting in 1884, coherent monthly series of standardized 100-Yen loan rates appear in the series *Ginkokyoku Nenpo* published by the bank of Japan. The particular appeal of this evidence stems from the fact that lending rates reflect region-specific risk perceptions, transaction costs and expected costs of contract enforcement. Hence, they provide an opportunity to get an idea of the contemporary agents' assessment of institutional quality. Unfortunately, though, this advantage does not come without shortcomings. Given that data on lending rates are unavailable prior to the 1880s, the period of most fervent restructuring (i.e., the mid-to-late 1870s) remains uncovered. In consequence, the years where institutional differences were probably most acute drop from the econometric analysis which might cause the effects of $Native_{jt}$ and RNH_{jt} to be downwards biased.

To close the econometric model, I introduced a set of control variables which account for regional heterogeneity in non-institutional determinants of economic growth. In particular, the corresponding covariates include measures of the economy's sectoral structure, pre-Meiji prefectural output and human capital formation. In addition, I controlled for population density, distance to Tokyo and Osaka, the country's main industrial centers, each prefecture's access to ports and periods of political unrest and wars. Regarding these variables, it should be noted that the need to account for the sectoral composition of the economy stems from the fact that late Tokugawa and early Meiji agricultural development varied across prefectures. Hence, whereas backward cultivation techniques survived in some regions, others featured a highly productive first sector or became centers of burgeoning proto-industrial activities. However, while agricultural progress and proto-industrialization arguably increased prefectural per capita income, both factors might be underestimated by my development indicators. The reasons are first that agrarian households probably used part of their production for personal consumption (reducing "Relativetax"), second that agricultural businesses, not to mention small-scale craft producers and manufacturers, were less likely to be registered at the stock exchange (driving "Startups" downwards) and finally that

[27] Considering this indicator, critical spirits might be concerned with the fact that $Startup_{jt}$ reflects only a part of overall economic activity as businesses which failed or were not listed at the Tokyo stock exchange in 1984 are not included in the database. However, issues of survivor bias are mitigated by the inclusion of defunct firms whose assets were transferred to direct successors (Tang 2011). Moreover, since favorable institutional conditions like administrative rationality should have similar effects both on the establishment and on the long-term success of businesses, this shortcoming does not seem to jeopardize the overall validity of the development indicator $Startup_{jt}$.

agricultural and proto-industrial finance frequently relied on other sources than credit from regular banks.[28] To account for these possibilities, I constructed the control variable "AgriShare_{jt}" which denotes the fraction of agricultural households in each prefecture and thereby measures the importance of agriculture for the local economy.[29]

The additional control variables were constructed as follows: *Gakko,* the measure of human capital formation is derived from the nine volumes of the *Nihon Kyouiku-shi Shiryou* edited by the Ministry of Education in the early 1890s. This source reprints the results of a government survey which specified the number of primary schools that existed before 1872. Using this information, I approximated prefectural human capital by the adjusted ratio of primary schools to inhabitants.[30] Regional differences in pre-Meiji per capita income are captured by a second control variable, *Kokudaka,* i.e, the product of a territory (e.g., a domain) as assessed for taxation purposes.[31] The construction and interpretation of the remaining control variables is straightforward: The variable Distance_j gives the kilometer distance between the capital of prefecture *j* and Tokyo or Osaka depending on which of the latter two was closer. Popden_{jt} renders the population density of prefecture *j* in year *t* and relies on data from the *Kokudaka and Population Tables of the Domains and Prefectures* and the Statistical Yearbooks of the Japanese Empire. Landlocked_j and Conflict_t are binary variables, the first of which takes the value of one if prefecture *j* had no access to ports (i.e., if costal shipping was impossible) and zero otherwise. Conflict_t

[28] In fact, there is a substantial literature (see p.e. Francks 2002; Howell 1992; Smith 1988) showing that agricultural and commercial development in rural areas relied strongly on better-off merchants and land owning families (sometimes referred to as "Meibôku") who combined traditional cultivation and landowning activities with money lending, trading or the establishment and operation of putting-out systems. If those agents were better able to assess the prospects of investments and the solvency of their local clientele, they might have picked the "plums" from the prospective borrowers while leaving more risky enterprises to other financial intermediaries. Hence, ordinary banks might have increased their risk premia which in turn would cause Lendingrate to be higher in the corresponding prefectures.

[29] "AgriShare" arguably also reflects agricultural productivity: As shown in the seminal work of Nakamura (1966), agricultural innovation tended to be labor-intensive throughout the early Meiji period such that prefectures that derived most of their per capita income from agriculture should exhibit high realizations of "AgriShare." Hence, "AgriShare" is unlikely to underestimate the relative importance of agriculture in highly productive regions as it would be the case if innovation had been primarily labor-saving.

[30] Note that the necessity for modifications in the number of primary schools stems from the fact that historical records contain only data of private primary schools (shingakkô) and temple schools (terakoya). These schools were mainly frequented by commoners while most samurai had their children privately educated. Hence, measuring human capital formation by the unadjusted number of schools per capita is likely to bias the estimates in favor of regions with relatively large commoner populations. In order to correct for this problem, I proceeded as follows: Schooling rates are estimated to have reached about 40 % for commoners and 100 % for samurai. Using 1868/1869 census data, I identified regional samurai population p_{samurai} at the time of the Meiji Restoration. Based on these material, I multiplied the number of schools in each prefecture by a correction factor of $(1 + p_{\text{samurai}} = p_{\text{total}} \times (1 = 0/4))$. For an exhaustive account of measuring school attendance in pre-modern Japan see: Dore, R.P., Education in Tokugawa Japan, London 1984, S. 317–323.

[31] The *shôgunate* and the *daimyo* collected lists of their fiefs' taxable *kokudaka* output. These data can be found in *Hansei Ichiran* (A Summary of the Domain System) edited by the governmental department for historical research as well as in the official *Kokudaka and Population Tables of the Domains and Prefectures* issued in 1869 and hold by the Okuma Shigenobu online library of Waseda University.

controls for political turmoil and takes the value of one for the years 1877 (Satsuma Rebellion) and 1894/1895 (Sino-Japanese War). Finally, exogenous time effects are accounted for by the variable Trend, a linear time trend starting in the first observation year 1875.

3.2 Estimation and results

The basic specifications of the econometric model concentrate on the contemporaneous effects of Native and RNH on the three development indicators. The estimated relationship is stated in the following equation:

$$z_{jt} = \beta_1 \cdot \text{Native}_{jt} + \beta_2 \cdot \text{RatioNativeHigh}_{jt} + \gamma_1 \cdot \text{Gakko}_{jt} + \gamma_2 \cdot \text{Kokudaka}_{jt} + \gamma_3 \cdot \text{Distance}_j + \gamma_4 \cdot \text{Popden}_{jt} + \gamma_5 \cdot \text{Landlocked}_j + \gamma_6 \cdot \text{Conflict}_t + \gamma_7 \cdot \text{Trend} + \text{cons} + \varepsilon_{jt}$$

In terms of the direction of the partial effects of Native and RNH, the historical evidence suggests that high proportions of native staff were associated with inertia and delayed organizational reform. Consequently, the coefficient β_1 ought to carry a negative sign in the regressions on Startup and RelativeTax as institutional deficiencies discourage business activity and reduce regional per capita income. Concerning lending rates, the opposite holds because efficient administration reduces transaction costs and fosters mutual trust.

Proceeding to the expected relationship between RNH_{jt} and the development indicators, both the influence of established elites and the Meiji leaders' reliance on strategic recruitment need to be born in mind. In particular, in the early Meiji years, it was not uncommon for accomplished governors and their high-ranking staff to literally hop from one troubleshooting prefecture to the next. Regions that managed to *oppose* the central's attempts at administrative rationalization were accordingly unlikely to be headed by non-native governors and high-ranking bureaucrats. In combination, these considerations imply that small values of RNH_{jt} should be associated with institutionally advanced prefectures that exhibited high occupational permeability at all levels of the bureaucracy. In terms of econometric estimates, one would therefore expect the coefficient β_2 to be negative in the regressions on RelativeTax and Startups and positive in the regression on Lendingrate.

As it turns out, the expectations formed from the qualitative historical evidence are supported by the quantitative analysis. Table 1 shows the estimation results for the basic specification of the model. In order to take account of the dataset's panel structure, the simple OLS regressions in column (1) were subsequently subject to the following modifications: Column (2) introduces regional controls and panel robust standard errors, Column (3) further includes lacked values of the respective dependent variable, and Column (4) relies on the Arellano–Bond technique to address possible serial correlation.

No matter which estimation technique is chosen, the coefficient of Native almost always conforms to the expectations, that is, β_1 is negative and significant in the regressions on RelativeTax and Startups and positive in the regressions on Lendingrate. The results are less conclusive for RNH as many estimates end up insignificant albeit mostly carrying the expected signs. Also, concerning

Table 1 Effects of strategic recruitment on development indicators

	Dep. variable: RelativeTax				Dep. variable: Startup			
	OLS	Regional controls	Lagged values	Arellano Bond	OLS	Regional controls	Lagged values	Arellano Bond
Native	−0.2500***	−0.1480*	−0.0972**	0.0077	−0.2830*	−0.1060	−0.2370*	−0.4250
	(0.0817)	(0.0869)	(0.0412)	(0.0609)	(0.1670)	(0.1810)	(0.1160)	(0.5400)
RatioNative	0.0093	0.0305	−7.11e−06	−0.0089	0.0259	0.0346	−0.0215	−0.0659
	(0.0269)	(0.0264)	(0.0175)	(0.0212)	(0.0545)	(0.0550)	(0.0587)	(0.0971)
Relativetax_Lagged			0.7900***	−0.0987			−0.0009	−0.1860
			(0.0569)	(0.1210)			(0.0702)	(0.1310)
AgriShare	−0.6880***	−0.5420***	−0.1650	0.2280	−0.9550***	−0.9430***	−0.9670**	1.8690**
	(0.1110)	(0.1140)	(0.1260)	(0.3170)	(0.2240)	(0.2380)	(0.3970)	(0.8200)
Popden	0.0002**	0.0001	−3.26e−05	−0.0014***	0.0019***	0.0019***	0.0020***	−0.0021
	(0.0001)	(0.0001)	(9.44e−05)	(0.0005)	(0.0002)	(0.0003)	(0.0005)	(0.0038)
Kokudaka	−0.0009***	−0.0013***	−0.0003***		5.11e−05	−0.0003	6.16e−05	
	(0.0002)	(0.0003)	(8.72e−05)		(0.0004)	(0.0005)	(0.0002)	
Distance	−0.0008***	−0.0009***	−0.0002*		−0.0003	−0.0005**	−0.0003*	
	(8.33e−05)	(0.0001)	(8.34e−05)		(0.0002)	(0.0002)	(0.0001)	
Gakko	0.0693***	0.0716***	0.0072		0.0935**	0.1060**	0.0941***	
	(0.0192)	(0.0216)	(0.0105)		(0.0386)	(0.0445)	(0.0294)	
Landlocked	−0.0334	−0.1010**	−0.0144		0.0332	0.0018	0.0334	
	(0.0355)	(0.0416)	(0.0237)		(0.0711)	(0.0859)	(0.0638)	
Trend	−0.0013	−0.0004	−0.0003	0.0095	0.0006	0.0006	−0.0003	0.0627**
	(0.0020)	(0.0019)	(0.0009)	(0.0081)	(0.0040)	(0.0040)	(0.0045)	(0.0278)
Conflict	−0.0072	−0.0074	−0.00186		0.1990***	0.1980***	0.1980***	
	(0.0320)	(0.0312)	(0.0125)		(0.0660)	(0.0660)	(0.0700)	
Constant	−0.6970***	1.743***	0.437***		0.6710***	0.6550**	0.6940*	
	(0.1110)	(0.137)	(0.134)		(0.2330)	(0.2550)	(0.3810)	
Observations	723	723	646	557	801	801	767	687
R-squared	0.344	0.387	0.407	–	0.283	0.289	0.290	–

Table 1 continued

Dep. variable and coefficients	OLS	Regional controls	Lagged values	Arellano Bond
Lendingrate				
Native	0.3620	0.3450	0.4040	0.4500
	(0.6000)	(0.6210)	(0.3940)	(0.5480)
RatioNative	0.1680	0.2270	0.0062	0.0414
	(0.6550)	(0.6460)	(0.6850)	−1.400
Lendingrate_Lagged			0.7140***	0.3640***
			(0.0345)	(0.0485)
AgriShare	2.3090***	2.3590***	0.0143	2.8150
	(0.8030)	(0.8410)	(0.5660)	(2.8440)
Popden	0.0002	−0.0004	−0.0005	0.0002
	(0.0009)	(0.0009)	(0.0005)	(0.0027)
Kokudaka	0.0064***	0.0024***	0.0010**	
	(0.0014)	(0.0016)	(0.0005)	
Distance	0.0032***	0.0053***	0.0011***	
	(0.0006)	(0.0009)	(0.0004)	
Gakko	0.1520	0.1070	−0.0165	
	(0.1300)	(0.1460)	(0.0875)	
Landlocked	−0.1320	−0.6160**	−0.0144	
	(0.2350)	(0.2910)	(0.1730)	
Trend	−0.3760***	−0.3760***	−0.1170***	−0.2650***
	(0.0209)	(0.0205)	(0.0199)	(0.0347)
Conflict	1.5320***	1.5310***	1.4310***	
	(0.2590)	(0.2550)	(0.1380)	
Constant	15.300***	15.740***	5.0770***	
	(0.8780)	(0.9380)	(0.8350)	
Observations	553	553	516	445
R-squared	0.453	0.489	0.750	–

* Indicates significance at 10 %, ** indicates significance at 5 %, *** indicates significance at 1 % level

Lendingrate, β_1 and β_2 indicate the expected relationship between administrative rationalization and loan rates, but the estimates fail to reach statistical significance at the conventional levels. Admittedly a caveat, this finding does not come entirely unexpected as due to the discussed data restrictions, the sample size is smallest in the interest rate regressions and covers only the final years of reform activity—a period in which institutional quality had arguably already started to converge. Finally, the control variables are largely significant and carry plausible signs among other things, suggesting a beneficial impact of high levels of pre-Meiji human capital accumulation, closeness to the country's economic centers or high population density.

While the econometric analysis provides quantitative backbone to the historical sources, the dataset contains further intriguing insights if temporal dynamics are accounted for: Since the provinces that had supported the *shôgunate* experienced the most rapid and comprehensive personnel restructuring after 1868, regions that had supported the restoration managed to maintain more independence and lower levels of occupational permeability. Yet, in the course of time, this relationship probably changed. As appointment to public posts ceased to depend on status or personal ties and became a matter of achievement in standardized examinations, having many native officeholders no longer needed to be an unambiguous indicator of institutional backwardness. Rather it might point to a prefecture's ability to provide its citizens with the educational infrastructure they needed to excel at the national level. In this case, the effect of Native and RNH ought to weaken if not to change signs. To investigate this possibility, I generated the additional variables Native_time and RNH_time by interacting Native and RNH with a linear time trend. Both variables capture the temporal change in the impact of the two independent variables. Concretely, in the regressions on Startups and RelativeTax, negative signs on the interaction terms imply that the positive effect of strategic recruitment policies declined over time, whereas positive interaction coefficients suggest increasing benefits from personnel restructuring. Again, the inverse is true for the regressions on Lendingrate. The results of the extended estimation are displayed in Table 2.

The estimates confirm that the unfavorable impact of regional dominance in the prefectural administration became less palpable during the reform years. This is most evident in the regression on RelativeTax and with regard to the pair Native and Native_Time with both variables being significant and carrying opposite signs. The same pattern emerges in the cases of Startup and Lendingrate, although more coefficients fail to reach significance. These observations are not only worth noting because they confirm the qualitative historical evidence almost by the letter, but also carry general policy implications, namely first that strategic replacements constitute a powerful instrument to overcome institutional inertia and to initiate necessary reforms. Second, given that RNH and RNH_time turn out insignificant in many cases, it seems that beneficial growth effects were largely due to increased administrative permeability per se, whereas the question whether personnel restructuring occurred at the top or the lower strata of the bureaucracy was less important. Finally, no matter its instantaneous impact, this strategy seems to lose its strength over time because it becomes superfluous when meritocratic principles are

Table 2 Estimation with time effects

	Dep. variable: RelativeTax				Dep. variable: Startup				Dep. variable: Lendingrate			
	OLS	Regional controls	Lagged values	Arellano Bond	OLS	Regional controls	Lagged values	Arellano Bond	OLS	Regional controls	Lagged values	Arellano Bond
Native	−0.4940***	−0.4830***	−0.0969	0.0244	−0.2630	−0.0376	−0.3360	−0.1940	6.4100***	6.5870***	0.1830	−0.0877
	(0.1490)	(0.1540)	(0.0666)	(0.1450)	(0.3130)	(0.3330)	(0.3620)	(0.1310)	(1.9180)	(1.9530)	(1.8830)	(3.2630)
RatioNative	0.0755	0.0698	−0.0220	−0.0730	0.1760	0.1740	−0.1970	0.2370	1.3720**	1.2930**	0.8430	−0.0683
	(0.0572)	(0.0560)	(0.0329)	(0.0512)	(0.1190)	(0.1190)	(0.1630)	(0.1820)	(0.6550)	(0.6460)	(0.6850)	−1.400
Native_time	0.0202*	0.0231**	0.0003	0.0007	0.0005	−0.0078	0.0057	−0.0656	−0.4070***	−0.3890***	−0.0394	0.0192
	(0.0109)	(0.0108)	(0.0040)	(0.0093)	(0.0226)	(0.0231)	(0.0292)	(0.0770)	(0.1140)	(0.1150)	(0.0999)	(0.1600)
RatioNative_time	−0.0046	−0.0023	0.0017	0.0054	−0.0114	−0.0110	−0.0130	−0.0222	−0.0757**	−0.0679*	−0.0484	0.0014
	(0.0039)	(0.0039)	(0.0017)	(0.0039)	(0.0081)	(0.0082)	(0.0118)	(0.0143)	(0.0382)	(0.0378)	(0.0360)	(0.0713)
Relativetax_Lagged			0.7910***	0.0530			−0.0050	−0.0050			0.7090***	0.6460***
			(0.0572)	(0.1240)			(0.0702)	(0.0702)			(0.0345)	(0.0621)
AgriShare	−0.6970***	−0.5530***	−0.1620	0.1430	−0.9720***	−0.9690***	−0.9960**	2.0070**	2.4610***	2.4880***	0.1190	3.7660
	(0.1110)	(0.1140)	(0.1260)	(0.3170)	(0.2240)	(0.2380)	(0.4070)	(0.8500)	(0.7960)	(0.8340)	(0.5650)	−2.8440
Popden	0.0003**	0.0002	−3.66e−05	−0.0014***	0.0020***	0.0019***	0.0020***	−0.0018	0.0002	−0.0002	−0.0004	0.0004
	(0.0001)	(0.0001)	(9.59e−05)	(0.0004)	(0.0002)	(0.0003)	(0.0005)	(0.0038)	(0.0009)	(0.0009)	(0.0005)	(0.0032)
Kokudaka	−0.0009***	−0.00121***	−0.0003***		5.62e−05	−0.0003	7.80e−05		0.00613***	0.0052***	0.0014***	
	(0.0003)	(0.0003)	(8.97e−05)		(0.0005)	(0.0005)	(0.0002)		(0.0014)	(0.0016)	(0.0005)	
Distance	−0.0008***	−0.0008***	−0.0002*		−0.0003	−0.0005**	−0.0002*		0.00322***	0.0024***	0.0011***	
	(8.37e−05)	(0.0001)	(8.31e−05)		(0.0002)	(0.0002)	(0.0001)		(0.0006)	(0.0009)	(0.0004)	
Gakko	0.0682***	0.0531***	0.0074		0.0902**	0.1020**	0.0913***		0.1430	0.0923	−0.0057	
	(0.0192)	(0.0203)	(0.0105)		(0.0386)	(0.0447)	(0.0295)		(0.1290)	(0.1440)	(0.0886)	
Landlocked	−0.0283	−0.0502	−0.0148		0.0377	0.0088	0.0387		−0.1660	−0.6350**	−0.0380	
	(0.0354)	(0.0373)	(0.0239)		(0.0713)	(0.0862)	(0.0644)		(0.2340)	(0.2880)	(0.1750)	
Trend	−0.0100	−0.0118*	−0.00115		0.0052	0.0093	0.0024	0.1050*	−0.1240*	−0.1380*	−0.0758	−0.1660*
	(0.0065)	(0.0064)	(0.00271)		(0.0134)	(0.0136)	(0.0185)	(0.0637)	(0.0707)	(0.0709)	(0.0664)	(0.0995)
Conflict	−0.0072	−0.0074	−0.00186		0.1920***	0.1920***	0.1910***		1.4640***	1.4710***	1.1110***	
	(0.0320)	(0.0312)	(0.0125)		(0.0662)	(0.0662)	(0.0703)		(0.2590)	(0.2550)	(0.1400)	

Table 2 continued

	Dep. variable: RelativeTax				Dep. variable: Startup				Dep. variable: Lendingrate			
	OLS	Regional controls	Lagged values	Arellano Bond	OLS	Regional controls	Lagged values	Arellano Bond	OLS	Regional controls	Lagged values	Arellano Bond
Constant	−0.6970***	1.743***	0.437***		0.6270**	0.5740**	0.6760		11.230***	11.860***	4.3770***	
	(0.1110)	(0.137)	(0.134)		(0.2720)	(0.2900)	(0.4180)		(1.3960)	(1.4460)	(1.4300)	
Observations	723	723	646	557	801	801	767	687	553	553	516	445
R-squared	0.350	0.386	0.728		0.284	0.291	0.752		0.477	0.500	0.755	

* Indicates significance at 10 %, ** indicates significance at 5 %, *** indicates significance at 1 % level

successfully introduced or powerless when traditional elite networks are replaced by a new clique with other identifying links.

4 Robustness and extensions

As a matter of fact, policy lessons as those outlined above cannot draw only unless the Japanese case stands up to a close assessment of the estimation results' validity. Readers might be especially concerned with two points. For one, at least during the early Meiji years, rapid economic development and industrialization revolved mainly around the country's traditional metropolitan areas Tokyo, Kyoto and Osaka. At the same time, the three cities had been political, cultural and administrative centers from s*hôgunate* times onwards—a role they continued to play after the restoration. As a result, the administrative apparatus of the urban prefectures was regularly charged with over-regional tasks, implying an increased demand for specialist staff. Moreover, given the cities' strategic importance, the central government had strong reasons to staff administrative positions only with such individuals whose loyalty was beyond doubt. This implies that Tokyo, Kyoto and Osaka were not only the nuclei of the country's industrialization, but also the prefectures where permeability for non-native officeholders was arguably highest. Consequently, the prior results concerning the favorable impact of strategic replacement policies might be largely driven by the three cities' influence and not pertain to other prefectures. To address this concern, I re-estimated the model and excluded the metropolitan prefectures. The corresponding results are provided in Table 3. They show that for none of the three development indicators, the restriction of the sample leads to a change in the direction of the predicted relationships as far as the estimates remain significant. That this observation does not apply to the variable Startup is probably due to the fact that business activity clustered in the urban prefectures until well into the observation period with other regions becoming economic centers only at later points in time. As a result, variation in Startup becomes substantial only at a time when administrative structures had already been the subject of modernizing reforms for several years—i.e., when the influence of Native had started declining. As they arguably stem from ex-*ante* structural conditions, it seems therefore permissible to consider the insignificant results in the regression on Startup a minor drawback which does not fundamentally change the impression given in the previous sections.

The second potential problem relates to the possible endogeneity of population growth and administrative appointments. Concerning the former, readers might suspect that successful economic performance boosted population growth in the early reforming prefectures and thereby increased the ratio of native to non-native officials. Yet, given the statistical evidence, this was not the case as the correlation between population growth and the variable "Native" is small and negative (the correlation coefficient being merely −7.99 %). Counterintuitive at first, this result is less surprising once the timing of events is considered: Even if fast administrative rationalization fostered growth and even if economic success translated directly into higher fertility, this nexus would not be reflected in the ratio of native to non-native

Table 3 Estimation after sample reduction

	Dep. variable: RelativeTax				Dep. variable: Startup			
	OLS	Regional controls	Lagged values	Arellano Bond	OLS	Regional controls	Lagged values	Arellano Bond
Native	−0.2320***	−0.1620*	−0.0877**	−0.0066	−0.0064	−0.0006	0.0146	0.0175
	(0.0754)	(0.0800)	(0.0412)	(0.0627)	(0.1160)	(0.1250)	(0.1420)	(0.4740)
RatioNative	0.0324	0.0463*	−7.11e−06	−0.0292	−0.0305	−0.0168	−0.0479	−0.0980
	(0.0276)	(0.0268)	(0.0175)	(0.0196)	(0.0417)	(0.0415)	(0.0481)	(0.0951)
Relativetax_Lagged			0.8620***	−0.2980***			0.2020**	−0.0577
			(0.0516)	(0.0566)			(0.0917)	(0. 0726)
AgriShare	−0.5260***	−0.2580***	−0.1700	0.5710	0.1740	0.6270**	0.1330	2.6520***
	(0.1390)	(0.1640)	(0.1010)	(0.4190)	(0.2120)	(0.2580)	(0.2540)	(0.8910)
Popden	0.0006***	0.0006***	0.0001	−0.0012***	0.0006**	0.0002	0.0005*	0.0009
	(0.0002)	(0.0002)	(0.0001)	(0.0005)	(0.0003)	(0.0003)	(0.0003)	(0.0022)
Kokudaka	−0.0009***	−0.0011***	−0.0001***		−0.0004	−0.0005	−0.0003***	
	(0.0002)	(0.0002)	(7.93e−05)		(0.0003)	(0.0004)	(0.0001)	
Distance	−0.0007***	−0.0007***	7.895 − 05		−0.0004***	−0.0003*	−0.0003*	
	(7.76e−05)	(0.0001)	(7.77e−05)		(0.0001)	(0.0001)	(0.0001)	
Gakko	0.0741***	0.0724***	0.0049		0.0503*	0.0866***	0.0414*	
	(0.0178)	(0.0196)	(0.0098)		(0.0270)	(0.0302)	(0.0249)	
Landlocked	−0.0186	−0.0763*	−0.0005		0.0737	−0.1690***	0.0577	
	(0.0334)	(0.0401)	(0.0211)		(0.0506)	(0.0620)	(0.0464)	
Trend	0.0027	0.0039**	−0.0010	0.0195	0.0036	0.0060**	−0.0025	0.0195**
	(0.0019)	(0.0019)	(0.0008)	(0.0101)	(0.0029)	(0.0030)	(0.0028)	(0.0120)
Conflict	−0.0317	−0.0341	−0.0130		0.1810***	0.1770***	0.1570***	
	(0.0300)	(0.0290)	(0.0098)		(0.0466)	(0.0460)	(0.0547)	
Constant	−1.4920***	1.2440***	0.3300***		−0.0980	−0.4240*	−0.0570	
	(0.1280)	(0.1510)	(0.0978)		(0.1950)	(0.2380)	(0.2720)	
Observations	667	667	596	494	740	740	709	634
R-squared	0.220	0.281	0.726	–	0.057	0.091	0.085	–

Table 3 continued

Dep. variable and coefficients	OLS	Regional controls	Lagged values	Arellano Bond
Lendingrate				
Native	0.0094	0.3730	−0.3640	0.2370
	(0.6310)	(0.6610)	(0.4080)	(0.5550)
RatioNative	0.2520	0.3580	0.0451	0.0273
	(0.2210)	(0.2210)	(0.1320)	(0.2630)
Lendingrate_Lagged			0.7200***	0.3700***
			(0.0364)	(0.0477)
AgriShare	3.4350***	2.3660	−0.0591	0.6760
	(1.410)	(1.4220)	(0.7800)	(2.7670)
Popden	−0.0008	−0.0010	−0.0006	−0.0189
	(0.0014)	(0.0015)	(0.0009)	(0.0090)
Kokudaka	0.0061***	0.0052***	0.0014**	
	(0.0014)	(0.0017)	(0.0006)	
Distance	0.0031***	0.0024**	0.0010**	
	(0.0007)	(0.0010)	(0.0004)	
Gakko	0.1370	0.1210	−0.0230	
	(0.1360)	(0.1510)	(0.0915)	
Landlocked	−0.2640	−0.6600**	−0.0085	
	(0.2580)	(0.3230)	(0.1870)	
Trend	−0.3710***	−0.3760***	−0.1140***	−0.2430***
	(0.0225)	(0.0225)	(0.0205)	(0.0379)
Conflict	1.5320***	1.5260***	1.050***	
	(0.2740)	(0.2700)	(0.1450)	
Constant	14.490***	15.760***	4.9930***	
	(1.0770)	(1.3550)	(0.8930)	
Observations	515	515	481	415
R-squared	0.448	0.470	0.752	–

* Indicates significance at 10 %, ** indicates significance at 5 %, *** indicates significance at 1 % level

officials until the turn of the century when the first post-restoration cohorts reached the age to assume administrative positions. Since the observation period ends in 1895, endogenous fertility is therefore unlikely to seriously bias the previous estimates. To the extent this effect might become detectable in the last observation years, it would moreover distort the coefficients downwards. In other words: if high economic performance drove population growth and increased the number of suitable native applicants, "Native" would assume higher values in faster developing regions. Yet, this works in the opposite direction as the observed *inverse* relationship between "Native" and the development indicators. Similar considerations apply to the case where non-native administrators were deliberately appointed to "trouble provinces." In that case, high fractions of non-native staff coincide with weak economic performance and the endogeneity bias works against the positive effect of strategic replacement. The estimates obtained so far would then constitute a lower bound for the true impact of Native and RNH, and the direction of the relationship between the dependent and independent variables would remain unbiased. Things are more involved if booming regions were subject of within-country immigration or if they systematically attracted non-native staff. This scenario is conceivable if the central government used appointment to developed, wealthy regions as a means to reward outstanding personnel, if the administration of more advanced prefectures called for more specialist (non-native) staff or again, if prospective officeholders actively sought appointment in the highly developed prefectures. Given the historical record, the last two scenarios probably occurred because at the time, civil servants typically retired early and continued their careers as entrepreneurs or high-ranking managers in private companies.[32]

In order to avoid bias from these sources, I employed an instrumental variables approach and instrumented the variable Native. To follow this strategy, it was necessary to find an instrument which was correlated with Native and otherwise exogenous to the development indicators. The instrument used in the present estimation is the variable Governor_{jt}. Governor is a binary variable which takes the value of 1 if a new governor was appointed to prefecture j in year t and 0 otherwise. That the requirements for a valid instrument are met is apparent if the following historical details are taken into consideration: Governors were typically appointed according to fixed office terms or—in case of extraordinary circumstances like the Satsuma Rebellion—on special order of the central government. In any case, the staffing decision was taken in Tokyo with no indication of local economic characteristics having had any systematic influence on the timing or the choice of native and non-native governors. These appointment cycles warrant attention because new governors were regularly followed by a number of other new (predominantly non-native) appointees at the lower levels of the administrative hierarchy. Hence, years with a change at die prefecture's leadership offer a source of variation in the ratio of native to non-native officials that is independent of current economic performance. Repeating the analysis using Governor as an instrument leads to the estimation results depicted in Table 4.

[32] This phenomenon was widespread enough to enter contemporary Japanese under a special expression: "amakudari"—literally translated "descending from the heavens".

Table 4 Instrumental variables estimation

	Dep. variable: RelativeTax			Dep. variable: Startup			Dep. variable: Lendingrate	
	OLS	Lagged values		OLS	Lagged values		OLS	Lagged values
Native	−0.2580***	−0.0976*	Native	−0.2460*	−0.2520*	Native	−0.0362	0.4040
	(0.0821)	(0.0567)		(0.1270)	(0.1280)		(0.6000)	(0.4030)
RatioNative	0.0082	−0.0006	RatioNative	0.0240	0.0194	RatioNative	0.1680**	0.0062
	(0.0270)	(0.0182)		(0.0546)	(0.0569)		(0.1850)	(0.1260)
Relativetax_Lagged		0.7900***	Startup_Lagged		0.0003	Lendingrate_Lagged		0.7140***
		(0.0263)			(0.0390)			(0.0277)
AgriShare	−0.6910***	−0.1650***	AgriShare	−0.9690***	−0.9740***	AgriShare	2.3090***	0.0143
	(0.1110)	(0.0772)		(0.2380)	(0.2390)		(0.8030)	(0.5580)
Popden	0.0002*	−3.28e−05	Popden	0.0019***	0.0020***	Popden	0.0002	−0.0005
	(0.0001)	(8.23e−05)		(0.0003)	(0.0003)		(0.0009)	(0.0006)
Kokudaka	−0.0010***	−0.0003	Kokudaka	4.55e−05	5.58e−05	Kokudaka	0.0064***	0.0014
	(0.0003)	(0.0002)		(0.0005)	(0.0005)		(0.0014)	(0.0009)
Distance	−0.0008***	−0.0002***	Distance	−0.0003	−0.0003	Distance	0.0032***	0.0010**
	(8.33e−05)	(6.02e−05)		(0.0002)	(0.0002)		(0.0006)	(0.0004)
Gakko	0.0696***	0.0073	Gakko	0.939**	0.0945**	Gakko	0.1520	−0.0165
	(0.0192)	(0.0132)		(0.0386)	(0.0404)		(0.1300)	(0.0871)
Landlocked	−0.0313	−0.0143	Landlocked	0.0364	0.0367	Landlocked	−0.1320	−0.0144
	(0.0355)	(0.0242)		(0.0713)	(0.0745)		(0.2350)	(0.1580)
Trend	−0.0012	−0.0003	Trend	0.0008	−0.0001	Trend	−0.3760***	−0.1170***
	(0.0019)	(0.0014)		(0.0040)	(0.0044)		(0.0209)	(0.0181)
Conflict	−0.0023	−0.0027	Conflict	0.2020***	0.2020***	Conflict	1.5320***	1.1430***
	(0.0321)	(0.0211)		(0.00662)	(0.0675)		(0.2590)	(0.1720)
Constant	1.7710***	0.4290***	Constant	0.6820**	0.7060***	Constant	15.300***	5.0770***
	(0.1110)	(0.0930)		(0.2330)	(0.2500)		(0.8780)	(0.7310)
Observations	722	645	Observations	800	767	Observations	553	516
R-squared	0.345	0.728	*R*-squared	0.283	0.290	*R*-squared	0.462	0.755

* Indicates significance at 10 %, ** indicates significance at 5 %, *** indicates significance at 1 % level

As in the case of the sample reduction, none of the effects changes signs and most estimates which had been statistically significant in the non-IV specification remain so if Native is instrumented for. Overall, it thus seems save to conclude that the prior results stand up both to changes in panel composition and to cases where endogeneity is controlled for.[33]

Keeping this encouraging conclusion in mind, there are still two additional ways to corroborate the robustness of the proceeding sections and to assess the developmental significance of the observed effects. The first builds on the notion that if personnel restructuring was beneficial, prefectures where strategic recruitment policies were implemented earlier should have exploited their institutional head start to outperform late-coming regions. Following this argument, I introduced the variable mintime$_{jt}$. Mintime$_{jt}$ takes advantage of the fact that the temporal evolution of Native almost universally conformed to a degenerated U-shaped pattern. In particular, high ratios of native staff were followed by a decrease until regional dominance attained some minimum. Then, the process usually reversed with Native increasing again but typically not reaching its initial level. Mintime$_{jt}$ is therefore constructed as a count variable which renders the number of years until—or from—the point in time when Native reached its first local minimum. Thus, mintime takes the value of 0 in the minimum year and is increasingly negative for points in time *before* and increasingly *positive* for points in time after the minimum has been attained. Accordingly, if the previous results were to be confirmed, the estimated coefficients on mintime ought to take positive values in the regressions on Startup and RelativeTax and negative ones in the case of Lendingrate. This pattern would point to the fact that low realizations of the development indicators occurred before traditional networks had lost their influence and administrative rationalization had set in.

[33] Another piece of corroborative evidence can be derived from a direct measure of corruption. In its account of the late Tokugawa and early Meiji bureaucracies, Sect. 2 has outlined that administrative failure and questionable if not illegal activities of local officials were particularly frequent in the tax assessment of cultivated land. To reduce the tax losses from the common practice of undermeasuring the true size of cultivated fields, the imperial government conducted two large-scale cadastral surveys in 1872 and 1899. Unfortunately, the accuracy of these surveys is unknown. It is, however, known that they were mainly undertaken by local officials whose work was subjected to on-the-spot controls from centrally appointed surveyors. Hence, if undermeasurement survived into the twentieth century, it most likely did so in prefectures where traditional local influences were strong and administrative rationalization retarded. The first reliable quantitative assessment of undermeasurement was undertaken in the course of a comprehensive countrywide sampling survey in the 1950s (survey data are reported in Nakamura (1966) statistical annex). The corresponding reports include an Undermeasurement Index (UI) which denotes the ratio of the true size of a cultivated area relative to its reported size (i.e., if there is underreporting, the UI is larger than 1). Under the admittedly strong assumption that cross-prefectural differences were perpetuated for a long time or vanished very slowly, a high prefectural UI in the twentieth century would suggest a lack of institutional quality in the nineteenth century. I investigated this possibility by, respectively, correlating the two reported UIs (one for paddy and one for upland fields) with the institutional indicator "Native." The resulting correlation coefficients were both positive and amounted to 12 % for paddy and 29 % for upland fields. With all necessary precaution regarding the long time span between the observed institutional data and the undermeasurement indices, the estimates obtained in section 3 are therefore also supported by direct indicators of corruption. (For corresponding data see report in Nakamura (1966) statistical annex).

I report the estimates of the modified regressions with mintime as explanatory variable in Table 5. Once again, the statistical analysis confirms the expectations and the estimates are largely significant with the postulated signs. Moreover, the predicted effects are fairly substantial. For depending on which specification of the model is chosen, prefectures which reached their minimum say 3 years before the national average would be expected to exhibit a 1–2 % higher per capita income (as captured by RelativeTax) and 3–9 % lower lendigrates.

The final implementable robustness check provides an extension to the previous estimations by introducing an additional dependent variable. As stated before, lack of annual prefectural GDP data necessitated the recourse to the indirect development indicators. Yet, for single years at the end and after the observation period, comprehensive GDP estimates at the prefectural level have been compiled by Fukao et al. (2009, annexes). Again, the special appeal of this dataset stems from the fact that it provides an idea of the persistence and the magnitude of the economic damage inflicted by delayed institutional reform. Both of these aims are pursued in Table 6. The depicted results give the estimated coefficients of 4 regressions in which the average values of Native and RNH during the periods 1875–1879, 1880–1884, 1885–1889 and 1890–1894 were, respectively, regressed on the GDP estimates for 1890 and 1909, the two years for which corresponding data were available.

In terms of robustness, the estimates reconfirm what has been found before as, to the extent they are significant, the coefficients on Native and RNH all carry the expected signs. Moreover, the temporal evolution of the estimated coefficients suggests the following: First, one finds that the repercussions of early reform activity—i.e., the realizations of Native in the first periods—are detectable even in the much later GDP estimates of 1909. The impact of rapid bureaucratic modernization at the beginning of the Meiji era therefore seems to have retained its economic significance until well into the twentieth century. This remarkable persistence goes along with the observation that the explanatory power of variation in Native and RNH declines over time. Hence, as the process of administrative rationalization went on, as incentives changed and meritocratic principles became dominant, the importance of strategic recruitment policies as a short-term development strategy apparently declined. Yet, that early attempts at administrative modernization are nevertheless likely to translate into considerable developmental benefits becomes clear if the magnitude of the effect of reform activity on later GDP outcomes is accounted for. In particular, the estimated coefficients of Table 6 suggest that increasing the share of native officeholders in the first period by 1 % (i.e., increasing Native by 0.01) would have been reflected by a decrease of 0.69 Yen in the real prefectural per capita GDP of 1890 as measured in constant prices of 1934. The same pattern applies to 1909 with the corresponding estimate amounting to 0.37 Yen. Given that the national average per capita GDP for the two years was 133 Yen in 1890 and 168 in 1909, this implies that late-reforming prefectures like Kagoshima or Ishikawa with their shares of native officials, respectively, exceeding the national average by roughly 40 and 18 % paid dearly for the maintained dominance of traditional regional elite networks: In plain numbers, the econometric estimation predicts that their failure to participate in early strategic recruitment

Table 5 Timing of reforms and economic performance

	Dep. variable: RelativeTax				Dep. variable: Startup			
	OLS	Regional controls	Lagged values	Arellano Bond	OLS	Regional controls	Lagged values	Arellano Bond
Mintime	0.0040**	−0.0045**	−0.0120**	0.0088	−0.0050	−0.0010	−0.0067	−0.0395
	(0.0017)	(0.0020)	(0.0073)	(0.0074)	(0.0034)	(0.0040)	(0.0243)	(0.0381)
Relativetax_Lagged			0.0127	−0.1110			0.0116	−0.1800
			(0.0090)	(0.0902)			(0.0258)	(0.1120)
AgriShare	−0.6150***	−0.4840***	−0.4810*	0.0959	−0.8750***	−0.9200***	−0.8610*	1.6570**
	(0.1060)	(0.1110)	(0.2630)	(0.4110)	(0.2070)	(0.2240)	(0.4720)	(0.7120)
Popden	0.0004***	0.0002	0.0014	−0.0011***	0.0020***	0.0018***	0.0017***	−0.0024
	(0.0001)	(0.0001)	(0.0002)	(0.0004)	(0.0002)	(0.0002)	(0.0006)	(0.0032)
Kokudaka	−0.0009***	−0.0011***	−0.0015***		0.0001	−0.0003	−0.0002	
	(0.0003)	(0.0003)	(0.0003)		(0.0004)	(0.0005)	(0.0002)	
Distance	−0.0008***	−0.0008***	−0.0014***		−0.0003**	−0.0006**	−0.0004*	
	(7.60e−05)	(0.0001)	(0.0003)		(0.0001)	(0.0002)	(0.0002)	
Gakko	0.0674***	0.0745***	0.0783		0.0856**	0.1050**	0.0864**	
	(0.0181)	(0.0206)	(0.0482)		(0.0352)	(0.0411)	(0.0339)	
Landlocked	−0.0250	−0.1150**	−0.2280**		0.0285	0.0296	−0.0056	
	(0.0339)	(0.0401)	(0.0984)		(0.0657)	(0.0799)	(0.0787)	
Trend	−0.0062**	−0.0056**	0.0028	0.0688	−0.0056	−0.0012	−0.0042	0.0034
	(0.0026)	(0.0027)	(0.0058)	(0.0730)	(0.0049)	(0.0054)	(0.0067)	(0.0358)
Conflict	−7.01e−05	−0.0036	−0.0185		0.2050***	0.2030***	0.2020***	
	(0.0312)	(0.0304)	(0.0190)		(0.0624)	(0.0623)	(0.0685)	
Constant	1.5950***	1.5640***	1.8550***		0.5530***	0.6640***	0.6060	
	(0.0967)	(0.1070)	(0.2760)		(0.1890)	(0.2150)	(0.4280)	
Observations	777	777	775	601	881	881	875	797
R-squared	0.338	0.379	0.702	–	0.270	0.277	0.269	–

Table 5 continued

Dep. variable and coefficients	OLS	Regional controls	Lagged values	Arellano Bond
Lendingrate				
Mintime	−0.0117	0.0335	−0.0112	−0.1360**
	(0.0129)	(0.0153)	(0. 0088)	(0.0596)
Lendingrate_Lagged			0.6970***	0.7140***
			(0.0356)	(0.0430)
AgriShare	2.570***	2.1710***	0.0774	5.0070*
	(0.7900)	(0.8290)	(0.5500)	(2.6810)
Popden	0.0003	−0.0003	−0.0004	−0.0024
	(0.0008)	(0.0009)	(0.0005)	(0.0029)
Kokudaka	0.0065***	0.0057***	0.0015***	
	(0.0013)	(0.0015)	(0.0005)	
Distance	0.0031***	0.0025***	0.0007*	
	(0.0006)	(0.0009)	(0.0004)	
Gakko	0.1920	0.1480	−0.0019	
	(0.1280)	(0.1430)	(0.0862)	
Landlocked	0.0272	−0.5500*	0.0054	
	(0.2370)	(0.2940)	(0.1710)	
Trend	−0.3400***	−0.3190***	−0.1020***	0.0445
	(0.0240)	(0.0248)	(0.0221)	(0.0608)
Conflict	1.4140***	1.4260***	1.0780***	
	(0.2630)	(0.2570)	(0.1360)	
Constant	14.810***	15.590***	4.9090***	
	(0.7790)	(0.8450)	(0.7880)	
Observations	515	515	481	415
R-squared	0.448	0.470	0.752	–

* Indicates significance at 10 %, ** indicates significance at 5 %, *** indicates significance at 1 % level

Table 6 Persistence of reform effects

	Period 1 (1875–1879)	Period 2 (1880–1884)	Period 3 (1885–1889)	Period 4 (1890–1895)
Dep. variable: GDP 1890				
Native	−69.280**	−118.80***	−75.970**	–
	(32.690)	(38.990)	(34.070)	
RatioNative	31.540	−12.990	24.520	–
	(20.240)	(16.410)	(14.730)	
AgriShare	−92.360*	−105.40**	−118.50**	–
	(52.920)	(46.560)	(49.830)	
Popden	0.1260**	0.1000**	0.0815*	–
	(0.0500)	(0.0400)	(0.0400)	
Kokudaka	6.5600	−0.7500	0.3900	–
	(12.580)	(8.9300)	(0.2600)	
Distance	−0.0400	−0.0400	−0.0200	–
	(0.0300)	(0.0300)	(0.0300)	
Gakko	12.170	15.530**	12.760*	–
	(7.8800)	(7.2700)	(7.1000)	
Landlocked	3.7900	10.870	4.8300	–
	(13.650)	(13.260)	(13.440)	
Constant	182.60***	246.90***	214.80	–
	(53.860)	(46.650)	(43.170)	
Observations	40	43	43	
R-squared	0.683	0.701	0.708	
Dep. variable: GDP 1909				
Native	−36.830**	−52.800**	−15.980	−14.110
	(15.170)	(19.410)	(17.190)	(17.250)
RatioNative	16.000*	−2.8150	11.320	6.7950
	(9.4000)	(8.1700)	(7.4300)	(7.7600)
AgriShare	−48.800*	−48.120**	−59.140**	−52.020*
	(24.570)	(23.170)	(25.150)	(26.260)
Popden	0.0831**	0.0777***	0.0701***	0.0723***
	(0.0200)	(0.0200)	(0.0200)	(0.0200)
Kokudaka	−3.2620	−4.0670	0.1600	8.96e−05
	(5.8400)	(4.4500)	(0.1300)	(0.0300)
Distance	−0.0312*	−0.0370**	−0.0332**	−0.0402**
	(0.0200)	(0.0200)	(0.0200)	(0.0200)
Gakko	4.4430	7.0460*	4.4590*	4.7810
	(3.6750)	(3.6200)	(3.5820)	(3.6760)
Landlocked	3.6200	3.6200	2.7800	3.1400
	(6.3360)	(6.6000)	(6.7800)	(7.0460)
Constant	120.90***	136.60***	116.40***	113.30***
	(33.6191)	(40.510)	(19.440)	(10.540)
Observations	40	43	44	44
R-squared	0.789	0.761	0.752	0.737

* Indicates significance at 10 %, ** indicates significance at 5 %, *** indicates significance at 1 % level

policies caused prefectural per capita GDP to lack behind the national average by approximately 21 and 10 % (1890) and 9 and 4 % (1909).

Put together, the two facts provide clear-cut implications both about the workings of the institutional reforms that facilitated Japan's staggering development record during the Meiji era and about possible modernization policies in today's developing countries. In fact, what they suggest is first that early efforts at eradicating bureaucratic sleaze have a sizeable and persistent impact on future economic performance. And second, strategic recruitment and personnel replacement at all levels of the administrative hierarchy provide a powerful—and relatively inexpensive—strategy to reach this aim.

5 Conclusion

As Japan embarked on its rapid modernization course after the Meiji Restoration, it underwent a process of profound administrative rationalization. The essence of this evolution has arguably been grasped most cunningly by nineteenth-century novelist Soseki Natsume who somewhat exasperatedly stated:

> We eat for the *Nation*, wash our faces for the *Nation* and even go to toilet for the *Nation.* (Soseki 1914: 313)

Put less sarcastically, Soseki's statement hints how traditional region- and status-based institutions gave way to increasingly impersonal, professionalized and centrally designed forms of governance. These structures were deliberately implemented by the imperial government to support its *Fukoku Kyohei* development strategy. As a result, Meiji officials other than their *shôgunate* predecessors faced strongly merit-based incentives and a close internal supervision system. Yet, this profound organizational change occurred only to the extent that the new central government was able to impose its *de iure* monopoly of power *vis-à-vis* traditional local elites. In particular in the early post-restoration years, overcoming regional dominance came largely down to replacing native officials in the prefectural bureaucracies by centrally appointed staff from other parts of the country. Using the ratio of native to foreign officeholders and their hierarchical distribution as indicators of administrative rationalization, the present paper has aimed at assessing the long-run benefits from increased institutional permeability and the suppression of administrative sleaze. Its basic results are as follows: First, applying canonical econometric techniques, one finds that prefectures where strategic personnel restructuring was delayed performed significantly and consistently worse with regard to measures of regional per capita income, business activity and financial market development. What is more, the corresponding coefficient estimates not only prove fairly robust against modifications of sample composition and potential endogeneity, but are also of considerable size with, e.g., a 10 % increase in the ratio of native officials in the 1870s translating into an estimated 6 % penalty on 1890 prefectural per capita GDP. Finally—arguably due to historical path dependency—the impact of differences in the timing of administrative modernization exhibits a striking persistence over time. As long as well into the first decade of the twentieth

century, the composition of the prefectural bureaucracy—i.e., the ratio and the distribution of native versus non-native officials—in the 1870s and 1880s, constitutes a potent predictor of regional economic development.

From a historical perspective, these observations lend empirical backbone to the notion that administrative rationalization acted as a major catalyst for economic growth. Moreover, the present results provide an intriguing vantage point for research in current development studies. Once agreed upon the fact that personnel restructuring apparently provides a *prompt* and relatively *inexpensive* means to initiate institutional change, understanding under which socioeconomic conditions strategies similar to those embraced by the Meiji government facilitate public sector modernization in today's developing countries seems a worthwhile task. Lastly, since the Japanese evidence points to a diminishing significance of recruitment policies as compared with meritocratic, incentive- and surveillance-orientated *structural* modifications, identifying the optimal mix and the optimal timing of these interventions carries lessons which deserve as much attention from development scholars and present-day policy planners as from economic historians.

References

Printed sources and statistical compendia

Bank of Japan First Hundred Years' Editing Committee [ed.], Bank of Japan: The First Hundred Years: Materials, Tokyo 1986

Gabe, Masao [ed.], Meiji-15nen-Meiji-16nen-chiho-junsatsushi-fukumeisho (Collected Reports of Local Police Officers 1882–1883), reprint of original source, Tokyo 1980

MacLaren, Walter W. [ed.], Japanese government documents, Tokyo, 1914

Meiji-6-society (Meirokusha) [ed.], Meiroku Zasshi (Journal of the Meiji-6-society), Tokyo 1874–1876

National Statistical Office (Naikaku-Tôkei-Kyoku) [ed.], Nihon Teikoku Tôkei Nenkan (Statistical Annuals of the Japanese Empire), Tokyo various years

National Print Office at the Ministry of Finance [ed.], Kan-in Roku (later renamed Shoku-in Roku) (List of Government Officials), Tokyo various years

Statistical Offices of the Japanese Prefectures (Fu-Ken Tôkei-Kyaku) [ed.], Fu-Ken Tôkeisho (Statistical Annuals of the Prefectures), various years 1881–1902

Ministry of Education [ed.], Nihon Kyouiku-shi Shiryou (Documents on the History of Education in Japan), Tokyo 1890

Ôkawa, Kazushi, Shinohara Miyohei [ed.], Shôhi Shijutsu (Private Consumption), Tokyo 1967

Ôkawa, Kazushi [ed.], Chôki-Keizai-Tôkei—Bukka (Long Term Economic Statistics of Japan, Prices), Tokyo 1992

Ôkawa, Kazushi [ed.], Chôki-Keizai-Tôkei—Suikei to Bunseki (Long Term Economic Statistics of Japan, Regional Statistics), Tokyo 1992

Special Government Committee for Historical Research [ed.], Hansei Ichiran (A Summary of the Domain System), reprint Tokyo 1967

Literature

Acemoglu D, Johnson S, Robinson JA (2008) Persistence of power, elites, and institutions. Am Econ Rev 98(1):267–293

Aoki K (1967) Meiji Nômin Sôjô no Nenjiteki Kenkyû (chronological research on peasant uprisings in the Meiji period), Tokyo Shinsensha

Aoki K (1974) Hyakushô Ikki no Nenjiteki Kenkyû (chronological research on peasant uprisings), Tokyo Shinsensha
Baxter JC (1994) The Meiji unification through the lens of Ishikawa prefecture. Harvard University Asia Center, Cambridge
Besley TJ, Ghatak M (2009) Property rights and economic development. CEPR Discussion Papers 7243
Beasley WG (1972) The Meiji Restoration, 1st edn. Stanford University Press, Stanford
Befu H (1966) Duty, reward, sanction, and power. The four-cornered office of the Tokugawa Village Headman. In: Silberman BS, Haroothunian HD (eds) Modern Japanese leadership, 1st edn. University of Arizona, Tucson, pp 25–50
Befu H (1968) Village autonomy and articulation with the state. In: Hall JW, Hansen MB (eds) Studies in the institutional history of early modern Japan, 1st edn. Princeton University Press, Princeton
Buchheim C (2006) What causes late development? Insights from history. S Afr J Econ Hist 21(1):52–83
De Long JB, Shleifer A (1993) Princes and merchants: European city growth before the industrial revolution. J Law Econ 36(2):671–702
DeVere B (1966) Sidney, Okubo Toshimichi and the first home ministry bureaucracy 1873–1878. In: Harootunain HD, Silberman B (eds) Modern Japanese leadership. Transition and change, 1st ed. University of Arizona, Tucson, pp 195–232
Epstein L (2000) Freedom and growth, markets and states in Europe, 1300–1750, 1st edn. Routledge, London
Francks P (2002) Rural industry, growth linkages, and economic development in nineteenth-century Japan. J Asian Stud 61(1):33–55
Fukao K et al (2009) Zensôki Nihon no Ken-nai Sôseisan to Sangyô Kôzô, Hitotsubashi University Global COE Hi-Stat Discussion Paper Series 071
Harootunian HD (1970) Toward restoration, 1st edn. University of California Press, Berkeley
Honmi Y (1898) Satsuma Kenbun-Roku (Memoirs of Satsuma), Tokyo
Howell DL (1992) Proto-industrial origins of Japanese capitalism. J Asian Stud 51(2):269–286
Izumi S (2001) Bakufu no Chiki Shihai to Daikan, 1st edn. Koeibunkasha, Tokyo
Jansen MB, Roszman G (1986) Japan in transition, 1st edn. Princeton University Press, Princeton
Kasuya M (2005) Continuity and change in the employment and promotion of Japanese white-collar employees: the case of the house of Mitsui. Enterp Soc 6(2):224–253
Koh BC (1991) Japan's administrative elite, 1st edn. University of California Press, Berkeley
La Porta R, Lopez-de-Silanes F, Shleifer A, Vishny R (1999) The quality of government. J Law Econ Organ 15(1):222–279
Machlup F, Taber M (1960) Bilateral monopoly, successive monopoly, and vertical integration. Econ New Ser 27(106):101–119
McLaren WW (1916) A political history of Japan during the Meiji era 1867–1912, 1st edn. Frank Cass & Co., London
Matsushita S (2006) Kagoshima Han no Minshu to Seikatsu, 1st edn. Nanpoushinsha, Kagoshima
Mizu no Tomenaga (1980/1981) Yoshi no Sasshi, vol 8–9. Choukaironsha, reprint of historical documents, Tokyo
Mosk C (2011) Traps embraced or escaped: elites in the economic development of modern Japan and China, 1st edn. World Scientific, Singapore
Motoyama Y (1997) Proliferating talent: essays on politics, thought, and education, in the Meiji Era, 1st edn. University of Hawai Press, Honolulu
Murakami T (1997) Edo Bakufu no Daikan, 1st edn. Dohsei, Tokyo
Nakamura JI (1966) Agricultural production and the economic development of Japan, 1873–1922. Princeton University Press, Princeton
Nishizawa A (2004) Daikan no Ichijou Seikatsu, 1st edn. Kodansha, Tokyo
North D (2009) Violence and social orders: a conceptual framework for interpreting recorded human history, 1st edn. Cambridge University Press, Cambridge
Ohkawa K (1957) The growth rate of the Japanese economy since 1878, Tokyo
Ohshima M (1994) Meiji Kokka to Chiiki Shakai, 1st edn. Iwanami Shoten, Tokyo
Ono T (1932) Kinsei Chihou Keizai Shiryou, Tokyo
Ôtsu J (1881) Nihon Kanri Ninyôron, Tokyo
Paik C, Steele A, Tanaka S (2012) Rebellion and taxation: evidence from early modern Japan (working paper)
Ramseyer JM, Rosenbluth FM (1998) The politics of oligarchy: institutional choice in imperial Japan, 1st edn. Cambridge University Press, Cambridge

Ravina M (1998) Land and lordship in early modern Japan, 1st edn. Stanford University Press, Stanford
Shleifer A, Vishny RW (1993) Corruption. Q J Econ 108(3):599–617
Smith TC (1988) Native sources of Japanese industrialization, 1750–1920, 1st edn. University of California Press, Berkeley
Soseki N (1914) Kokoro, Maddison Books (1992, Reprint of 1914); New York
Spaulding RM (1967) Imperial Japan's higher civil service examinations, 1st edn. Princeton University Press, Princeton
Steenstrup C (1996) A history of law in Japan until 1868, 1st edn. Brill, Leiden
Steiner K (1965) Local government in Japan, 1st edn. Stanford University Press, Stanford
Tacke V (1975) Die Entwicklung der japanischen Landwirtschaft in der Tokugawa- und Meji-Periode. Dissertation University of Basel, Basel
Tang J (2011) Technological leadership and late development: evidence from Meiji Japan, 1868–1912. Econ Hist Rev 64(1):99–116
Westney ED (1987) Imitation and innovation, 1st edn. Cambridge University Press, Cambridge
Wilson GM (1992) George, patriots and redeemers in Japan: motives in the Meiji Restoration, 1st edn. University of Chicago Press, Chicago
Yamamura K (1974) A study of samurai income and entrepreneurship, 1st edn. Cambridge University Press, Cambridge

Future Lessons from Past Epidemics?

The Economics of Smallpox Vaccination in 19th Century Germany

Abstract:

In the last century, medical innovations like antibiotics and antiviral therapy promised to render the horrors of epidemics a matter of the past. Yet, recently new infections emerge in worrying frequency, while diseases which had long been under control make a comeback in many countries. Given the enormous increase in international trade and mobility, the identification of suitable strategies to prevent the spread of epidemics has therefore become an urgent challenge for social and economic politics.

The present paper argues that a country's epidemiological environment is a key factor in explaining the long-term evolution and efficiency of national health policies. Its vantage point is a simple question: How do biological, geographical and institutional conditions determine health strategies? And: Do decision-makers adjust institutional structures to changing exogenous conditions or does historical path-dependency produce persistent inefficiencies? To address these questions, the paper combines epidemiological modelling with a case study of smallpox prevention in 19th century Germany.

Its central results hold first, that international disease externalities are substantial and increasing in the *connectivity* between countries and their heterogeneity in epidemiological conditions. Historical evidence supports these predictions as policymakers in states which were highly exposed to imported epidemics tended to create institutions which facilitated rapid and coordinated response to outbreaks. In contrast, comprehensive prevention policies were not imposed unless initial conditions put decision-makers under the pressure to see for a high degree of institutionalization. The salient feature of this result is that institutional inertia apparently prevented states like e.g. Prussia and Wurttemberg from reforming their decentralized or uncoordinated policies.

With all due caution, history therefore suggests that the disease environment of a country has a lasting impact on its long-term health policies. Thus, there might be a lot to gain if the heighted interest in current health threats like Ebola was channeled into initiatives for multilateral health institutions, binding commitments to certain immunization and documentation targets and cooperation in the detection and containment of outbreaks.

Introduction:

When the armistice of January 28th 1871 ended the Franco-Prussian war, neither of the belligerents anticipated that the worst was yet to come: Lurking from the battlefields and the barricades of the Paris Commune, an aggressive strain of smallpox dogged the steps of soldiers and refugees, slipped from one country to the next and finally turned into a pandemic which left more than 500.000 Europeans dead. Smallpox have been eradicated for good. But other diseases quickly replaced it and seeing that Polio is ripping through Afghanistan, Pakistan and Syria, while Ebola keeps preying in conflict-torn Africa, and the MERS virus infests South Korea, there is no reason to put the epidemics of the past to the history books. This is all the more true as even the most careful evaluations of current policies have little to tell about the long-term effects of specific health policies. Hence, when it comes to conceiving efficient prevention programs, stepping back in history is probably worth the effort.

The pandemic of the 1870s is particularly intriguing because the disease appeared strangely picky as it continued its raid throughout Europe: Originating from France it killed more than 70.000 of Prussia's 20 million inhabitants but only 5000 of 5 million Bavarians. It raged in Austria, Belgium and Holland but failed to take foothold in Sweden. And while it claimed roughly 1000 lives per million in England it spared Scotland where the fraction approximated only 500 casualties per million.(Kotar, 2013, pp. 170-191). The mixed record largely stemmed from the fact that vaccination against smallpox was strongly encouraged or mandatory in some but not all countries of 1870 Europe. Stating the obvious does – however – miss the central question behind the epidemic: How could it be that many states failed to pass or to enforce comprehensive immunization laws although vaccination was inexpensive and its protective capacities widely known?

As it attempts to solve this puzzle, the present paper will go beyond the history of public healthcare in Europe and touch on matters which are crucial for the economical control of today's epidemics. Its vantage point is a simple question: How do biological, geographical and institutional conditions determine what kind of epidemiological policies are adopted in different countries? Seeing that microorganisms care little for political borders, it will subsequently ask: are the decisions of national health authorities inefficient at the international level? If not, what are the common characteristics of strategies which produce superior results in terms of long-term cost-efficiency and disease prevalence? And finally: Do decision-makers adjust institutional structures to changing exogenous conditions or does historical path-dependency produce temporal and inter-regional inefficiencies?

Given this broad set of questions, it is necessary to draw on various fields of research. The first – theoretical – cornerstone of the paper is the methodological toolbox of mathematical epidemiology and optimal control theory. While these methods capture the biological "mechanics" of an epidemic, economic theory is needed to model the optimization problem faced by health planners. The paper therefore also builds on canonical dynamic optimization models like Solow (1956), Ramsey (1928), Dixit (1990) and Romer (2001) as well as work on behavioral health economics and vaccination including Brito et al. (1991), Francis (1997), Gersovitz (2003) and Gersovitz and Hammer (2004). Based on this analytical framework it will be possible to create links to a second area of *theoretical* literature namely the optimal provision

of public goods and *empirical* work on actual health interventions like Baldwin (2001), Sköld (1996) and Rigau-Pérez (1989).

Even with these fundamentals established, some basic knowledge of the historical epidemiology of smallpox is needed to render the past epidemics a useful precedent for today. While this issue will be dealt with in the following section, the reminder of the paper is organized as follows: Section 3 introduces the structural framework of the model. Sections 4 to 6 discuss the corresponding historical evidence. Finally, section 7 concludes.

2. Historical Overview:

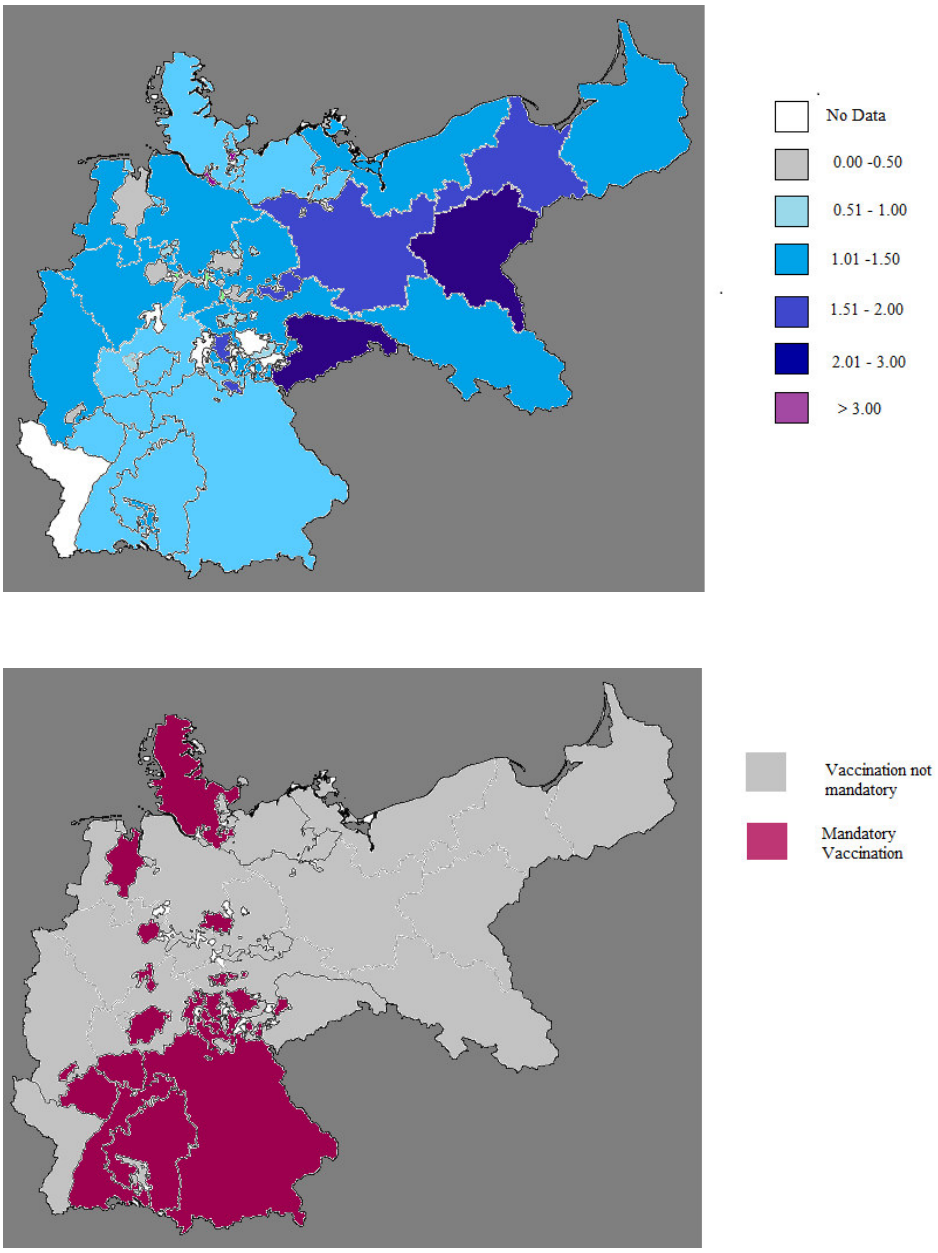
The year 1798 brought about a stir in the scientific community when English country doctor Edward Jenner published his findings about the possibility to induce immunity against smallpox by inoculating patients with cowpox, a disease similar to smallpox, but much less virulent. Knowledge about the new method spread rapidly from England to the continent and by the turn of the 19th century doctors all over Europe had started to vaccinate people with the pus from cowpox blisters. The tremendous impact of Jenner's modest 75-pages essay was no surprise. For at the time it appeared in print, smallpox was one if not *the* most aggressive and devastating disease. In fact, its death toll was so overwhelming that 18th century scholars claimed no-one would make a difference between the ravages of smallpox and the plague, were it not for the habit to count children's lives only once they had survived the former (Struve, 1802, p.1). Concerning smallpox prevalence, the picture looks even gloomier as contemporary sources from France and Germany estimated that approximately 80% of each cohort would catch the disease – a number which present-day research only slightly corrected to 66-75% (Wolff, 1998, p. 101).

As a result, getting a grip on smallpox promised such considerable economic and strategic benefits that most European governments encouraged the use of the lancet by means of information campaigns and free public vaccinations. Some of them like the German states of Baden and Wurttemberg even passed legislation which required the universal vaccination of all infants. Things changed markedly when the temporal limits of the vaccination-induced immunity became apparent: As more and more individuals who had been inoculated with cowpox as children caught the disease at a later point in time, even the most ardent vaccinators had to abandon the view of the vaccination's absolute protective capacity. Admitting that its strongest weapon was not invincible constituted a major blow for the newly forming science of medicine and created the breeding grounds for various schools of "alternative medicine". The academic dispute exerted noticeable influence on immunization policies because it ignited popular skepticism and cooled political enthusiasm for vaccination. What was more: after the first wave of massive vaccination, smallpox incidence had rapidly declined in many countries and since the risk of epidemics was low, policymakers of the revolutionary 1830s and 1840s were reluctant to provoke popular protest by newly introducing vaccination laws that bourgeois and liberal circles would largely reject as a token of despotic ancient regime policy.

Thus, the states which were to form the German Empire in 1871 might have been almost free of smallpox when the war broke out. Yet, their capacity to deal with a large-scale outbreak

varied drastically. The South-Western states as well as most of Thuringia, Oldenburg and Hannover required the vaccination of all infants and often encouraged the revaccination of school-age children. By contrast, Prussia followed a midway path by mandating only the vaccination of soldiers but not the civilian population, whereas the Hanseatic cities in the North passed no regulations at all. The institutional patchwork of different public health provisions was reflected in the impact of the 1870-74 pandemic. Figure 1 shows the geographical mortality rates of the epidemic. It is evident that the states which saw for the comprehensive immunization of their civilian populations fared far better than those which did not. This notion is further corroborated by the fact that the mortality difference between vaccinating states and laissez-faire states is statistically highly significant and stands at roughly 30%.

Figure 1: Smallpox Mortality during 1870-74 Pandemic and Vaccination Laws in Germany



Source: Imperial Health Department (ed.) Mitteilungen aus dem Reichsgesundheitsamt, various issues, 1881-1884

While different vaccination policies explain the bulk of regional variation in disease incidence, they cannot answer the question why the German states embarked on such different policy trajectories in the first place. In order to understand what initially determined the decisions of Germany's health planners, the next section will introduce a standard epidemiological model which will frame the choice problem of the historical policymakers in the subsequent analysis. This framework is necessary, because it will help us to determine how different conditions affected a country's expected epidemic costs and how these costs created incentives for policymakers to adapt the institutional structure in a particular way.

3. Modeling Epidemics: The SIR Model:

To analyze, how geographical, socio-economic and biological factors affected disease costs and optimal countermeasures, we replicate the decision problem of the German health authorities in the following stylized world: Epidemiological policies are determined by a single benevolent decision-maker who seeks to minimize the damage from infectious diseases. By assumption, the costs of an epidemic are proportional to the total number of infections, i.e. $C_{epidemic} = \vartheta I_{total}$.

The policymaker influences the spread of the disease by means of a single health input h which reduces the population's susceptibility and therefore the final rate of morbidity and mortality. In order to concentrate on the optimal provision of *prophylaxis*, the effect of h is assumed purely preventive and treatment – even if providing some relief to the sick – has no significant impact on the epidemic process.

We account for possible interaction at the international level by assuming that the spread of an epidemic in country i does not only depend on its own health investment $h_i$but also on the prevention effort of adjacent countries, h_j. The problem of the national decision-maker is then to minimize the sum of health expenditures and expected disease costs for given levels of foreign prevention. In formal terms this objective can be stated as follows:

$$\min_{h_i} \mathfrak{C} = (\pi_i \cdot C^i_{epidemic}(h_i, h_j) + c^i_h h_i)$$

The parameter π_i thereby denotes the natural prevalence rate of a disease in country i as a function of environmental characteristics like climate or the presence of animal hosts required by certain pathogens. Country i is epidemiologically more exposed, the higher π_i. The other terms capture the per-unit costs of prophylaxis c^i_h and the expected costs of an epidemic outbreak, $C^i_{epidemic}(h_i, h_j)$. To complete the formal set-up of objective function, we impose the realistic assumption that the marginal returns to prevention are decreasing in both countries with $\frac{\partial C^i_{epidemic}(h_i,h_j)}{\partial h_i} < 0$ and $\frac{\partial^2 C^i_{epidemic}(h_i,h_j)}{\partial {h_i}^2} > 0$.

Given the structure of the problem, the health authorities in country i face the tradeoff between the *certain* costs of prophylaxis $c^i_h h_i$ and the benefit from limiting the *expected* size of an outbreak $C^i_{epidemic}(h_i, h_j)$. The optimal choice of prevention will therefore equate the marginal benefit of reduced disease incidence with the marginal costs prophylaxis. In other words: in order to derive the policymaker's choice of optimal prevention effort h_i, we have to specify explicitly how the epidemic – and with it its expected costs – react to variation in environmental and institutional preconditions as well as to adjustments in the intervention variable h_i.

A classical approach to model the biological "mechanics" of infectious diseases is the *compartmental* SIR model by Kermack and McKendrick (1927). The term *compartmental* stems from the fact that the total population (N) is divided in the three compartments of those susceptible (S), infected (I), and recovered (R) with movements between the different states being represented by the below set of differential equations.

$$\dot{S} = -\beta\varphi(h_i, h_j)SI \quad \text{(i)}$$

$$\dot{I} = \beta\varphi(h_i, h_j)SI - (\gamma + \delta)I \quad \text{(ii)}$$

$$\dot{R} = (\gamma + \delta)I \quad \text{(iii)}$$

The interpretation of this system is straightforward:

(i) As hitherto healthy individuals contract the disease, the fraction of **susceptibles (S)** declines and that of infectives increases by the same amount. The corresponding rate of infection depends on the transmission rate β, which captures the biological virulence of the pathogen and the (institutionally determined) intensity at which susceptibles interact with potential carriers of the disease. One can therefore think of β as a measure of the epidemic's inherent ability to spread whereas its sensitivity to health interventions is depicted by the function $\varphi\left(h_i, h_{j_i}\right)$ which denotes the effect of prevention on susceptibility with $\frac{\partial\varphi\left(h_i, h_{j_i}\right)}{\partial h} \leq 0$ and $\frac{\partial\varphi^2\left(h_i, h_{j_i}\right)}{\partial^2 h} \geq 0$ for h_i and h_j.

(ii) Mirroring the dynamics of the susceptible class, the fraction of **infected** individuals **(I)** grows due to contagion and declines as infected individuals recover at rate γ or succumb to the disease at rate δ. Finally, individuals who have recovered from the disease are temporarily immune against reinfection. Just as those who have died, they cannot spread the pathogen any further and enter the compartment of those **"removed" (R)** from the epidemic process.[1]

Attempting to employ the above set of dynamic equations to derive the optimal prevention level, the health planner encounters an obstacle: What she needs to plug into the objective function are the values of S and I as functions of time and *not* the temporal derivatives determined by the SIR-system. Unfortunately, solving this task is not trivial because the corresponding transcendental equations cannot be solved analytically for S and I. To arrive at the desired result, it is useful to distinguish between different prototypical outcomes of the epidemic: To understand this approach, note first that S(t) and I(t) have to be nonnegative. Hence, whenever either of the two reaches zero, the epidemic process terminates. Since $\dot{S} < 0$ for all S and I, the spread of the disease will follow one of three possible patterns depending on the decision taken by the health planner:

(i) A sizeable outbreak in which the infected population share grows as long as $S > \frac{\gamma}{\varphi\left(h_i, h_{j_i}\right)\beta}$ (i.e. as long as there are enough susceptibles to ensure that the rate of new infections exceeds the recovery rate) and then decreases and ultimately approaches zero.

(ii) The occurrence of some sporadic cases if the initial size of S is small with $S(0) < \frac{\gamma}{\varphi\left(h_i, h_{j_i}\right)\beta}$ which leads to *herd immunity* and prevents the disease from affecting a sufficient number of victims to sustain an epidemic outbreak. And finally,

[1] To keep things mathematically tractable, the model abstracts from demographic effects by assuming that the population is stable in absence of diseases and that deaths occur only due to infection. While simplifying matters considerably, this assumption does not appear overly restrictive since the timeframe of the sort of epidemic infectious diseases considered in this paper is far shorter than that of demographic change.

(iii) A catastrophic epidemic which terminates only if there are no susceptibles left, i.e. as S converges to zero.

The last two scenarios correspond to "corner solutions" with respect to the optimal health intervention h: In the second case, prevention is so comprehensive that the disease fails to spread within the population and the number of infections remains negligible. With $I_{total} \approx 0$, total costs $\mathfrak{C}$ are equal to prevention costs $c_h^i h_i$. This outcome is only optimal if the costs of permitting *one single* infection $\vartheta \cdot 1/N$ are higher than the costs $c_h^i \overline{h_\iota}$ of setting prevention to the level $\overline{h_\iota}$ which ensures *herd immunity*, i.e. $\varphi\left(\overline{\left(h_\iota, h_{J_\iota}\right)}\right) < \frac{\gamma}{\beta S(0)}$.[2] The equilibrium is then given by $S_\infty = S(0)$, $I_\infty = 0$ and $h_i, h_{j_i} = \overline{\left(h_\iota, h_{J_\iota}\right)}$.

Another corner solution occurs if the costs of prevention are prohibitively high, that is, if purchasing one unit of h is more costly than the expected damage caused by an epidemic which affects the entire population. In this case, the optimal prevention investment is equal to zero and the spread of the epidemic will remain unaffected by policy interventions.

It is evident that none of these two scenarios applied to 19th century Germany – or to any current health threat for that matter. For, neither had smallpox been eradicated prior to the 1870-74 pandemic (case ii) nor did any state refrain completely from attempting to stop the assault of the epidemic (case iii). This leaves us with the epidemiologically relevant cases, in which a part of the population will be affected while a strictly positive fraction of individuals escapes infection.

Mathematically, a limited epidemic outbreak is characterized by the conditions $I_\infty = 0$ and $I_{total} = \mathrm{S}(0) - S_\infty$ with $S_\infty > 0$. Using these relationships and the first two dynamic equations of the SIR system, it is possible to derive I as a function of S – albeit not of t directly. To do so, it suffices to divide the two differential equations to obtain the relationship

$$\frac{dI}{dS} = \frac{(\beta\varphi\left(h_i, h_{j_i}\right)S - \gamma)I}{-\beta\varphi\left(h_i, h_{j_i}\right)S} = -1 + \frac{\gamma}{\beta\varphi\left(h_i, h_{j_i}\right)S}.$$

Integrating this expression with regard to S yields:

$$I = -S + \frac{\gamma}{\beta\varphi\left(h_i, h_{j_i}\right)} \cdot \ln S + K \quad \text{(iv)}$$

where K is a constant of integration which is determined by the initial values S(0) and I(0).[3] Moreover, since the epidemic must eventually abate, the number of infectives will decrease over time such that $\lim_{t\to\infty} I(t) = I_\infty = 0$. Plugging this into (iv) and solving for S, one finds:

[2] In this case, a single infection produces less than one secondary infection such that each removed individual is replaced by less than 1 new infection. Hence the infected compartment shrinks monotonously and the outbreak abates before growing to a full-blown epidemic.

[3] Note that If the environment allows for an epidemic outbreak (i.e. if $I(0) \geq 0$ and $\frac{\gamma}{\beta\varphi\left(h_i, h_{j_i}\right)} > 1$), K will necessarily be larger or equal to 1.

$$S_\infty = e^{-\mathcal{W}\left(\frac{\beta\varphi\left(h_i,h_{j_i}\right)e^{\frac{-K\beta\varphi\left(h_i,h_{j_i}\right)}{\gamma}}}{\gamma}\right)+\frac{K\beta\varphi\left(h_i,h_{j_i}\right)}{\gamma}} \quad \text{(v)}$$

where $\mathcal{W}(...)$ denotes the Lambert W function. Since S(0) is exogenous, $C_{epidemic}$ is given by:

$$C_{epidemic} = \vartheta\left(S(0) - e^{-\mathcal{W}\left(\frac{\beta\varphi\left(h_i,h_{j_i}\right)e^{\frac{-K\beta\varphi\left(h_i,h_{j_i}\right)}{\gamma}}}{\gamma}\right)+\frac{K\beta\varphi\left(h_i,h_{j_i}\right)}{\gamma}}\right) \quad \text{(vi)}$$

With this result, the policymaker can move on to determine the optimal prevention level by plugging the above expression into the objective function. As she solves the optimization problem, the decision-maker in a country i will rely on the following first order condition:

$$\frac{\partial \mathfrak{C}}{\partial h_i} = \pi_i \frac{\partial C^i_{epidemic}(h_i,h_j)}{\partial h_i} + c^i_h = 0 \text{ (vii)}$$

This relationship is central because it depends only on h_i , h_j and exogenous parameters. It will therefore serve our purpose of characterizing the behavior of h_i and $C_{epidemic}$ under varying environmental conditions. To do so, it suffices to consider the total derivative of (vii):

$$\mathrm{d}\mathfrak{C} = \pi_i \frac{\partial^2 C^i_{epidemic}\left(h_i,h_j\right)}{\partial {h_i}^2} dh_i + \pi_i\left(\frac{\partial^2 C^i_{epidemic}}{\partial h_i \partial\beta}\right)d\beta + \pi_i\left(\frac{\partial^2 C^i_{epidemic}}{\partial h_i \partial\gamma}\right)d\gamma + \pi_i \frac{\partial^2 C^i_{epidemic}\left(h_i,h_j\right)}{\partial h_i \partial h_j} dh_j = 0$$

The following subsection will scrutinize this expression. In doing so, our aim will be the same as the purpose of the hypothetical policymaker: Assessing which circumstances require comprehensive preventive strategies and which conditions allow for less intensive effort.

3.2. Identifying Optimal Prevention Strategies and Assessing the Costs of Epidemics:

In order to assess how prevention strategies will change when they are adapted to different diseases and different socio-economic environments, one has to determine the signs of the terms $\frac{d\,h_i}{d\beta}$, $\frac{d\,h_i}{d\gamma}$ and $\frac{d\,h_i}{dh_j}$. Performing this exercise first for the two health inputs we find:

$$\frac{d\,h_i}{dh_j} = -\frac{\frac{\partial^2 C^i_{epidemic}}{\partial h_i \partial h_j}}{\frac{\partial^2 C^i_{epidemic}\left(h_i,h_j\right)}{\partial {h_i}^2}} \quad \text{(viii)}$$

Since returns to prevention are decreasing, the denominator of the expression is positive and indicates, that health planners cannot expect large additional gains from reinforced prophylaxis if they are already pursuing a rigorous preventive strategy. More importantly, because prevention activities at home and abroad are imperfect substitutes in the "production" of immunity, increased prophylaxis in one country reduces prevalence abroad.

The term $\frac{\partial^2 C^i_{epidemic}}{\partial h_i \partial h_j}$ will therefore be positive such that the total expression (viii) is strictly negative and the optimal level of domestic prophylaxis falls. Keeping this result in mind is important because it implies that countries can free-ride on the prevention effort of their neighbors and thereby create negative disease externalities at the international level. Proceeding to the partial effects of changes in two epidemiological parameters γ and β yields:

$$\frac{d\,h_i}{d\gamma} = -\frac{\frac{\partial^2 C^i_{epidemic}}{\partial h_i \partial \gamma}}{\frac{\partial^2 C^i_{epidemic}(h_i,h_j)}{\partial {h_i}^2} + \frac{\partial^2 C^i_{epidemic}}{\partial h_i \partial h_j}\frac{d\,h_j}{dh_i}} \quad \text{(ix)}$$

$$\frac{d\,h_i}{d\,\beta} = -\frac{\frac{\partial^2 C^i_{epidemic}}{\partial h_i \partial\,\beta}}{\frac{\partial^2 C^i_{epidemic}(h_i,h_j)}{\partial {h_i}^2} + \frac{\partial^2 C^i_{epidemic}}{\partial h_i \partial h_j}\frac{d\,h_j}{dh_i}} \quad \text{(x)}$$

The denominators of the two expressions are larger than zero because at an interior optimum the second-order effect $\frac{\partial^2 C^i_{epidemic}}{\partial h_i \partial h_j} \cdot \frac{d\,h_j}{dh_i}$ is necessarily smaller in magnitude than the direct effect $\frac{\partial^2 C^i_{epidemic}(h_i,h_j)}{\partial {h_i}^2}$. Since $\varphi(h_i, h_j)$is homogenous and independent of γ and β, the fact that the principal branch of the Lambert W function is strictly concave, causes $\frac{\partial^2 C^i_{epidemic}}{\partial h_i \partial \beta}$ to be strictly negative and $\frac{\partial^2 C^i_{epidemic}}{\partial h_i \partial \gamma}$ to be strictly positive.

These mathematical relationships translate into unambiguous policy directions for actual health planners: First, since $\frac{d\,h_i}{d\beta}$ is strictly positive, the optimal reaction to higher epidemic spreading capacity β is an increase in the level of prevention. The necessary increase in $h_i$will thereby be larger if health planners in neighboring countries take advantage of reinforced prevention in country i by reducing their own efforts h_j.[4]

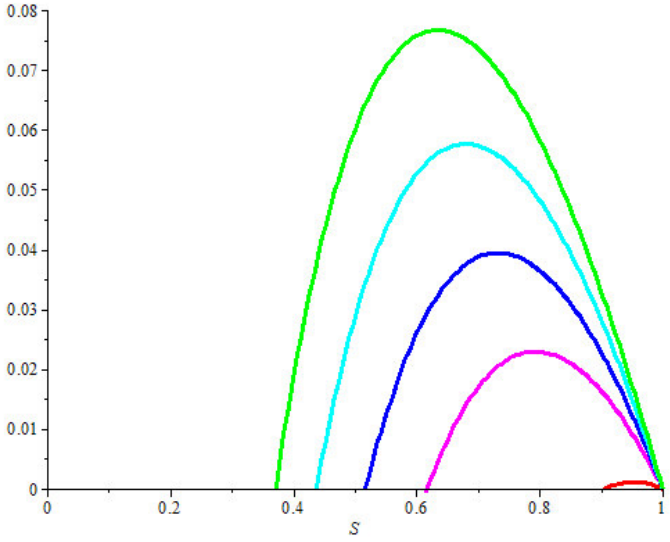
Conversely, the impact of epidemics is mitigated if the infective spell $\frac{1}{\gamma}$ is reduced because infectives will transmit the disease to fewer susceptibles the shorter the infective period. Hence, to be optimal, preventive activities should be reinforced in case of low recovery rates γ, whereas shorter infective periods and fast recovery necessitate less intensive prophylaxis. In sum, the

[4] The reason for this is that *ceteris paribus* the total level of infections - and hence $C_{epidemic}$ - rises as β grows larger. In order to ensure that the optimality condition is fulfilled, $\varphi'(h)$ therefore has to decline to offset the increase in $C_{epidemic}$. (Note that $\frac{\partial\, C_{epidemic}}{\partial \beta}$ is >0 because K>1).

damage of an uncontrolled epidemic outbreak – and hence the expected return prevention – is greater, the higher the so-called effective contact rate $R = \frac{\beta}{\gamma}$.

This fact is illustrated in Figure 2 which shows the trajectory of the epidemic in the S/I space for different values of R.

Figure 2: Disease trajectories for different effective contact rates R



The respective directions of the relationships between a population's socio-environmental conditions and the optimal epidemic strategy may be intuitive. Yet, the model can be used to investigate a broad range of additional questions. In the interest of conciseness we will concentrate on two central aspects of the epidemiological process: First to assess whether health planners did (and arguably still do) adapt their prevention strategies to the prevailing environment, it is important to understand under which conditions the costs of disease outbreaks were highest. For, the higher the expected costs of an epidemic, the stronger the incentives to take political action. This aspect is particularly evident with regard to time: If a country's characteristics punish failure to detect and contain outbreaks severely, the health authorities will be more inclined to establish a reliable epidemiological infrastructure. What is more: Besides for their long-term impact through historical path dependency, the expected costs of epidemics are also important because they impose a current burden which might be unbearable for countries with limited financial and expertise resources. The next section will address the policy implications of these issues by retracing the risks and incentives of different expected epidemic patterns and by discussing the costs of delayed response to disease outbreaks.

To begin with, the size of actual outbreaks is strictly decreasing in exposure and increasing in a country's exposure and prevention costs. This is the case because – as discussed previously – the health authorities in country i determine their epidemiological strategy by equating the marginal decline in the *expected* size of an outbreak with the *certain* marginal costs of prevention. If prevention becomes more burdensome or if the risk of a serious outbreak declines, policymakers will therefore be increasingly reluctant to enforce comprehensive

epidemic policies.[5] Thus, if effective countermeasures are hard to implement or again, if the prevalence of the disease declines, epidemic outbreaks will be rare events. But when they do occur, the risk of large-scale epidemics overwhelming a country's capacities is far higher than in the case of endemic or frequently recurring infections.

The public health infrastructure might – however – not only collapse if low prevalence rates lull the authorities into a false sense of security. The same applies to increases in the effective contact rate $= \frac{\beta}{\gamma}$. Relying on graphical analysis, this relationship is illustrated in figure 3 which depicts total health costs of country i as a function of domestic prophylaxis $h_i^*(h_j, R)$ and the effective contact rate R.

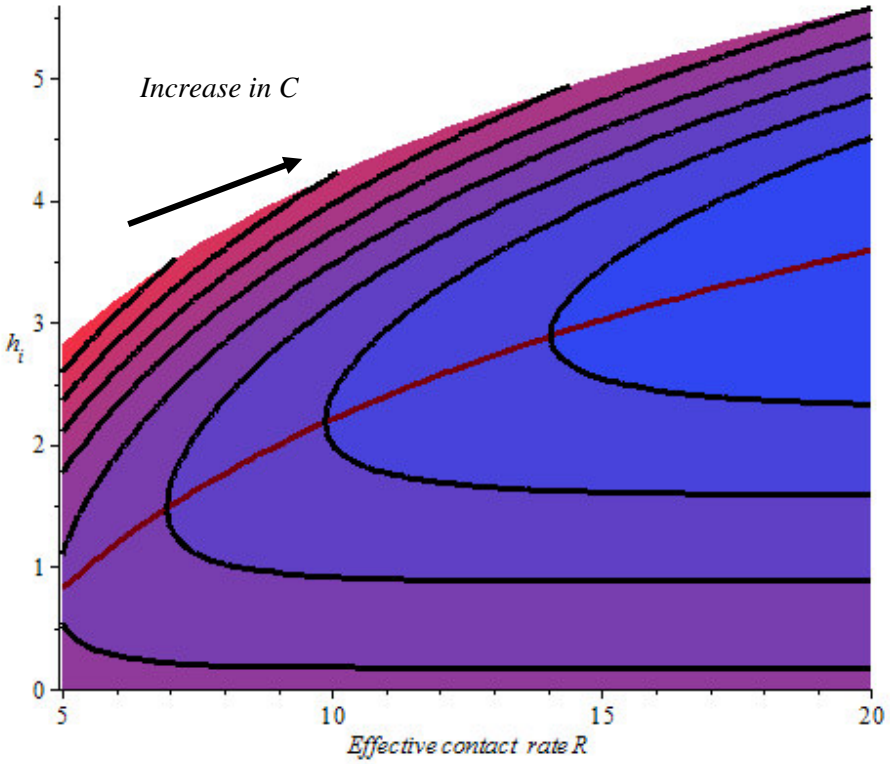


Figure 3: Epidemic costs for different values of R

Taking foreign epidemic policies as given, the red curve shows the optimal choice of domestic prevention $h_i^*(h_j, R)$if adjacent countries impose the fixed prevention level h_j. The total costs faced by country i, $\mathfrak{C}(h_i, h_j, R)$ are then determined by the intersection of the red curve and the black iso-costcurves. That infectious diseases cause more harm, the faster their spread follows from the fact that the red reaction curve $h_i^*(h_j, R)$ intersects higher cost-curves the larger R.

It is important to note in this context that the cost function is eventually concave because the total number of infections is limited by the size of the population. In other words: No matter

[5] The reason is that the derivative of the maximum number of infections, $I_{max} = S(0) - I(0) - \frac{\gamma}{\beta\varphi(h_i,h_{j_i})}\ln(S(0)) - \frac{\gamma}{\beta\varphi(h_i,h_{j_i})} + \frac{\gamma}{\beta\varphi(h_i,h_{j_i})}\ln(\frac{\gamma}{\beta\varphi(h_i,h_{j_i})})$ with respect to h_i is larger in absolute terms than the derivative of S_∞ with respect to h_i.

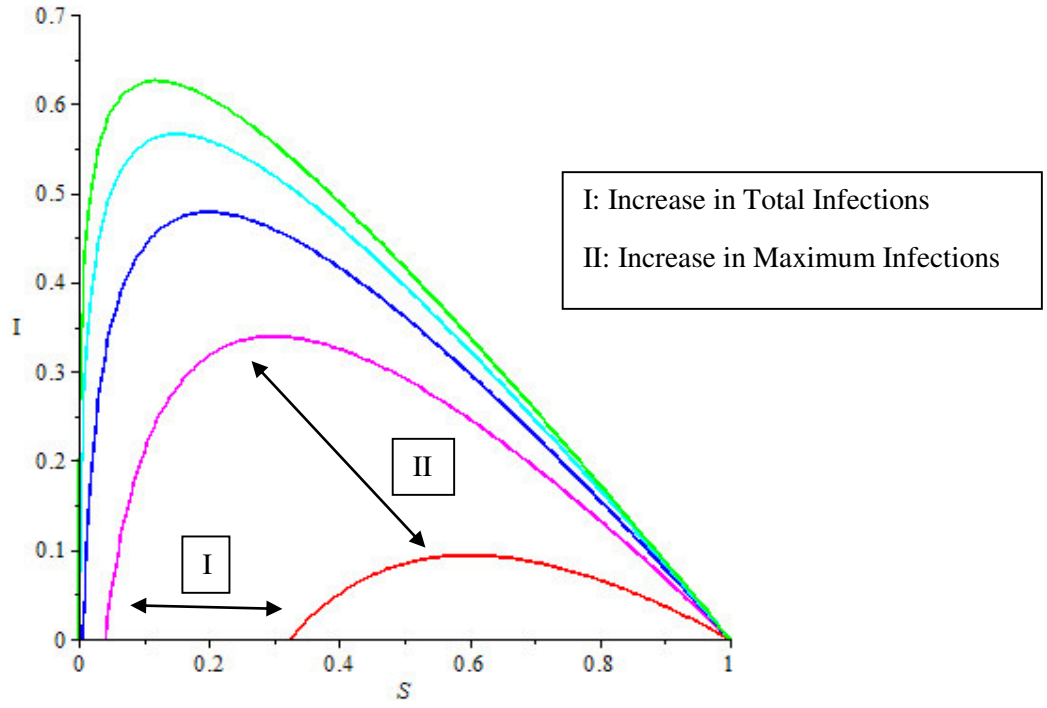
how fast the pathogen spreads and how devastating its impact, it cannot infect more than *all* individuals within its reach. Thus, once the epidemic runs out of potential victims, further increases in R have no effect on the final costs of an outbreak.

Unspectacular at first glance, this finding has profound implications for the interpretation of our historical evidence and for the coordination of current prevention strategies: It implies, that marginal changes in an epidemic's spreading capacity are more harmful when the contact rate R starts to increase from low initial levels. Thus, if changes in socio-economic or environmental conditions (e.g. mobility, climate) remove the natural barriers of pathogen transmission, the countries which experience the strongest marginal changes – and hence the most urgent need to adapt – are not the most isolated and backward ones. Nor will it be highly developed industrial nations. Rather, the disease's assault will be most harmful – i.e. the elasticity of the cost function with respect to R will be largest – for societies whose health institutions are in their infancy while mobility and contact rates begin to increase. This scenario arguably captured the situation of Europe during its industrialization and – more importantly – it continues to apply to today's emerging economies.[6] The important message from the epidemiological planning model is that - given the costs of a potential outbreak are concave - the marginal impact of these changes is *ceteris paribus* maximized when β reaches the level at which the effective contact rate R crosses the epidemic threshold R = 1. Then, the population's initial spreading capacity is still low and its variation hard to notice for the authorities. But if unforeseen events - like the Franco-Prussian War - cause small and transient increases in β, the corresponding change in infection numbers and hence the strain on the health sector will be considerable. What is more, as mobility and interregional contacts intensify, the maximum number of infections which occur at the height of the outbreak increases by more than the final size of the epidemic as figure 4 shows.[7]

[6]Adding to that, due to climate change and population growth, initially low epidemic spreading capacities β can grow imperceptibly in any society no matter whether developed or not. The reason is that most diseases with pandemic potential are *zoonotic*, that is, the causative pathogen spills over from animal to human hosts or requires an animal vector. Due to environmental change, human populations which used to be hard to infect for a certain pathogen might then end up facilitating its spread and produce large-scale epidemic outbreaks. Such increases in β take currently place in temperate regions like Europe where global warming paved the way for the insect hosts of formerly "tropical" diseases like Dengue or Yellow fever. On the other hand, population growth typically results in the intensification of agriculture and in the advancement of human settlements into hitherto untouched ecosystems. In both cases, zoonoses like coronaviruses (SARS) and H5N1 (avian flu) – which typically originate from poultry farms – or HIV and fibroviruses (Ebola and Marburg fever) – which used to be diseases of wild primates – benefit from strongly increased spreading possibilities. See: Lindgren et. Al. Monitoring EU Emerging Infectious Disease Risk Due to Climate Change, in: Science, vol.336, 2012, pp. 418-419

[7] As with changes in the conditional costs of an outbreak, this is the case because the derivative of the maximum number of infections I_{max} is strictly larger than the derivative of the epidemic's final size.

Figure 4: Epidemic trajectories for different spreading capacities



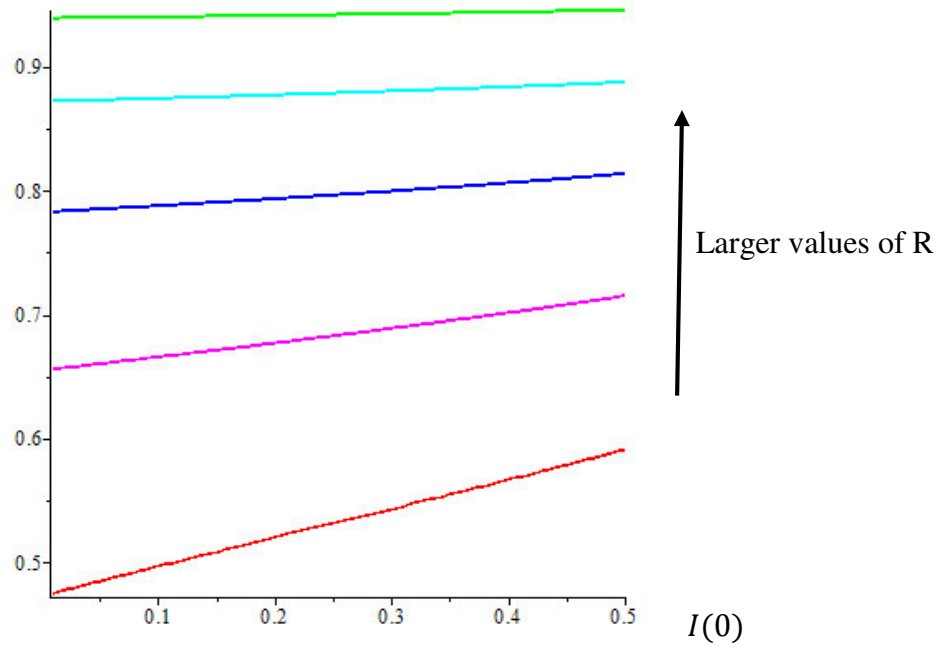
This epidemiological characteristic is a serious challenge for past and present health authorities alike because sudden peaks in prevalence threaten to overwhelm the capacities of the existing prevention infrastructure. For example, when Germany was hit by the 1870-74 pandemic, numerous regions ran out of vaccine and especially big cities like Berlin and Hamburg were forced to treat their sick in provisional isolation wards – sometimes even in tents – because hospital beds were lacking. And, the problematic is by no means resolved as limits in the production capacity and delays in the supply of vaccine occurred as recently as the 2009 H1N1 pandemic.[8]

The above sketch has shown to what considerable extent social and biological conditions boost the expected costs of epidemics. Since the final size of actual outbreaks depends crucially on the time lag between the occurrence of the first cases and the implementation of appropriate countermeasures, the health planners choice of prevention strategies will finally be guided by the potential harm of delayed response. In fact, the model has so far assumed that health authorities are able to detect disease cases and take appropriate countermeasures with no delay. In the real world, there will – however –be a time lag between an outbreak and the implementation of containment policies. During this "reaction period", the disease will spread unimpeded from one country to the next. The number of imported infections will therefore be larger the more time it takes to detect and contain the outbreak. The adverse effect from higher numbers of initial infections will – however – vary according to a country's environmental conditions. The public health response is particularly crucial the higher the population's

[8] See: Abramson, John Stuart, Inside the 2009 Influenza Pandemic, New York 2012, chapter 1. Moreover, given that even an industrial country like Germany disposes nationwide of 5(!) ICU wards to treat Ebola victims, the possibility that hospital capacities might fall short of the needs in case of an outbreak does not seem too far-fetched

epidemic spreading capacity β and the longer the infective spell $\frac{1}{\gamma}$. The reason is that - if the system allows for an outbreak - the introduction of a newly infected individual results in more than one secondary infection. In this case, the initial growth rate of the infected compartment obeys to the condition $\frac{dI}{dS} = -\frac{\beta\varphi(h)S(0)-\gamma}{\beta\varphi(h)S(0)} > 1$, which ensures that a marginal increase in the number of imported infections I_0 will produce *more* than one additional infection. In a less technical way, the vulnerability of populations whose characteristics facilitate epidemic spread (high β) or extend infectious periods (low γ) can be illustrated graphically. This approach is followed in figure 5, which plots total epidemic costs for countries with different R as a function of the initial number of infections. Longer reaction periods – and hence larger values of I_0 – always result in higher realizations of $C_{epidemic}$. The potential benefits from reinforced prevention are therefore greater the higher the probability of imported infections and the weaker a country's capacity to detect and contain outbreaks.

Figure 5: Total Epidemic Costs (y-Axis) as a Function of Delay in Response (Measured by the Number of Infected Persons I(0) at the Point in Time When Countermeasures are initiated)



We have just seen what determines the costs and benefits of national epidemic policies. Yet, in connected economies diseases spread freely across borders. Whether or not this will give rise to additional inefficiencies will be discussed in the following section.

3.3. Deadly Negligence? Accounting for International Disease Spillovers:

So far, the problem of the health planner has been considered in isolation. In fact, epidemic strategies used to be – and still are – often determined at the national level with no consideration

for the effect of local health initiatives on prevalence and disease costs abroad. Since germs pay no attention to political borders when spreading across populations, one country's prophylaxis is an unpaid factor in its neighbors' "production" of immunity. In order to assess the size of this Meadean externality (Meade, 1952), consider the following modification of the first order condition in which health planners fully internalize the effect of their prevention activities.

$$\frac{\partial \mathfrak{C}}{\partial h_i} = \pi_i \frac{\partial C^i_{epidemic}}{\partial h_i} + \pi_j \frac{\partial C^j_{epidemic}}{\partial h_i} + c^i_h = 0$$

The optimality condition for the international allocation differs from the national problem by the term $\pi_j \frac{\partial C^j_{epidemic}}{\partial h_i}$, i.e. by the beneficial impact of prevention on other countries' expected disease costs. Following the tradition of public economics, the size of the prevention externality in country i can be measured as the Pigouvian subsidy which would induce its policymakers to impose the internationally optimal health investment h*.

Considering $\pi_j \frac{\partial C^j_{epidemic}}{\partial h_i}$ constant as a national policymaker would do and integrating with respect to h_i, we find the size of the per-unit subsidy required by country i to be equal to

$s = -\pi_j \frac{\partial C^j_{epidemic}}{\partial h^*_i}$.

As a matter of fact, this will always be positive unless the populations under study are completely isolated – case in which $\frac{\partial C^j_{epidemic}}{\partial h_i} = 0$ and prophylactic efforts in one country have no impact abroad. [9]

Using our previous results on the way epidemics spread, there is – however – a lot more to learn about the potential size of the externality under varying initial conditions. First, we observe that the subsidy s* is higher, the higher $\frac{\partial C^j_{epidemic}}{\partial h_i}$' for given values of h. This implies that prevention externalities are increasing the closer the *connectivity* between two countries. Moreover, since the decentralized planners will react to high exogenous disease risk by reinforcing national health efforts, the relative importance of the externality for country i declines if its neighbor's exposure π_j – and hence foreign prevention effort – increases. Turning to the epidemiological parameters β and γ, it is useful to assume that the biology of the pathogen does not vary across regions (i.e. that γ is the same in all countries). Then, limiting the duration of the infective spell will mitigate the marginal impact of health interventions no matter whether imposed at home or abroad. As a result, the partial derivatives of $C_{epidemic}$ with respect to h are smaller for given prevention levels and both countries will reduce their corresponding effort. Matters are more involved with respect to spreading capacity β_i which has a different impact on domestic and foreign prevention policies. Thus, if health efforts are substitutes, countries whose conditions facilitate rapid contagion suffer more from disease spillovers when they border regions where

[9] Remember that increased prophylaxis reduces the expected costs of an outbreak

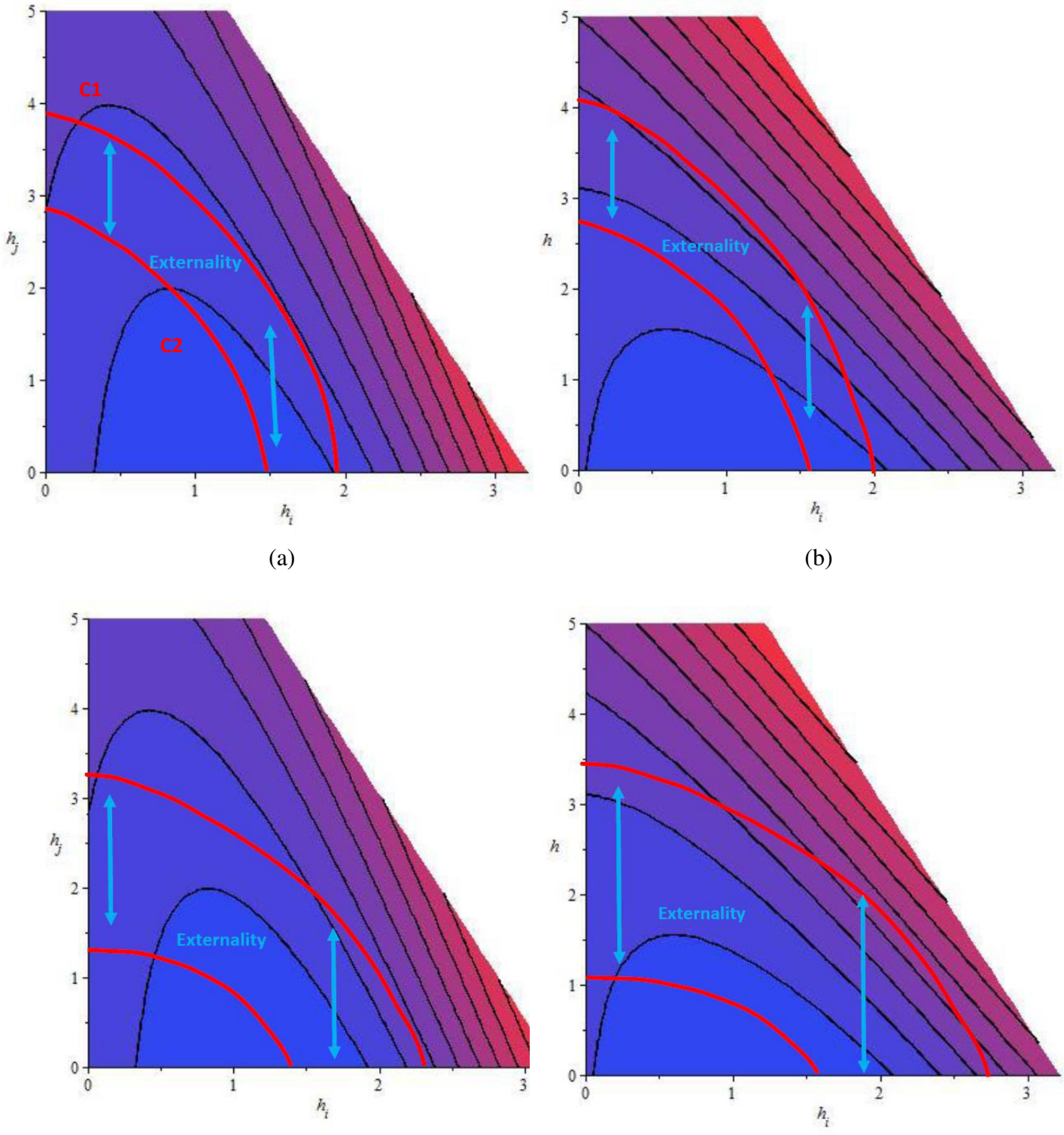
epidemics hardly spread. For again, ignoring the potentially disastrous impact of imported infections, their neighbors will exert little effort in prevention which - given that $\frac{\partial C^{j}_{epidemic}}{\partial h_i}$ is higher for small h - will cause the distortion to grow.[10] Conversely, countries which are adjacent to regions with high epidemic risk will benefit from the other states' efforts at detecting and containing outbreaks and experience only minor prevention externalities. In sum, the adverse impact of national prevention activities falling short of the international optimum is larger the closer the connectivity between countries and the greater their disparities in epidemiological conditions.

As before, the intuition of the mathematical results can be conveniently illustrated by means of numerical and graphical analysis. Figure 6 shows the size of the externality under varying environmental conditions. Panels (a) and (b) capture the effect of increased connectivity on the health costs of country i. The red curves correspond to all pairs of h_i and h_j which satisfy the optimality condition in country i and j if the policymakers act exclusively at the national level (C1) or if they account for the marginal impact of domestic prevention in adjacent countries (C2). Put differently, C1 and C2 denote all possible combinations of prevention strategies which would be chosen under international cooperation (C2) or in absence thereof (C1). The avoidable costs which arise from the international coordination failure can be read from the different intersection points between C1 and C2 and the country's iso-costcurves. As the graphs show, disease spill-overs create excess costs in both cases, but since the distance between C1 and C2 is lager in (b) than in (a), the additional harm is obviously increasing in connectivity.

This is also true if countries are more heterogenous with respect to their epidemiological characteristics. The corresponding effect is depicted in panels c-d. In both cases, country i's own epidemic spreading capacity β_i is fixed to some constant level. Yet, in the first case, country i is assumed to border a region whose epidemic risk β_j is sufficiently *high* to promote prophylactic activities in j and – consequently – to reduce the scope for international prevention externalities. The scenario in the last figure (d) is different. Here, β_j is low in comparison to β_i. Country i's efforts to cope with its permanent epidemic threat will therefore reduce the number of infections imported into country j to negligible levels. At the same time, the low spreading capacity of the population in j will discourage its policymakers from taking serious – and costly – health action. While this epidemiological free-riding entails only modest changes in the expected size of an outbreak in j, the consequences for its epidemic shield, country i, will be severe with larger disease spillovers and dramatically increased health costs.

[10] Note that for interior equilibria, the second-stage adjustments $\frac{d\,h_i}{dh_j}$ remain of second order in magnitude.

Figure 6: Changing Size of Externality in Response to Increased Connectivity (Panels a-b) and Increased Epidemiological Heterogeneity Panels (c-d)



Note: The cost differential between any two iso-cost curves is the same across all graphs. Hence, the externality is larger the more iso-cost curves lie between C1 and C2.

As we close the theoretical analysis and move back to the historical case of Germany, two points warrant closer attention. First, connectivity is not stable over time. For, as trade and labor mobility increase, commuters, migrant workers or refugees create epidemiological links between formerly unconnected regions. This is all the more important as the marginal return to prevention, $\frac{\partial \pi}{\partial h}$ and with it the degree to which countries are affected by epidemiological connectivity, decreases if health efforts are reinforced. Conversely, this implies that the impact of rising integration and mobility will be maximized in regions with a large number of heterogenuous countries and underdeveloped public health systems. In particular rapidly growing urban centers are then at risk of falling victim to negative disease externalities of dramatic dimension: As the interdependency between metropolitan areas and the periphery becomes closer, health policies undertaken in backward rural regions can influence the epidemic risk of densely populated industrial areas long before urban communities establish institutions which ensure appropriate preparedness and prompt reaction. Thus, low population density and mobility limit the immediate harm of lacking capacity or expertise and meager commitment of local authorities. Yet, this check will be removed once the disease is carried into a population which facilitates its epidemic spread.

A second factor relates to the possibility of eradicating diseases - or at least - of preventing them from getting endemic. As the analysis of the SIR model has shown, these aims can be attained if and only if prophylaxis is at a level which ensures that the basic reproduction rate R is strictly smaller than 1. However, if prevention is costly and marginal returns are decreasing, the fact that health policies are taken by national decision-makers can easily cause the levels of prevention h_i and h_j to remain below the critical threshold of herd immunity. In consequence, mitigating the prevention externality by creating avenues to coordinate epidemiological strategies at the international level carries potentially large benefits for all countries. For, only if the under-provision of prophylaxis is reduced, getting rid of the epidemic threat will become possible.

The previous section might contribute to our understanding of past epidemics. But even those who do not care for history might be left with awkward feelings: Seeing that the volume of international trade and the speed of mobility rise at staggering rates, *not* being concerned by the potential impact of rising epidemic spreading capacities in insufficiently protected populations would require a lot of carelessness – if not downright fatalism. Still, these risks would be mere mathematical artifacts if real policymakers found ways to avoid them. Let us therefore return to the case of smallpox prevention in to assess whether the historical evidence supports the notion of sizeable prevention externalities.

4. Coordination Failure in History: Empirical Analysis of Germany's Smallpox Prevention:

If the predictions of the model applied to the German case, one would expect states which were highly exposed or whose socio-economic conditions facilitated the spread of epidemics to adopt more comprehensive prevention policies than less vulnerable countries. The following section tests this conjecture by addressing a simple question: Under which conditions were German

governments willing to introduce and enforce mandatory vaccination laws? To investigate this issue, let us consider a standard probit model which estimates the probability of a state requiring its citizens to get immunized as a function of epidemiological exposure.

$$VAC_{i,t} = x_1 River_i + x_2 Exposure_{i,t} + x_3 Shield_{i,t} + Regional\ Controls + Time\ Controls + \varepsilon_{i,t}$$

Besides the dependent variable VAC[11] which takes the value of one if mandatory vaccination laws are in force in country i in year t and zero otherwise, the model relies on the following measures of exposure:

- **River** which denotes the length of navigable rivers within the state's territory. The intuition behind this variable is that throughout most of the 19th century the bulk of interregional trade relied on rivers as the dominant means of transportation. Since trade was a frequent catalyst for the spread of epidemics, regions which were connected to the river trade network were arguably more exposed to imported epidemics than states which were not.

- **Exposure** which gives the length of a state's borders in relation to its surface. High values of border thereby indicate increased exposure because the risk of infectives entering the country and widely spreading the pathogen before the epidemic can be detected and contained should be highest for relatively small, densely populated countries with long borders. As in the case of the previous variable "River", "Exposure" is constructed based on historical GIS data.[12]

-**Shield** which captures the possibility of states freeriding on the health efforts of their neighbors. Shield is a binary variable which takes the value of 1 if an adjacent state of country i has implemented mandatory vaccination laws and zero otherwise.

[11] The Necessary information was drawn from the final report on the 1870-74 pandemic contained in*: Imperial Health Department (ed.) Mitteilungen aus dem Reichsgesundheitsamt, various issues, 1881-1884.* These publications include a history of smallpox in Germany and state whether and when a state had introduced mandatory vaccination laws.

[12] The corresponding source are own computations based on the HGIS database of historical geo-information data constructed by the university of Mainz. Available online at www.hgis-germany.de

The summary statistics of my sample and the results of the estimation are depicted in tables 1 and 2:

Summary Statistics:

Observation Period	***1800 - 1870***			
Level of Observation	German states and Prussian Provices			
		River	**Exposure**	**Shield**
Number of States with Mandatory Vaccination Laws during Observation Period	18	665.38 (Std. Dev. = 997.6)	235.7 (Std. Dev. = 166.55)	0.44 (Std. Dev. = 0.51)
Number of Non-Vaccinating States	20	314.5 (Std. Dev = 524.2)	158. 67 (Std. Dev. = 217.10)	0.7 (Std. Dev. =0.47)

Estimation Results: Probit Estimation, Dependent Variable = VAC

Variable	***Coefficient***	***Std. Err.***	***z***	***P>|z|***
Exposure	0.0033	.0012	2.51	0.01
Shield	-1.2628	.5467	-2.31	0.02
River	0.0008	.0004	2.00	0.03
Constant	-0.3774	.4216	-0.89	0.37
Pseudo R^2	0.22			

The signs of all coefficients are in line with the predictions of the theoretical model: In particular, it seems that epidemiological exposure put some German states under pressure to strive for universal immunization whereas their exogenous conditions let others grow lax on

prevention. Corroborating this notion, the coefficient of "shield" is statistically highly significant and carries a negative sign. Hence, policymakers were apparently aware of other countries' prevention efforts and gladly seized the opportunity to rely on well protected neighbors instead of strictly enforcing - often unpopular - vaccination laws themselves. This observation is not only important because it suggests that epidemiological free-riding *did* take place in reality. Given that most states had adopted their vaccination policies in the first half of the century it also points to the remarkable persistence of the phenomenon. The final account of the 1870-74 epidemic in Germany is therefore a sobering message. For indeed, the *Reichsimpfgesetz* ("National Vaccination Law") eliminated prevention externalities and robbed the smallpox virus off its potential reservoirs within German territory. But this failure to coordinate had existed for decades and it seems that only the catastrophe of the epidemic could break the ubiquitous inertia and pave the way for a serious modernization of the prevention infrastructure.

5. Making Prevention Work: The Case of Baden and Wurttemberg

The fact that variation in exposure and epidemiological free-riding are key factors in explaining international differences in immunization policies is an important insight from the German smallpox epidemic. Yet, the case offers more lessons to learn. For one, safe and effective vaccines are not available for many diseases. And even when vaccination is possible, it would be naïve to believe that mandatory vaccination laws *alone* could guarantee high immunization rates. Let us therefore scrutinize the experiences of the German states in more detail and address two additional questions: How does the success of the vaccination laws depend on the way the regulations are implemented? And second: Given that health institutions were tailored to specific environmental and socio-economic needs, does the historical case suggest there are exogenous conditions which facilitate the emergence of institutions which are sufficiently flexible to deal with *all* sorts of different infections?

Addressing these issues requires an in-depth analysis of statistical and archival sources. Performing this analysis for the entire sample of all German states is not only beyond the scope of this paper. It is simply impossible because corresponding evidence has not survived in many cases. We will therefore follow a more modest approach by considering only the case of the two South-Western states of Baden and Wurttemberg.

The experiences of these states lend themselves to this purpose for various reasons: First, regional and local archives in the present state of Baden-Württemberg feature a wide range of statistical and qualitative records which are exceptional both by the standards of the 19th century and in comparison with other German regions. Moreover, with regard to initial conditions the case is an almost perfect miniature version of the later German Empire. Thus, Baden was arguably highly exposed to contagion from abroad while Wurttemberg was in the position to free-ride on the prevention efforts of its neighbor. Adding to that, the two were not only representative of the German Empire in the 19th century. Their health authorities were also familiar with a number of difficulties which keep bothering their present-day peers

including problems to detect and contain outbreaks in a region with extensive trade and labor mobility.

In order to learn, what insights Baden's and Wurttemberg's struggle with smallpox carry for present-day prevention strategies, the next section will briefly sketch the two states' epidemiological environment and the historical evolution of their vaccination systems. With this knowledge at hand it will be possible to analyze how their different initial conditions determined the long-term characteristics – and the ultimate success – of their anti-epidemic policies.

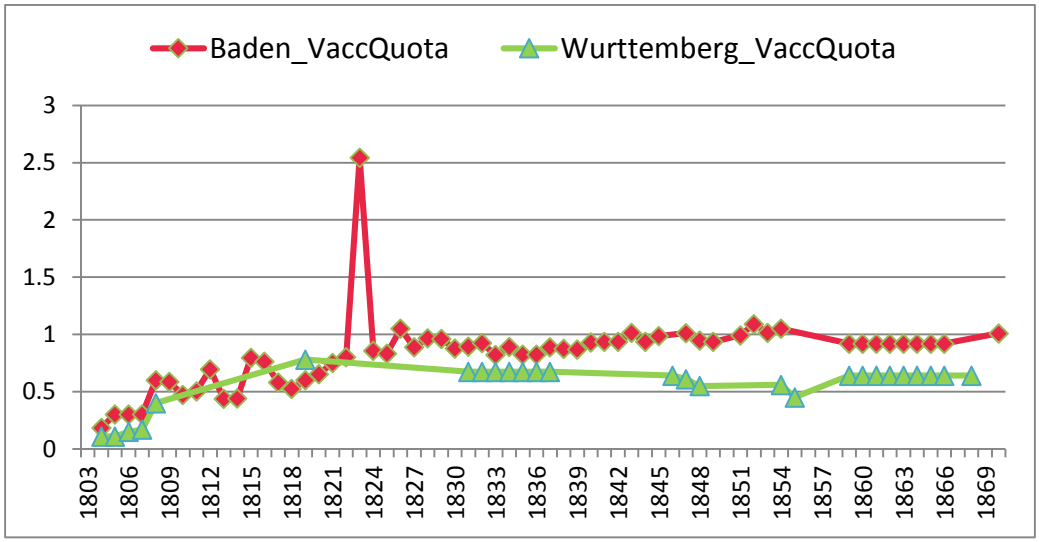
5.1. The Epidemiological Conditions of Baden and Wurttemberg:

Vaccination had a promising start in Baden and Wurttemberg alike: In Baden for instance, the first cowpox vaccinations had been performed in 1803 and as soon as 1804, *officially confirmed* inoculations amounted to a respectable 14.773 with public health officials estimating a total number of operations of about 25.000 (Maler, 1804, p. 62). Similar success was recorded in Wurttemberg, where numerous doctors offered free vaccinations with some of them spending enormous time and effort to expand the measure to remote hamlets and tiny villages.[13] However, it became soon apparent that indifference, ignorance, poverty, free-riding or sporadic cases of fundamental opposition would prevent universal immunization if the measure remained entirely voluntary.[14] Both states therefore first introduced a number of indirectly coercive measures like the mandatory vaccination of recruits, orphans and students in public schools. As with the *Reichsimpfgesetz* of 1874, the final trigger for the general introduction of compulsory vaccination came with a smallpox epidemic which hit both states in 1815 and created the political climate for the vaccination laws of 1815 (Baden) and 1818 (Wurttemberg).

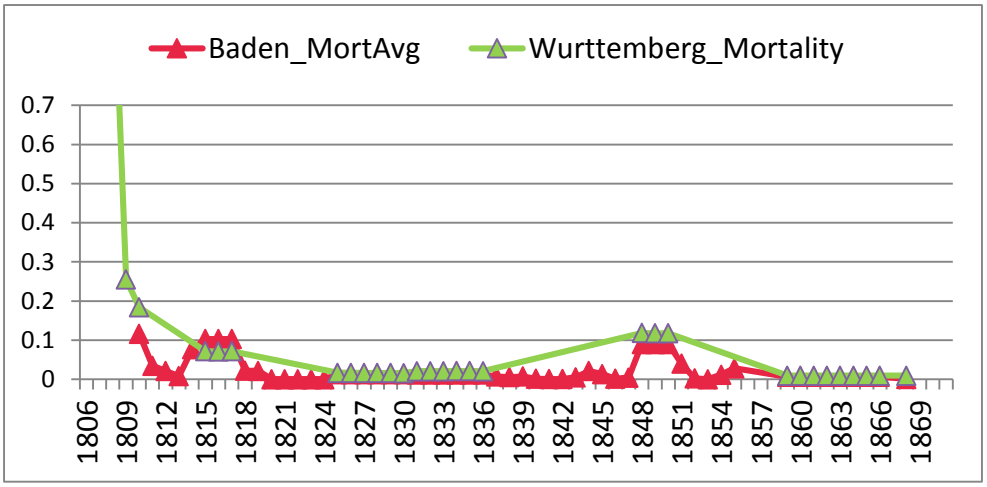
[13] Schwäbische Chronik, 16. 10. 1801, p. 365; M. Lächlin, Geschichte der allgemeinen Einführung der Kühpockenimpfung in Botnang in den Jahren 1801, 1802 und 1803, unpublished manuscript 1804, reprint in: G. Cless, Impfung und Pocken in Württemberg, Stuttgart 1871, p. 105-108

[14] In Baden the corresponding entity was termed *Sanitary Commission*, in Wurttemberg *Collegium Archiatrale.*

Diagram 1: Smallpox Vaccination Quotas, Prevalence and Mortality in Baden and Wurttemberg: (A) Vaccination Quotas: Ratio of Number of Vaccinations and Number of Births,



(B) Mortality: Smallpox Deaths per 1000 Inhabitants:



Note: For 1814-1817 and 1848-1851, Wurttemberg sources indicate only average mortality. To ensure comparability, the annual data from Baden have been converted to averages for these periods.

The depicted data reveal that the initial impact of the mandatory vaccination laws was similar in Baden and Wurttemberg. In particular, vaccination rates did not respond immediately to the new laws but went through a transition period before adjusting on higher levels. Be it for institutional constraints, opposition or lack of infrastructure, *mandatory* vaccination had no sizeable epidemiological impact but with a considerable time-lag. However, in later periods variation in immunization rates and disease prevalence was surprisingly strong for two states with almost identical legal framework. Put differently, if the diverging long-term experiences

did not stem from the laws themselves they must have been caused by differences in the epidemiological or institutional environment. This fits well with the observation that the vaccination rates kept rising in both states as long as the disease was endemic and sufficiently present in collective consciousness. Yet, as the policies translated into decreasing infection rates and as it became apparent that the protective capacities of vaccination were temporarily limited, public and private incentives for prevention diminished and increased the importance of both the two countries' natural conditions and of effective governance.

We have argued, that institutional differences largely emerge from variation in environmental conditions. But how did the two states' epidemiological environment differ *concretely*? The most important differences stemmed from the geography of South-Western Germany. First, Wurttemberg extended over a larger, often mountainous territory, while Baden was almost entirely made up by the valley of the Rhine and adjacent plains. As a result, extending medical coverage to the countryside was harder in Wurttemberg than in its western neighbor. In addition, Wurttemberg was in the fortunate position to be almost exclusively surrounded by states which promoted comprehensive immunization if they did not impose mandatory vaccination themselves. Baden lacked such an epidemiological "buffer zone" and faced neighbors whose prevention policies were at best lax by the country's own standards. Thus, in Switzerland vaccination was neither mandatory nor particularly encouraged by the state. And in the adjacent French departments of the "Haut Rhin" and the "Bas Rhin", doctors were in principle required to vaccinate free of charge, but immunization rates in Alsatia nevertheless failed to reach the Baden level.

Taking the time to consider the Alsatian case is worth the effort because certain deficiencies of the French system were well-known to the Badenese health department (the "Sanitary Commission") and strongly influenced its attempts at implementing more efficient institutional solutions. To begin with, the Sanitary Commission observed that none of the two départements set incentives for vaccinators to fulfill their duties with great zeal. In the southern departement du "Haut Rhin"for example, medical practitioners received neither travel allowances nor a fixed salary or a fee for inoculations. As in most other parts of France, only very few vaccinators who surpassed a threshold number of operations were awarded a so-called *medaille de vaccine* and a prime which ranged from 200-300 Fr. (Darmon (1986), p. 259). Still, seeing that a doctor could earn about 300 F per month if he practiced only in his resident community, this was a meager incentive especially in rural regions where the numbers of vaccinees were small and the ways to reach them long and exhausting (Darmon (1986), p. 259)
The incentive situation looked slightly better in the northern "Departement du Bas-Rhin" where the government installed public health officials (*Médecins Cantonaux*) who received a fixed salary and were in charge of vaccinating and providing fundamental health care to the poor. However, the salaries of the Alsatian *Médecins Cantonaux* were grossly inferior to those of German public doctors like the Badenese *Physicus* or the Wurttemberg *Amtsarzt*.[15] As a result, the Médecins Cantonaux felt understandably little inclination to practice in remote areas, let alone to document their vaccinations with sufficient diligence. In consequence, the French authorities were unable to attain high immunization rates and to contain recurrent epidemic

[15] A *Physicus* or *Amtsarzt* received between 350 and 500 fl per year, which was more than twice the 600 - 1000 Fr (= approximately 140 - 234 fl15) of the Alsatian *Médecins Cantonaux.*

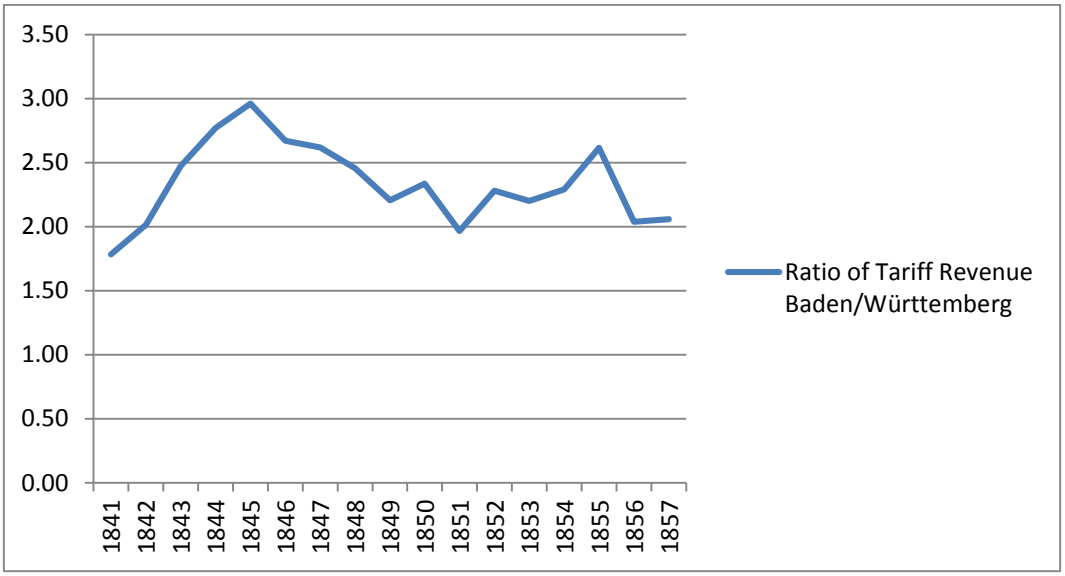
outbreaks. At least in the eyes of Badenese observers France was therefore a constant health threat and a notorious hotbet for imported smallpox infections.

That this situation convinced the Sanitary Commission of the necessity to install institutions which promised almost universal immunization and immediate detection of eventual outbreaks was also due to the country's economic interests: The Badenese fisc took enormous benefit from the tariff revenues generated by the international trade on the Rhine. Moreover, the well-being of the local economy in the border region depended critically on the integration of the goods' and labor markets on both sides of the river. Countering epidemic threats by isolationist policies was therefore no option for Baden's decision-makers.

In sum, it is evident that Baden *was* more exposed than Wurttemberg. In contrast to the sample of *all* German states, Baden's and Wurttemberg's epidemiological exposure can - however - be quantified by a more specific measure than the length of their borders. To do so, consider the two states' foreign trade: For Baden, trading activities came down to interacting with countries that followed an epidemiological laissez-faire approach like France or Switzerland. On the other hand, the vast majority of travelers and goods entered or left Wurttemberg via another vaccinating German state, be it Baden, Bavaria or Hessen. If at all its citizens therefore interacted with foreigners who were either vaccinated or had traveled across the epidemiological "buffer zone". Chances were consequently high that potential carriers of the disease fell ill (and were quarantined!) long before they reached Wurttemberg territory. These different levels of exposure can be quantified from the 1840s onwards when Baden and Wurttemberg entered the German Customs Union (Zollverein). The Zollverein issued annual statistics which give the tariff revenues of all import-, export- and transit- tariffs collected by the single member states.[16] These data provide a relatively accurate basis of comparison because the Zollverein had removed all *domestic* custom barriers such that economic exchange between Wurttemberg and its well protected, neighbors does not appear in the statistics.[17] Based on this indicator, Baden's contact with insufficiently immunized foreign countries – and therefore its exogenous epidemic risk – was almost consistently twice as high as the corresponding Wurttemberg values. This result is even more striking if the fact is taken into account, that the Wurttemberg territory was about twice the size of Baden and its population far more numerous.

[16]It is important to note that among all the statistics which were published by the Customs Union, only these data can be used to construct an indicator of a country's interdependence with external trading partners. This is the case because the tariff *incomes* received by each member state did not reflect the actual tolls collected at its frontiers. Rather, customs at foreign borders were collected on joint account, and the proceeds were distributed in proportion to the population and resources of the members.

[17] In addition, the Customs Union levied specific (per unit of quantity) tariffs. Revenue was therefore proportional to the amount of goods traded – and thereby probably also to the intensity of border traffic..Finally, the statistics reveal that the composition of Badenese and Wurttemberg imports and exports strongly resembled each other. Hence differences in the two states' revenues did not stem from variation in tariff rates but reflected indeed different quantities.



Source: *Statistik des Verkehrs und Verbrauchs im Zollverein*, various years

That these conditions determined the choices of the health bureaucracies in both states emerges clearly from archival sources. Thus, in 1830 the Wurttemberg health department declined an initiative for reinforced controls and more intensive vaccination coverage in the border regions by pointing to the fact that the small number of imported smallpox cases did not justify the costs of the intervention. On the other hand, Baden promoted revaccination in its entire periphery – a policy that was generally accepted and – in cases of epidemic outbreaks – even explicitly demanded by the local population.[18]

In fact 19th century Baden and Wurttemberg provide almost a textbook example of historical path-dependency and institutional inertia causing decision-makers to stick with organizational arrangements even if they had been tailored to conditions which had become obsolete long ago. This aspect was particularly evident in the area of vaccine procurement and the every-day implementation of the vaccination laws. Both these fields were critical for the ultimate success of public prevention efforts and – surprising it might seem – it was precisely for its unfavorable starting point that the Badenese health system ended up on a superior development path than the Wurttemberg one.

High immunization rates were unfeasible in both countries unless their health authorities managed to enlist the support of physicians and patients. To this end, they had to make sure the effort costs of all individuals involved in the process did not rise to prohibitive levels. In order to proceed, we will therefore need to assess which factors determined the vaccination costs of 19th century doctors and patients.

As a matter of fact, direct monetary costs exerted no decisive influence on the immunization choice of citizens in Baden and Wurttemberg alike. The reason was that vaccination fees were simply too low to effectively deter inoculations: Prior to the mandatory vaccination act of 1818, getting a child inoculated would cost a Wurttemberg family 15 Kreutzers. After 1818 this price was reduced to 12 kr. Compared to the estimated level of 1815 food prices, this implies that the fee of a vaccination amounted to no more than the price of 5 pounds rye bread or 1.5 pounds

[18] Multiple Examples are provided in: GLA Karlsruhe, 236-16042 (Sannitary Comission), „Epidemic outbreaks 1827-1871“

pork (Wolff(1998), p. 165). Performing the same exercise for Baden, one finds that a vaccination would cost between 27 and 18 Kreutzers which translates into 10 (7) pounds of bread or 1.2 (0.8) pounds of pork in 1806 Mannheim prices (Krauß, (1993), p. 89). Laws and ordinances in both states further stipulated, that vaccinations had to be offered free of charge to all families who received poor relief or were classified "indigent" by their municipality. Data from Baden reveal that depending on the region this classification applied to 10-30% of the population. This fraction was even higher in many rural districts where inhabitants of remote farms were routinely treated as "poor" which could bring the share of publicly financed vaccinations to more than 50 percent.[19] Hence, for those individuals for whom foregoing 10 pounds of bread constituted a serious constraint, vaccination was costless anyway. Turning to physicians, one finds piles of complaints and petitions which were submitted to the health departments of the two states. Yet, Baden's and Wurttemberg's vaccinators hardly criticized the size of their remunerations. If at all, they were bothered by the bureaucratic wrangling to obtain them: While wealthy individuals covered the vaccination fees themselves, those of the poor where paid out of communal or central state's treasuries. As a result, vaccinators received no public funds unless they listed all vaccinations by date, place and the vaccinee's income status. Still, even if this tedious task was carefully accomplished, there was no way that physicians could be sure to receive salaries and travel reimbursements in due time because the intricate appointment schemes were a frequent source of dispute between communal and central entities.

If monetary costs were no major obstacle to the functioning of the vaccination system, travelling expenses definitely were. What considerably trouble patients and doctors encountered when attempting to meet with each other is vividly depicted by the following mid-century testimony of a Baden country doctor:

> "Running a medical practice in the countryside presents far greater a challenge to the physical strength and professional abilities of the doctor than practicing in the city. Especially in the mountains this task calls for strong men who can bear the ailment of laboring day and night no matter how adverse the weather conditions.[...] There was rarely any night which passed without the door-bell ringing me out of bed and calling me to visit a patient in the town or some village; there was no way, I could have eaten my meals at regular hours and finally I resolved to wearing far too light clothes [for the cold season] because they would not impede my movement on the long horserides. In the end, even my horse could no longer bear the daily drudgery and I needed a second one [...] Two of the assistant doctors I later had in Freiburg, competent and diligent young men, became the victims of their profession and died after a short time in the mountains, one of them having practiced in the same district as I." [20]

Exhausting long-distance walks or horserides of this kind were by no means uncommon and – given that immunization was a matter of the active and healthy rather than the bedridden sick – many vaccinators spent hours on the way to some solitary hamlet just to learn at their arrival that the children they were supposed to vaccinate were in poor health or had been sent to help on the fields. Even successful visits in the countryside had little appeal for physicians because

[19]GLA Karlsruhe 236-16026(Sanitary Commission), Classification: "Medical Treatment for Indigent Persons"
[20] Adolf Kussmaul, Jugenderinnerungen eines alten Arztes, Reprint München 1960, p. 295

the fees for inoculating the few children of a sparsely populated area hardly justified the effort supply the service.[21]

As for parents, getting their offspring immunized implied a sizeable income loss. And this was not only the case because – unlike their brave country doctor – the inhabitants of remote villages had no claim to wage subsidies for a horse. What awaited those who took the trouble of a many-hour trip on more or less accessible field tracks had little the pain- and harmless kind of vaccination we know today. Rather, contemporary medical textbooks recommended to inflict no less than 5 crossing cuts of 3cm length and 1cm depth to both forearms of the inoculated person. The edges of the wounds would then be forced apart to introduce the vaccine – dried lymph from previous vaccinations, which was usually diluted in water, but in less lucky cases also in the vaccinator's saliva.[22] Finally, the vaccination-induced cow-pox infection typically entailed two to three weeks of painful local skin eruptions and flu-like symptoms like fever, nausea and headache, possibly coupled with acute conjunctivitis. At least one adult would therefore have to refrain from income-earning activities in order to take care of the vaccinated children. This represented a considerable burden especially during harvest times or in poor families, where all hands were needed to contribute to the household's living. Not surprisingly so, many documents of patient behavior refer explicitly to opportunity costs like the following testimony of a Baden peasant who justified his refusal to have his children vaccinated by claiming that "he could not afford having his children made sick in the middle of the harvesting season."[23]

Seeing the above, there is no denying, that a 19th century vaccination was no pleasant experience. Things could – however – take a much worse turn if vaccinators were poorly trained or failed to exert their duties with due care. In some cases, for instance other diseases – especially syphilis – were transmitted through contaminated vaccine or surgical instruments. Relatively rare events they were, these incidences inevitably attracted public interest and could prompt a veritable hysteria even if the presumed threat was unfounded.[24]

Lack of doctors' commitment did not only contribute to patients' vaccination costs by increasing the expected discomfort and the fear of harmful side-effects. It also bothered health authorities because unlike present-day vaccines, cowpox vaccinations relatively often failed to provoke a sufficient immune response. As a result, vaccinators needed to assess the success of the artificially induced infection 5-7 days after inoculation.[25] In regions like the mountainous tracts

[21]GLA Karlsruhe 236-16026 (Sanitary Commission), Report of the *Physicus* of Pfullendorf, April 13th, 1819
[22] For historical details on vaccination practices see:A. Bousquet, Traité de la Vaccine, Paris 1833
[23]Instruction of the Sanitary Commission, August 2nd , 1815 (Baden), Vaccination Instruction, June 25th 1818 (Wurttemberg)

[24]A typical example relates to a vaccinating surgeon in rural Wurttemberg: In the 1830s, he found no infant among his recent vaccinees whose parents would consent in having lymph taken from the pustules of their child. Having no vaccine stored he finally resolved to using the lymph of a child whose mother suffered from psychosis and periodical seizures. Fearing that the mother's mental disorder could be transmitted through the vaccine, the villagers refused any inoculation and chased the surgeon from their village. One year later, when the health officer from the district capital came to investigate the issue and to perform the missing vaccinations, he was welcomed with "stones and pitchforks" See: StAL E162-I-2133, Internal Report on Vaccination to the Ministry of Inner Affairs, Stuttgart 1831

[25]The need to supervise the success of the vaccination remained a problem of smallpox immunization programs until the WHO global eradication campaign of the 1970s.

of Wurttemberg where many physicians demonstrated little diligence in fulfilling this duty, a substantial number of failed vaccinations would therefore pass unnoticed. The official vaccination lists therefore gave a biased impression of immunization levels and rendered a prompt containment of disease outbreaks virtually impossible.[26]

In order to minimize the number of failed vaccinations and to reduce the risk of tampered vaccine getting into use, those who did the operation had to be provided with sufficient amounts of virulent cow-pox lymph. This was no simple task because techniques to preserve and cool the vaccine remained on rudimentary standards throughout most of the 19th century. The lymph used in the process was either derived from the dried material of some previous vaccination or was directly transferred from the pustules of one vaccinee to the next. In the first case, chances were high that the lymph degenerated during storage and failed to cause the immunizing outbreak of cowpox. In the second case, doctors needed an unbroken "supply" of vaccinees in order to maintain the process of vaccine transmission.[27] Due to these constraints, problems of vaccine procurement were particularly acute when vaccinations were performed on a private basis instead of comprehensive public vaccination campaigns.[28] In the latter case, vaccinations could be scheduled such that the recently vaccinated children of one community provided the lymph for the vaccinations in the next. By contrast, private vaccinations relied on the swayings of individual demand and were not precisely plannable in advance. The ensuing uncertainty increased effort costs of the vaccinator because patients' refusal to provide lymph from their (or their children's) pustules frequently cut off doctors from any supply of new vaccine.

6. The Design of the two German States' Epidemic Coping Strategies:

In sum, the health departments in Baden and Wurttemberg needed to conceive institutions which solved the following problems: First, the vaccination infrastructure had to be designed in a way that minimized the traveling expenses of vaccinators and vaccines, second it had to mitigate fears of adverse side-effects and finally it necessitated a constant supply of high quality vaccination lymph. Abstracting from minor details, the two states' responses to this challenge differed in three aspects all of which can largely be traced back to Baden's increased exposure to imported epidemics. For, what the country's authorities had to accomplish was to establish

[26]Especially in Wurttemberg, the impression of a rather perfunctory attitude towards documentation is further supported by various pieces of anecdotal evidence: In one case, a physician indicated that the vaccination reports of his predecessor were utterly unreliable if not forged because they indicated not a single unsuccessful vaccination - instead of the usual 10-20% failure rate. What is more, the lists also contained passages in which dozens of vaccinations and follow-up visits had apparently been filled in just a few days before their official submission to the health department. The second example involves a doctor who was reprimanded for shirking his documentation duties because he not only refused to revisit vaccinees in rural areas but also delegated this responsibility to his untrained 18 year old son whom he moreover entitled to collect a private fee of 1 gulden per visit.

[27] Timing was all the more of essence as the lymph could be collected only within a limited time-frame: If the pustules were opened too early, they leaked only small amounts of fluid with little infectious material. If on the other hand the operation took place at an advanced stadium, the lymph would already be mixed with pus and debris which reduced the power of the vaccine and increased the risk of transmitting other diseases when inoculating new vaccinees.

[28]This particularity is discussed both by Cless (1875), ch1 and Franz Heim, Historisch-kritische Darstellung der Pockenseuchen, des gesamten Impf- und Revaccinationswesens im Königreiche Württemberg, Stuttgart 1838.

a system which allowed them to attain almost universal immunization and ensured immediate response in case of an epidemic outbreak.

To integrate as many persons as possible in the program inoculations in Baden were performed annually in the course of large public vaccination campaigns in May and June and October. Since the introduction of compulsory vaccination in 1815, the dates for these campaigns were carefully set to avoid conflict with the sowing and harvesting periods of the main crops. Equal importance was placed on the schedule of the subsequent vaccination campaigns to maintain the line of vaccine transmission unbroken.[29] This course of action represented a suitable compromise between the conflicting needs of patients and vaccinators. Due to the mass campaigns, physicians could conveniently tour the larger communities of their district and combine the assessment of old vaccinations with the procurement of new lymph from the recently vaccinated. On the other hand, although the inhabitants of remote farms or villages incurred travel expenses, this was a minor inconvenience because the regular vaccination dates allowed parents to plan the necessary trip in advance and to choose the closest locality for the operation.

Limiting travelling costs for patients and doctors was indispensable to guarantee the vaccination laws high compliance rates. Yet, trusting in compliance alone was a fallacy unless the health department could be sure that the operations were properly done and carefully documented. Badenese policymakers were therefore anxious to supervise the vaccinators' activities and to ensure a smooth flow of information between health practitioneers and the Sanitary Commission. As a second important difference to Wurttemberg, Baden therefore restricted the right to perform vaccinations to the district "*Physicus",* a licensed doctor who was hired and paid by the government. Only in very large or inaccessible districts, the *Physicus* was allowed to appoint additional vaccinators, but he remained the sole responsible for the implementation and documentation of the vaccination program. Things were profoundly different in Wurttemberg which conferred the right to perform cowpox inoculations to any licensed doctor or surgeon. This high degree of centralization and standardization proved crucial in several aspects: First, since the *Physicus* was legally obliged to inoculate *all* children in his district, the laws ensured that inhabitants of remote areas and the city poor – in other words, the clientele doctors would normally not bother for – were sufficiently protected. Moreover, the Baden government relied not only on its network of district *doctors*, but was also among the first to see for comprehensive professional training of nurses and midwifes. As a result, even rural areas were relatively well equipped with academically trained medical personnel who provided fundamental healthcare and facilitated the detection and containment of epidemic outbreaks.[30] Finally, the fact that the art of inoculating cowpox was part of the standard curriculum for future doctors and midwifes, arguably reduced the risk of failed vaccinations and complications like the unintended transmission of other diseases. The different degrees of professionalization in

[29]The corresponding laws and ordinances can be found in: (Baden) Philipp Carl Baur von Eiseneck, Sammlung sämtlicher Gesetze, Verordnungen, Instructionen, Belehrungen und Entscheidungen, welche in dem Gross-Herzogthume Baden über Gegenstände der Gesundheitspolizey erscheinen sind, Karlsruhe 1830, (Wurttemberg): Albert Renscher (ed.), Vollständige, historisch und kritisch bearbeitete Sammlung der württembergischen Gesetze, Stuttgart 1830;

[30] In fact, archival evidence shows that it took rarely more than a day for the district health officials to report cases of smallpox to the Sanitary Commission and to perform revaccinations if necessary.

the Baden and Wurttemberg health sectors mattered little during the early stages of smallpox prevention. In fact, when smallpox was still endemic, patients cared little whether they received a vaccination from a doctor, a surgeon or even the schoolteacher or some local clergyman. Moreover, as initial resistance against the new method often stemmed from folk beliefs and mistrust against allopathic medicine, local vaccinators were in a better position to lobby for immunization than academic doctors from without.[31] However, with rising vaccination rates and decreasing smallpox prevalence the balance changed. For, although surgeons and laymen had vaccinated with satisfying results, the quality of their services was virtually unpredictable for patients and the health bureaucracy alike. Wurttemberg with its small number of publicly employed and supervised vaccinators therefore recorded a number of unpleasant incidents ranging from the use of badly stored and ineffective vaccine to complete lack of follow-up documentation and from the accidental transmission of other infections to the issuing of forged vaccination certificates.[32] At a time, where the microbiological foundations of vaccination were unknown even to medical scholars, such complications reinforced reliance on folk medicine and fostered all sorts of pseudo-scientific antivaccionism in bourgeois circles. Before long, the wide-spread uneasiness turned into a veritable political force as it found a charismatic and energetic leader in the person of the Stuttgart doctor Carl Georg Nittlinger. By 1874 Nittlinger had produced 25 voluminous tracts against vaccination and while Baden's doctors and patients silently accepted immunization if they did not genuinely support it, Wurttemberg became an epicenter of German antivaccionism. In the infant years of smallpox prevention, the decentralized Wurttemberg system might have seemed a cost-effective way to ensure sufficient immunization coverage in a country with moderate epidemic risk. Yet, as time passed, it proved increasingly problematic because controlling the large and heterogeneous group of vaccinators was difficult. By contrast, the vaccinating doctors in Baden were tenured officials, whose salary and career – including the possibility to move from unpopular remote districts to the metropolitan areas of Mannheim, Heidelberg or Karlsruhe – depended on the degree by which they fulfilled the expectations of the health department. Hence their incentives to meet immunization targets and documentation standards were sufficiently strong to guarantee that Baden's reliance on few government-appointed doctors yielded superior results than the fragmented solution of Wurttemberg.

Speaking of the benefits from centralization, brings us to the third major difference between the two German states: the costs of procuring suitable vaccination lymph. Here too, the need to counter the risk of imported epidemics resulted in the Baden authorities being more inclined to take matters in their own hands than their Wurttemberg counterparts. In the latter, the collection of fresh vaccine was largely confined to private channels and individual patients' willingness to provide lymph for subsequent arm-to-arm vaccinations. For the first third of the century, Wurttemberg had not passed any regulations on vaccine procurement but a law which exhorted the district doctors to collect and store fresh vaccine and awarded a prime of 1 fl to any peasant who reported a case of natural cowpox among his cattle.[33] In contrast to these piecemeal efforts, Baden was fast in providing its vaccination policies with an institutional infrastructure. In

[31] Corresponding evidence is described in Loetz (1993), pp. 227-46
[32] Examples were drawn from StAL, E 162-I-2137
[33] Albert Renscher (ed.), Vollständige, historisch und kritisch bearbeitete Sammlung der württembergischen Gesetze, Stuttgart 1830

particular, the country was by far the first to establish public vaccination institutes (*Impfinstitute)*. These institutions provided free vaccinations by specially appointed doctors to anyone who allowed the vaccinators to collect fresh lymph once the pustules had reached the proper stadium. Moreover, the vaccination institutes were assigned the tasks of seeing for the immunization of orphans and the city poor, of receiving and storing fresh strains of vaccine from medical institutions abroad, of collecting and preparing the lymph from occasional eruptions of cowpox and of promptly supplying doctors with vaccine whenever needs were voiced. All in all, Baden had four such institutes which were strategically located to serve the Northern parts of the country (*Impfinstitut* Mannheim), the capital and adjacent central regions (*Impfinstitut* Karlsruhe and *Impfinstitut* Freiburg) and the very South (*Impfinstitut* Meersburg). They operated in relatively modest premises – the Mannheim institute for instance consisting of no more than three rooms. That the costs of installing and maintaining the network of vaccination institutes remained virtually negligible for the state's treasury, can be inferred from the surviving cost calculations of the Mannheim *Impfinstitut,* Thus, the tables reveal that the costs of establishing the institute and paying the salaries of two physicians amounted to 293fl 54kr. Assuming that this sum applied also to the three other institutes, the total costs of setting up the system in 1808 would have come down to 1178fl 55kr. In the subsequent years, the Mannheim institute received a fixed budget of 192fl, whereas the smaller ones in Freiburg, Karlsruhe and Meersburg had to contend with 104fl each, which gives a total of 504fl for the four institutes.[34] Comparing these costs with other items on Baden's health bill shows that the vaccination institutes accounted for less than 0.5% of total health spending.[35] The numbers are – however – even more telling if the decentralized Wurttemberg system is taken as a basis for comparison: By law, Wurttemberg peasants were paid a prime of 2 fl for reporting cases of cowpox among their cattle. Once a case was reported, the district health official was required to collect lymph and to generate fresh vaccine by transmitting the animal lymph to a human vaccine. This process was to be repeated on at least two successive dates and if the operation was successful, the owner of the infected animal would be awarded an additional prime of 2 fl. The accruing travelling costs of the vaccinator were covered by the state at an average rate of 3fl. 30kr per day. In sum, obtaining a single dose of fresh vaccine would thus normally come at a cost of 11fl. After the introduction of the prime in the late 1820s, the yearly average of reported cases was 30, six of which would lead to a successful transfer of animal lymph to human recipients. In 1861/62, when the system had already reached a mature state of development, an average of 35 cases was reported, but again, the number of successful transmissions remained at merely 12. As a result, the sole procurement of original vaccine cost the Wurttemberg state between 282 fl. (average 1828-31) and 339 fl. (average 1861/62), which is equal to 56% and 67% of the costs Baden encountered for maintaining its four vaccination

[34]GLA 236-16034 (Sanitary Commission)

[35] For 1808 the sources list direct central government spending of 18730fl on poor relief and medical treatment for indigent individuals. Salaries for government-appointed health staff and costs for the maintenance of hospitals and asylums are not explicitly listed in the early issues of the national accounts but only figure in the section "public expenditure on police and health affairs". Yet, the numbers can be reconstructed using the more detailed budget bills of the Baden parliament. For selected years, these are available from 1819 onwards and depict health spendings that increased from slightly below 100.000fl in 1819 to 139324fl in 1831/32 and finally 163500fl in 1861/62.[35] Hence, vaccination institutes accounted for less than 0.5% of total health spending as stated above.

institutes.[36] Moreover, collecting vaccine was a loathed burden of district doctors, which reduced their commitment to the cause of vaccination and raised the state's supervision costs. This is not only reflected in a district doctor's sarcastic comment that he could find no point in "examining every single wart on the udders of the cows in the neighboring villages".[37] More seriously, the availability of vaccine remained a constant problem in Wurttemberg whereas in Baden it ceased to be a pressing concern after the establishment of the vaccination institutes. For, although complaints of vaccine shortage do appear in the early Baden sources, they vanish after 1810 and the highly detailed annual vaccination reports of the Sanitary Commission do not allow the matter more space than the standardized statement that the institutes successfully provided *all* district doctors with sufficient amounts of lymph.[38] How smoothly the system worked is strikingly apparent if actual outbreaks are considered. The Baden infectious disease regulations stipulated that all children or – in case of severe outbreaks – even the entire population was to be (re)vaccinated on the spot whenever outbreaks of smallpox were reported. The detection of disease cases therefore led to a sudden and unpredictable increase in the demand for vaccine. However, the network of the vaccination institutes proved able to deal with such challenges as a number of successful revaccination campaigns illustrates: A particularly impressive example for such interventions dates back to 1826/27 when Alsatia was hit by a large-scale smallpox epidemic that caused more than 12000 infections and claimed at least 2464 lives.[39] As Baden and Alsatia were economically far too integrated to suspend cross-country trade and labor mobility, the Badenese authorities implemented a comprehensive vaccination campaign in the entire border region. Due to this focused prevention effort the country managed to keep its 1826/27 smallpox mortality at the negligible level of 7 cases. In Wurttemberg on the other hand, lymph procurement not only remained a serious problem – with physicians' complaints of vaccine shortage numbering more than 30 for 1827 alone – but also prevented the fast containment of smallpox outbreaks. For, more than two decades after the 1826/27 revaccination campaign in Baden, in 1849/50, Wurttemberg was threatened by an epidemic wave entered the country from Switzerland and Baden. Yet, unlike its western neighbor 20 years before, Wurttemberg obviously failed to contain the outbreak as the sudden increase in the mortality rates for the years 1849-1851 as depicted in diagram 1 illustrates.

The success of the Badenese system was also due to the fact that the vaccination institutes and the public doctors were parts of a larger network which began to form in the 19th century and supplied the population with fundamental health services and provided the government with information on sanitary conditions and outbreaks of notifiable diseases. Besides the *Physicus*, each district disposed of licensed midwifes who had taken a centralized examination or even graduated from special schools that were linked to the medical departments of the universities in Heidelberg and Freiburg. Those who engaged in this work were however not only acquainted with technical skills. Since they were *elected* by the women of their district they also enjoyed the trust and support of the local population. And it was certainly no minor detail that the

[36] Corresponding bills from doctors and peasants' primes can be found in StAL, F177-II-387

[37] StAL, F177-II-387, letter of the Amtsarzt of Neeresheim, October 13th 1829

[38] For the period 1813-1861, these reports are available on an annual basis in: Ministry of Inner Affairs (ed.) Regierungsblatt für das Großherzogtum Baden, Karlsruhe various years.

[39] A very detailed contemporary account can be found in Victor Stoeber, Topographie et Histoire médicale de Strasbourg et du Département du Bas-Rhin, Paris 1864, pp.438-445

government urged midwifes to lobby for vaccination and required candidates at the schools in Heidelberg and Freiburg to study inoculation. Unlike Wurttemberg, where adherents of the anti-vaccination movement were no rarity even in the medical profession, Baden therefore successfully implemented a strategy which was just recently recommended by the WHO as a part of the organization's roadmap for the prevention of Ebola: Drawing on the expertise and enlisting the support of lay practitioners and traditional healers.[40]

When it came to the detection of outbreaks, the prevention networked reached sill lower: From midcentury onwards, mayors and local police officers in every Baden community disposed of standardized forms they only needed to fill in and submit to the Sanitary Commission to report presumable cases of smallpox. By doing so, they triggered a routine response which set in *simultaneously* at the local and the regional level: If the district doctor confirmed the diagnosis of the disease, his office conferred him the right to initiate emergency vaccinations at the state's expense and to impose revaccinations by police force if necessary. While vaccinations set in in the exposed community, the Sanitary Commission would contact the closest vaccination institute to ensure the continued supply of vaccine, informed the district doctors of adjacent communities and inquired whether they disposed of sufficient stocks of lymph to commence an own revaccination campaign if corresponding needs were voiced. The fascinating feature of this piece of evidence is the speed at which the entire process took place: Among the dozens of notification forms which survived in the Badenese Generallandesarchiv, most show that the Sanitary Commission responded on the same day or within a day's time to a potential outbreak and that vaccinations and containment measures were performed immediately. This swift flow of communication was evidently due to the fact that competencies were clearly defined and that local officials – medically trained or not – were informed about the course of action to be taken whenever new infections with smallpox occurred. The historical findings also shed light on the prospect of current epidemiological policies. In particular, they suggest that seemingly simple early warning systems like pre-written forms, check- and contact lists might give the health authorities a valuable head-start in the fight against epidemics. For, if a pile of printed notification forms and the effort of country-doctors who toured their districts on horsebacks sufficed to halt epidemics like the 1826 outbreak in Alsatia, cellular phones and online-communication might have an even larger impact if appropriately used.

A final note in this context relates to the health institutions' ability to cope with heterogeneous epidemic threats. There is little doubt that the Baden system represented an appropriate response to smallpox. This result would - however - be more intriguing if it was not confined to a specific disease but applied to other infections as well. In fact, there are reasons to assume this was the case. To start with, contemporary sources frequently refer to the country as an example of succesful disease prevention. Yet, there is more: In the post-war period of 1918-1920, Germany was hit by the global influenza pandemic of the "Spanish Flu". At the time, neither a vaccine nor antiviral medication was available and all doctors could do to oppose the disease was isolating patients and treating the symptoms. Unfortunately, statistical evidence on mortality and morbidity is patchy for the first wave of the pandemic. By contrast, corresponding data were reported by some - although not all - federal states during the second wave of the pandemic 1919/20. The exciting feature of this information lies with the fact that Baden again weathered

[40] See: WHO (ed), Ebola Response Roadmap, Published online, August 2014

the epidemic comparatively well although it was highly exposed and more densely populated than the German average. Thus, while mortality amounted to about 6 per 1000 nationwide, the corresponding figure for Baden stands at 4/1000. And even if anecdotal evidence is rare, it suggests that the district doctors played a substantial role in bringing this result about. For, it was typically the public doctors who pressed the administration into adopting quarantine measures like the temporal closure of schools. In some cases physicians' initiatives even succeeded in enforcing unpopular measures like the cancellation of festivals and the closure of cinemas and theaters.[41]

7. Conclusion:

In past and present alike, infectious diseases caused enormous harm both in terms of economic damage and human suffering. Notwithstanding the considerable costs of epidemics, prevention policies at the national - and in particular - at the international level are often deficient or poorly enforced. This is all the more intriguing as effective means of prophylaxis are well known and seldom expensive. The present paper has attempted to shed light on this puzzle by combining methods of mathematical epidemiology, theoretical economics and economic history.

In order to derive testable theoretical predictions, it has first introduced a compartmental model of disease spread. Next, this framework was integrated into an economic choice model to capture the relationship between the expected costs of an epidemic and the optimal response policies. The model predicts that national prevention efforts will always fall short of the internationally optimal level if countries are connected and policymakers do not coordinate their preventive activities. The reason is that prevention in one country is an unpaid factor in the "production" of immunity abroad. Regarding the total costs of an outbreak, the analysis showed that they are increasing in the spreading capacity of the disease and in the length of the infective period. Turning to the prevention externality, it was found that the national allocations will be more distorted the closer the connectivity between countries, the greater international variation in epidemiological conditions, the faster the spread of the epidemic or the longer the infective spell. Conversely, the externality shrinks if prevention efforts have little effect on the international level or if countries are sufficiently exposed to ensure comprehensive national health policies anyways.

In a next step, the theoretical results were tested based on the experiences of the 1870-74 smallpox pandemic which ravaged Europe after the Franco-Prussian war. The empirical analysis corroborated the theoretical findings as it found that German states were more likely to pass and enforce mandatory vaccination laws if they were highly exposed to the risk of imported epidemics. On the other hand, states which proved relatively lax in their prophylactic policies were typically bordered - and hence protected - by countries which had introduced mandatory immunization regulations before. Consequently, epidemiological free-riding was no mere academic scenario but a real danger which was not banned until the "Reichsimpfgesetz" made the vaccination of all infants mandatory throughout Germany in 1874.

[41] For a discussion of the pandemic in Baden: Witte, Wilfried, Erklärungsnotstand, die Grippeepidemie in Deutschland, Heidelberg 2004

Vaccination is certainly a cornerstone of prevention. Nevertheless, successful disease control depends at least as much on the implementation of immunization programs and on the rapid detection and containment of potential outbreaks. In consequence, the paper has attempted to analyze first, how preventive policies evolved under different socio-economic and geographical conditions and second, to identify the common characteristics of successful anti-epidemic strategies. Given the limitations of the surviving archival evidence, it has thereby focused on the well-documented case of the states of Baden in Wurttemberg in the 19th century.

The increased scrutiny provided support and several additions to the previous results. First, the case of Baden and Wurttemberg resembled the German situation in the sense that Baden was epidemiologically more exposed than its eastern neighbor. The country's policymakers reacted by creating institutions which allowed for a rapid and coordinated response in case of severe outbreaks. The common characteristics of these institutions were strict centralization, control of the medical personnel and a high degree of standardization at all stages of the vaccination system. When it came to translating these general principles into an actual prevention infrastructure, Baden's health authorities relied on a set of relatively simple but apparently effective measures. Thus, the country was among the first establish public vaccination institutes which were in charge of vaccinating the poor and maintaining stocks of high-quality vaccine. In spite of their modest budget, these institutions allowed for economies of scale and played a central role in facilitating rapid response – i.e. prompt mass-vaccination campaigns – in case of epidemic outbreaks. And even under less dramatic circumstances they were a key factor in maintaining the second pillar of the Badenese system's success: the acceptance and the support of a majority of the country's citizens. In contrast to other regions, where the use of contaminated or denatured vaccine had repeatedly spurred popular protest, the vaccination institutes allowed Baden's citizens to be relatively sure they received safe vaccination lymph. That inoculations were performed at the state of contemporary medical art – and that patients' pain costs were kept at acceptable levels – was also due to the fact that vaccination permissions were exclusively granted to public doctors. By setting minimum standards for the operation and obliging the public doctors to submit written reports about all vaccinations and follow-up visits they performed, the system reduced the risk of failed vaccinations and guaranteed that the reported inoculations reflected the true extent of immunization. In addition, the regular timing and standardized procedure of the public vaccination campaigns fostered a certain habit-formation and reduced resistance against the policy. More importantly though, the Sanitary Commission successfully addressed popular opposition – or ignorance – by enlisting the support of midwifes, teachers and pastors, of actors that is, who disposed of intimate local knowledge and particular trust in their resident communities. Finally, Baden's strive for detection and containment encouraged constant communication between the Sanitary Commission, medical practitioners, local officials and the general public. Being as much a source of conflict as of epidemiological information, this effort proved crucial because it initiated institutional learning within the public health sector. For, in the course of time, the health-authorities were able to draw on this knowledge and established efficient response mechanisms like checklists and standardized forms for doctors and local officials which simplified documentation and facilitated immediate and coordinated action if necessary.

In contrast, Wurttemberg's geographical and socio-economic conditions failed to put the country's decision-makers under the pressure to see for a similar degree of institutionalization. The central feature of this result lies not only with the fact that the possibility to free-ride on their neighbors' prevention effort eventually turned the Wurttemberg authorities into the fools of their own fortune. What is more important, historical path-dependency apparently trapped the country in a state of institutional inertia. For although more and more contemporary experts – even members of Wurttemberg's own health department – observed that Baden benefited considerably from its centralized system, the Wurttemberg authorities proved incapable of reforming the decentralized and uncoordinated vaccination system before the "Reichsimpfgesetz" of 1874 forced them to do so.

In sum, the epidemics of the past have a simple lesson to teach: Coordination pays off. Yet, since public health institutions inevitably reflect a complex set of biological, geographical and political conditions, even countries which face the same diseases may embark on totally different policy trajectories. Moreover, as a result of free-riding and historical path-dependency, institutions which facilitate cooperation at the international level are seldom created unless the horrors of a huge outbreak like the 1870 pandemic pave the way for concerted political action. With all due caution, history therefore suggests that there might be a lot to gain on a global scale if the heighted interest in current health threats like Ebola or pandemic influenza was channeled into initiatives for multilateral health institutions, binding commitments to certain immunization and documentation targets and constant and intensive cooperation in the detection and containment of outbreaks. This is particularly true as highly institutionalized health sectors like that of 19th century Baden seem to benefit from their coordination and response mechanisms when addressing both known threats like smallpox *and* newly emerging diseases like the "Spanish Flu". Given that rapid socio-economic development and environmental change raise epidemic spreading capacities world-wide, the capacity to deal with heterogeneous pathogens will be pivotal for future prevention success. The doctors of the 19th century would therefore certainly agree with their literary confrere that

"The plague never dies or disappears for good; that it can lie dormant for years and years in furniture and linen-chests; that it bides its time in bedrooms, cellars, trunks, and bookshelves; and that perhaps the day would come when, for the bane and the enlightening of men, it would rouse up its rats again and send them forth to die in a happy city.[42]"(Albert Camus, The Plague)

But they would also tell us that the plagues of the future – might they disguise as Ebola, influenza or some hitherto unknown pathogen – can be prevented by institutions which record the warning signs and trigger early action no matter which city in the world the rats chose to infest.

[42] Albert Camus (transl. by Stuart Gilbert), The Plague, London et al., 1980, p 150

References

Archival Resources:

1. Generallandesarchiv Karlsruhe (GLA Karlsruhe) 236-16019-16045

2. Hauptstaatsarchiv Stuttgart (HStASt), E-31

3. Staatsarchiv Ludwigsburg (StAL), E-161 to F-184

4. Archives Départementales du Bas-Rhin 5M67

5. Archives Départementales du Haut-Rhin 5M68

Historical Periodicals and Journals:

Annuaire Statistique du Département du Bas-Rhin

Annuaire Statistique du Département du Haut-Rhin

Hufelands Journal der practischen Heilkunde

Mitteilungen aus dem Reichsgesundheitsamt, various issues, 1881-1884

Rapport sur les vaccinations pratiquées en France pendant l'année

Statistik der Inneren Verwaltung des Großherzogthums Baden

Statistisches Jahrbuch für das Land Baden

Statistisches Jahrbuch für das Königreich Württemberg

Zeitschrift für Staatsarzneikunde

Literature:

Erwin Heinz **Ackerknecht**, Geschichte der Medizin, Stuttgart 1992

Leonid V. **Azarnert,** Child mortality, fertility, and human capital accumulation, in: Journal of Population Economics, Vol.19, No.2 (2006), pp. 285–297

Peter **Baldwin**, Contagion and the state in Europe, 1830-1930, Cambridge Massachusetts 1999

N.T.J., **Bailey**, The Biomathematics of Malaria., London (1982)

Philipp Carl **Baur** von Eiseneck, Sammlung sämtlicher Gesetze, Verordnungen, Instructionen, Belehrungen und Entscheidungen, welche in dem Gross- Herzogthume Baden über Gegenstände der Gesundheitspolizey erscheinen sind, Karlsruhe 1830

Sekhar **Bonu**, Manju **Rani**, Timothy D **Baker**, The impact of the national polio immunization campaign on levels and equity in immunization coverage: evidence from rural North India, Social Science & Medicine, Volume 57, Issue 10, November 2003, pp. 1807-1819

A. **Bousquet**, Traité de la Vaccine, Paris 1833

Fred **Brauer** et al., Mathematical Epidemiology, Heidelberg 2008

Dagobert L. **Brito**, Eytan **Sheshinski**, Michael D. **Intriligator**, Externalities and compulsory vaccinations, Journal of Public Economics, Vol 45, Issue 1, June 1991, pp. 69-90

Heinrich **Cless**, Impfung und Pocken in Württemberg, Stuttgart 1871

Pierre **Darmon**, La longue traque de la variole: Les pionniers de la medecine preventive, Paris 1986

Barbara **Dettke**, Die asiatische Hydra, Cholera 1830/31, Berlin 1995

Otto **Diekmann**, Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation, New York, 2000

R. **Feilden**, Immunization and health reform: making reforms work for immunization, Geneva 1998

F.E. **Foderé**, Mémoire sur la petite vérole vraie et fausse et sur la vaccine pour l'instruction des parens et des jeunes médecins, Strasbourg 1826

Varun **Gauri**,Peyvand **Khaleghian**, Immunization in Developing Countries: Its Political and Organizational Determinants, Geneva 2002

Mark **Gersovitz**, Infectious diseases, public policy, and the marriage of economics and epidemiology, World Bank Research Observer Volume 1, No.2, (2003) pp. 129-157

Mark **Gersovitz**, Jeffrey S. Hammer, The Economical Control of Infectious Diseases, The Economic Journal, Volume 114, Issue 492, January 2004, pp. 1–27

Ute **Frevert**, Krankheit als politisches Problem 1770 – 1880, Göttingen 1984

Edward **Jenner**, An inquiry into the causes and effects of the variola vaccine, a disease discovered in some of the western counties of England, particularly Gloucestershire, and known by the name of the cow pox, London 1798.

Robert **Jütte**, Geschichte der Alternativen Medizin

W. O. **Kermack** and A. G. **McKendrick**,Contributions to the Mathematical Theory of Epidemics. I-III, Procedures of the Royal Society London; vol's 115, 138, 141, London 1927; 1932)

M. **Lächlin**, Geschichte der allgemeinen Einführung der Kühpockenimpfung in Botnang in den Jahren 1801, 1802 und 1803, unpublished manuscript 1804, reprint in: G. Cless, Impfung und Pocken in Württemberg, Stuttgart 1871

Friedrich **Kluge**, Adolf Kussmaul, Freiburg 2002

S.L., **Kotar**, Smallpox: A History, Jefferson (N.C.) 2013

Paul **Kübler**, Geschichte der Pocken und der Impfung, Berlin 1901

Adolf **Kussmaul**, Jugenderinnerungen eines alten Arztes, Reprint München 1960

Francisca **Loetz**, Vom Kranken zum Patienten, Stuttgart 1993

A. **Maler**, Geschichte der Kuh-Pocken-Impfungen, Karlsruhe 1804

Elard Hugo **Meyer**, Badisches Volksleben im 19. Jahrhundert, Freiburg 1900

Kenrad E. **Nelson**, Carolyn Masters Williams, Infectious disease epidemiology:theory and practice, Sudbury, 2007

Md. Abdul **Quaiyum**, Jahanara Khatun, Impact of National Immunization Days on Poliorelated Knowledge and Practice of Urban Women in Bangladesh, Health Policy Planning, Volume 12, No. 4, 1997, pp. 363-371

Albert **Renscher** (ed.), Vollständige, historisch und kritisch bearbeitete Sammlung der württembergischen Gesetze, Stuttgart 1830

Steven **Riley** et al., Transmission Dynamics of the Etiological Agent of SARS in Hong Kong: Impact of Public Health Interventions, in: Science, vol. 300, no. 5627 (June 2003), pp. 19611966

James C. **Robinson**, Theory and Practice in the Design of Physician Payment Incentives, in: The Milbank Quarterly, Vol. 79 No 2 (2001), pp. 149-177

Peter **Sköld**, The two faces of smallpox: a disease and its prevention in eighteenth-and nineteenth-century Sweden, Umea 1996

Victor **Stoeber**, Topographie et Histoire médicale de Strasbourg et du Département du BasRhin, Paris 1864

Christian August **Struve**, Anleitung zur Kenntnis und Impfung der Kuhpocken, Breslau 1802

WHO, Smallpox and its eradication, Geneva, 1988

WHO State of the Art of New Vaccines, Geneva 2006

Wilfried, **Witte** Erklärungsnotstand, die Grippeepidemie in Deutschland, Heidelberg 2004

Eberhard **Wolff**, Einschneidende Massnahmen: Pockenschutzimpfung und traditionale Gesellschaft im Württemberg des frühen 19. Jahrhunderts, Stuttgart 1998

Darwin beats Malthus: Medicalization, Evolutionary Anthropology and the Demographic Transition

Abstract

For the better part of human history, life was most fragile and death most imminent during infancy and early childhood. The death of a child may be hardly bearable from a humanitarian perspective. Yet, certain currents in economic theory attach a silver lining to high mortality by claiming that the Malthusian check on population raises per capita income and facilitates the accumulation of capital. The present paper challenges this conventional wisdom. In essence, it argues that high levels of environmental risk produce genetic and behavioral adaptations which induce individuals to have many - in terms of parental investment - cheap offspring. Conversely, stable environments recast the tradeoff between child quantity and quality in favor of more quality-based reproductive strategies. Incorporating these biological relationships into the traditional Barro-Becker model of fertility, the paper finds that both declining extrinsic mortality and increased effectiveness of parenting effort potentially trigger a demographic transition. Thus, the economic benefits of Malthusian population checks are mitigated because high mortality endogenously produces high fertility whereas improved survival encourages human capital investment and fosters long-term growth. To assess whether the theoretical predictions conform with historical reality, I use smallpox vaccination in 19th century Germany as a natural experiment. Performing an econometric analysis of 67 districts in the Granduchy of Baden provides evidence, that comprehensive immunization and advanced medicalization came along with reduced mortality, significantly lower fertility and increased parental care. In sum, it therefore seems that Malthusian mechanisms are at least partly offset by countervailing biological adaptations.

1. Introduction

Death is old. A skeletal figure, wrapped in black, carrying a huge scythe. Be it on gravestones or paintings, the image is ubiquitous – and deeply flawed. For in fact, throughout centuries of human history, the face of death was the face of an infant: As late as the turn of the 20th century, 3 out of 10 European children died before school age. And although progress has been made, WHO statistics continue to report sobering 16000 infant deaths per day worldwide.[1]

The enormous amount of human suffering behind these figures is undeniable. Still, a substantial body of macroeconomic literature attaches a silver lining to the high mortality rates in underdeveloped nations. Following Malthusian arguments, research like Clark (2007) stresses the fact that mortality acts as a buffer to reduce population growth and raises per capita income. In the same vain, Young (2005), Lagerlöf (2003) and Brainerd and Siegler (2003) postulate a benign effect of epidemics by arguing that they foster human capital acquisition or raise wages by increasing the scarcity of labor. Voigtländer and Voth (2005) went as far as dubbing death through infectious diseases a "horseman of the riches" because high mortality led to higher income which – by Engel's Law – created a positive feedback effect on the demand of manufactures.

A direct implication of these results – albeit rarely spelled out *expressis verbis* – is a strong skepticism regarding the economic benefits of medical and social progress. Neo-Malthusians would therefore probably agree with Azarnert (2006) who claims that "a premature public health intervention to reduce early child mortality, although in harmony with humanitarian approach in the short run, generates a mechanism that works against children's survival chances in the long run." (Azarnert (2006), p. 294)

The present paper challenges this received wisdom. Based on recent advances in evolutionary biology, it argues that high mortality triggers physiological and psychological adaptations which foil the prospects for sustained economic growth. The dimensions and potential implications of these behavioral responses are largely ignored by both (Neo-)Malthusian theories and the canonical "altruism model" proposed by Barro and Becker (1986): In the world of Malthus, individuals follow the irresistible "passion between the sexes" without adapting to their environment (Malthus (1798), ch. 2). As for the Barro/Becker model and its numerous extensions, they typically assume the relative costs of reproducing, rearing and educating offspring unaffected by changes in the environment.[2] In other words, high expected mortality does not prevent parents from attaching closely to their children and low mortality does not foster increased investment in child "quality". Offspring quantity and quality being treated as normal goods, these models predict that improved survival reduces the price of children which in turn will translate into population growth and declining living standards unless productivity and wages increase accordingly.

The life sciences reach different conclusions: In the world of Darwin, fertility is not the outcome of uncontrolled emotions, but a function of ecological conditions. Parents who seek to

[1] UNICEF, WHO, World Bank UN DESA/Population Division (2014), Data Appendix

[2] The model reaches the empirically unfounded prediction that a decline in mortality reduces the expected costs of surviving children which in turn raises the demand for offspring. As a result, improved survival will translate into population growth.

maximize their *fitness* (the combination of life expectancy, socio-economic rank and reproductive success) are expected to counter high mortality by "hedging their bet" and adopting a "high fertility-low investment" strategy. The "Malthusian gift" of positive checks on population growth therefore turns out a cold comfort because the evolved behavioral response to high mortality leads to high fertility, sluggish human capital accumulation and socio-economic stagnation. Conversely, improved survival, medical progress and the dissemination of medical knowledge carry a "Darwinian gift" of low fertility, higher investment in child quality and rising living standards. Hence, rather than fearing a Malthusian trap, biologists point to falling mortality as a necessary precondition for the fertility transition and the acceleration of human capital accumulation.

To assess if Malthus was wrong where Darwin was right, the present paper ventures into various fields outside the focus of traditional economics including evolutionary anthropology, ecology, and psychology. Concentrating on the interface between economics and the life sciences, it contributes to different strands of research. Regarding theory, it is the first to incorporate the principles of evolutionary *life history theory* into an economic optimization model. Doing so, it ties up some loose ends in the economic literature on the fertility decline. In particular, it provides a parsimonious explanation for one of the central forces in endogenous growth models like Galor and Weil (2000), Galor and Moav (2002), Kalemli-Ozcan (2002) or Soares (2005): the emergence of preferences which favor a small family size and heavy investment in human capital. At the same time, accounting explicitly for ecological conditions sheds new light on empirical findings which would otherwise contradict the classical theory of the demographic transition. For, standard accounts of the process typically require either a mortality decline before the fertility decline or a strong increase in per capita income. Unfortunately, empirical support for these claims is mixed. In fact, most studies of completed demographic transitions detect a negative nexus between fertility and income but fail to establish an unambiguous timing of the rise of per-capita income and the onset of the fertility decline.[3] Regarding the hypothesis of incremental habit formation, historical records include a number of cases in which the fertility decline paralleled or - rarely - preceded the fall in mortality.[4] Just like the emergence of preferences for increased offspring quality, these initially puzzling observations can – however – be explained as evolutionary adaptations to changed environmental conditions. How these adaptations form in the course of natural selection and how they contribute to our understanding of the demographic transition will be the central questions in the following sections: Section 2 introduces the paper's biological background, namely life history theory (LHT) and the Darwinian mechanism of r/K-selection. Section 3 analyzes how the traditional economic models of fertility are affected if evolutionary fitness maximization is incorporated in the household's objective function. Seeing that LHT has not been merged with economic theory so far, section 4 discusses the empirical plausibility of the "Darwinian" fertility model proposed in section 3. In order to gauge the potential size of adaptations to variation in selective pressure, section 5provides a historical case study of smallpox vaccination in 19th century Germany. The case is particularly valuable because vaccination caused a sizable decline in

[3] For a survey see Hotz, Klerman, and Willis (1993). More recent contributions include: Tertilt (2008) and Shoven(2011)

[4] For an overview see Caldwell (2006)

extrinsic mortality[5] and improved parents' ability to influence the survival chances of their children to an unprecedented extent. Having analyzed the historical impact of medical progress on the demographic transition, section 5 discusses potential policy implications and concludes.

2. Understanding the Economics of Nature: Life History Theory

Grasping the fundamentals of life history theory (LHT) is easiest on those days of spring, when nature turns into a huge and bustling bazaar: Birds twittering in the trees under your window, insects buzzing in the air, the neighbors' cat displaying and caterwauling on the roof at four o'clock in the morning – in a sense they all sell or purchase genes in order to pass them down to future generations. Life history theory (LHT) seeks to understand the economic logic of this exchange. Based on the insight that all organisms face binding genetic and energetic constraints,[6] LHT considers reproductive success or early death as the results of a series of optimization problems. Nature has solved these problems in a variety of ways: Some species like Atlantic Salmons produce thousands of offspring in one suicidal effort and die immediately afterwards. Others like rabbits compensate the risk of being a simple a prey by fast maturation and notorious fecundity. Finally, suburban humans invest in violin, chess and Chinese classes in order to make sure, their adored little cherub succeeds in life.

The above *life history strategies* can all be defined as optimal combinations of so-called *life history traits*. A life history trait is a numerical indicator (e.g. size at birth, age at first reproduction) which corresponds to a specific tradeoff between the conflicting energetic needs of survival, growth and reproduction. Following the convention of evolutionary biology, both single life history traits and entire *life history strategies*, are ordered along a "slow-fast" continuum. For example, males may solve the tradeoff between mating and parenting effort by marrying one woman and have few offspring (slow) or by trying their luck at serial seductions (fast). Conversely, females can continue the quest for their "Mr. Right" (slow) or be happy with the "Mr Not So Right" they have found and begin reproduction (fast). Finally, parents may choose to invest in few, highly competitive offspring (slow) or to have many children who might – however – fail to survive or reproduce.

Which of these strategies maximizes the fitness of an individual – or an entire species for that matter – is a function of ecological characteristics like food sources, predation, disease stress or climate. Since natural selection favors optimal adaptations to these characteristics, life history traits are subject to *evolutionary plasticity*. This implies that reproductive choice and nurturing behaviors – hence, the biological underpinnings of the demographic transition we are interested in – are predominantly shaped by the selective pressure of infant and childhood mortality. Biologists typically formalize the relationship between mortality and optimal fertility by the concepts of r- and K-selection.[7] To highlight the logic behind these prototypical LHT-strategies, consider the following stylized example: The population consists of only two competing dynasties who face an *extrinsic* mortality rate of μ which is unaffected by parental

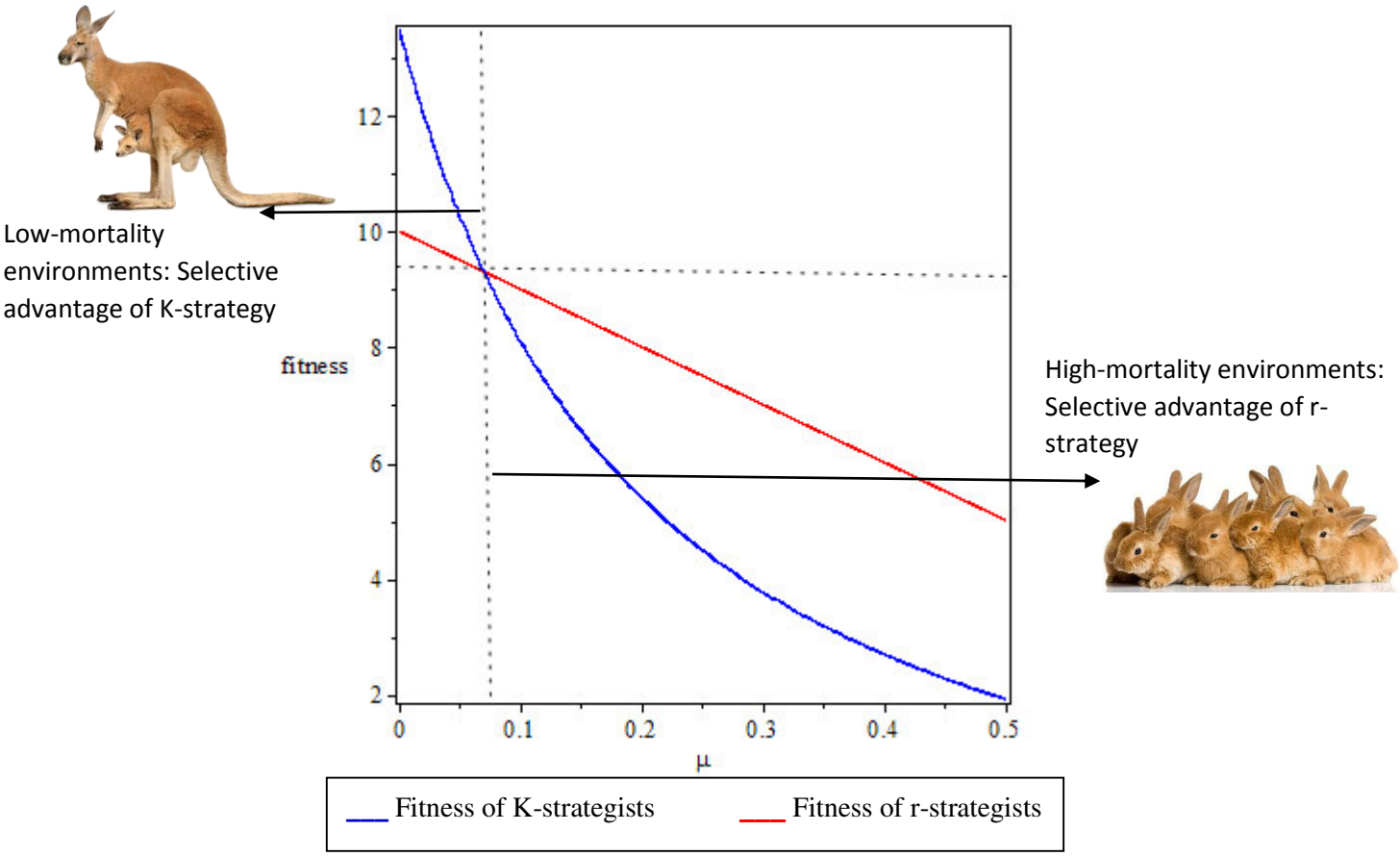
[5] The term "extrinsic mortality" is used by biologists to refer to mortality risks which are beyond human control.
[6] This precludes the emergence of *Darwinian Demons*, that is, of organisms which reproduce immediately after birth, have infinitely many offspring and live forever.
[7] For an overview see: Rockwood (2015)

care. Call the first lineage the "rabbits" or the r-lineage. Rabbit parents seek to maximize their fitness by producing as many descendants as biologically feasible but refrain from investing in their quality. By contrast, the second lineage, the "kangaroos" or K-lineage, reduces fecundity, and provides their few children with a "womb" of intensive care and education.[8] Which of the two dynasties will prove "fitter" in terms of passing down their genes to future generations depends on the environment. The dangers of highly unstable and unpredictable, r-selecting ecosystems are best countered by producing many (in terms of parental investment) cheap offspring. For even if many of them fall victim to predators, diseases, natural disaster or violence, there will always be some left to carry on the lineage. On the other hand, stable environments favor K-selective strategies because improved survival increases competition for resources and suitable mates. What is more, the marginal return to quality investment rises in the offspring's expected life span because quality investment – be it practical skills, academic attainment, reputation or socially preferable characteristics – is incorporated in individual children and cannot be transferred if the carrier of these traits dies.[9] Parents who seek to increase the future competitiveness of their descendants will therefore only collect the dividend of their investment if a sufficient number of offspring reaches adulthood and succeeds in the strive for mates and resources. These relationships are illustrated in Figure 1. The panels show the expected number of reproductively successful offspring under the "rabbit" and the "kangaroo" strategy. While μ is exogenous, total fitness – quantified by the reproductive success of future generations – depends on parenting effort and varies across the two dynasties. Thus, as mortality decreases, the fitness of the r-lineage increases proportionally because the expected number of survivors rises linearly as μ falls. By contrast, the fitness of K-strategists grows more than proportionally because in addition to the linear survival effect, their offspring also benefit from the rising competitiveness effect. As shown in figure 1, r-selective behavior provides a selective advantage in high mortality environments where the expected return to quality investments is low. Yet, as survival chances improve, this outcome is reversed such that fitness is maximized by parents who chose a K-selective strategy.

[8] In fact, the scholarly terms of r- and K- selection stem from r=rate and K=carrying capacity. Yet, thinking of rabbits and kangaroos is definitely more intuitive.

[9] In the animal kingdom, the ramifications of this fact are most palpable in the case of large mammals: Due to their long gestation and lactation periods, females of these species typically produce only one offspring per season. Their reproductive fitness therefore depends on the ability to survive multiple periods.

Figure 1: r-selection and K-selection under different environmental conditions



Before leaving biology and turning to economic theories of fertility, let us briefly summarize the gist of the previous section: To begin with, life history theory proposes that all living organisms solve economic optimization problems when allocating scarce energy resources. Moreover, whether fast (r-selective) or slow (K-selective) strategies are ultimately more successful depends on environmental characteristics, in particular on the prevailing mortality regime. In consequence, to be in line with evolutionary biology, an economic model of human reproduction must account for the fact that the relative rewards in the tradeoff between child quantity and quality change as mortality rises or falls. Keeping these results in mind, let us assess in more detail how the inclusion of LHT fits into the received economic theory of fertility.

3. From Biology to Economics – A Darwinian Fertility Model

It is not difficult to integrate adaptation to environmental risk into the standard economic theory of fertility. Yet, this simple modification leads to remarkable alterations in the predicted responses to declining mortality. In particular, it will turn out that evolutionary plasticity creates an endogenous feed-back mechanism which curtails Malthusian crises if the improvement in survival rates is substantial *and* sufficiently obvious to impact the households' family planning.

To support this claim, we rely on the classic Barro-Becker model, and tentatively integrate evolutionary fitness in the objective function.

As in the original Barro-Becker set-up, consider the problem of representative parents whose utility depends on the consumption $u(c)$ of energetic resources (including investment in own human capital) and on the value assigned to their offspring. The emotional bond between parents and children enters the model through the "altruism" parameter β which links parents' utility to the value of their children $f(N,q)$.[10] Of course, offspring "value" is determined by a variety of factors including survival, health, kinship support, mutual affection, wealth, success or social recognition. Still, irrespective of cultural particularities, these factors are nothing if not the socio-economic concomitants of evolutionary fitness. Thus, we can interpret the illusive "value" of children as their fitness which depends on parental quality investment q.

For mathematical convenience, utility from consumption $u(c)$ is assumed to exhibit the conventional functional characteristics, namely being twice continuously differentiable and exhibiting positive but decreasing marginal returns in *c*. The analytical treatment of fitness merits closer attention because it constitutes the major point of departure from the orthodox Barro-Becker framework. Following standard evolutionary models like Smith and Fretwell (1974), the fitness of each child is a sigmoid function of parenting effort.[11] At very low levels, the impact of quality investment is therefore modest but rapidly increasing. As parents increase their efforts, the corresponding marginal returns become smaller and eventually converge to zero as q approaches infinity. In the interest of analytical tractability, we will abstract from the possibility of birth order effects, sibling competition and correlated mortality risks. With this simplifications, the total fitness of the lineage is given by the expected number of survivors from N births multiplied by individual fitness, i.e.

$$f(N,\mu,q) = (1-\mu)\cdot Nf(\mu,q),$$

where μ denotes the exogenous part of mortality. Given that human infants require a minimum of care to survive at all, $f(N,\mu,0) = 0$ by construction. Furthermore, to capture the different evolutionary pressures of r- and K- selective environments, the fitness function meets the following criteria: First, the sensitivity of f with regard to q indicates the effectiveness of parental solicitude. Consequently, the elasticity of the fitness function ε_q can be considered a formal measure of parents' ability to control the fitness of their offspring. Next, mortality μ influences fitness via two different channels. To begin with, for given levels of q, higher mortality limits survival and therefore reduces $f(N,q)$ in *absolute terms*. In addition, high childhood mortality and morbidity prevent the accumulation of somatic capital (e.g. physical strength, beauty, academic and practical skills) and mitigate the expected impact of each unit of parental investment. Under high mortality regimes, K-selective strategies therefore loose their competitive advantage against r-strategies because a one-percent increase in q leads to smaller gains in offspring fitness than under more favorable conditions. We incorporate this

[10] For the complete model and the underlying assumptions see: Barro and Becker (1988), pp. 1-9
[11] This functional form was first introduced and discussed in: Smith and Fretwell (1974), pp. 499-503

selection effect into the model by requiring ε_q to be smaller the higher extrinsic mortality μ, i.e. $\frac{\partial \varepsilon_q}{\partial \mu} < 0$.

Translated into mathematical terms, the overall problem of a representative couple can now be expressed by the following objective function:

$$\max_{c,N,q} U(c,N,q) = \alpha ln(c) + \beta(1-\mu) \cdot Nf(\mu,q)$$

As parents attempt to choose their optimal reproductive strategy, the set of feasible solutions is constrained by the fact that rearing and educating children is a time consuming task and competes with income earning activities. Under the additional assumptions that (i) wages are exogenous, that (ii) all adults are endowed with one unit of time and (iii) income cannot be saved, parents' budget constraint takes the below form:

$$c = w(1 - Npq),$$

where w and p respectively denote the wage rate w and the "price" of parenting effort p. Using this expression, c can be eliminated from the objective function, such that the problem depends only on the two choice variables q and N. Normalizing wages to 1 and taking the derivatives of the modified function U delivers the following optimality conditions:

$$(i) \quad \frac{\alpha Np}{1 - Npq} = \beta(1-\mu) \cdot Nf'(\mu,q)$$

$$(ii) \quad \frac{\alpha pq}{1 - Npq} = \beta(1-\mu) \cdot f(\mu,q)$$

We can eliminate N from the system by dividing (ii) and (i) to obtain the following relationship:

$$(iii) \quad q^* = \left(\frac{f(\mu,q)}{f'(\mu,q)}\right) \leftrightarrow \varepsilon_{q*} = 1$$

The equation states the intuitive fact that at the optimum, fertility and parenting effort are set to the levels where the opportunity costs of having *one* more child or of providing *all* existing children with one additional unit of quality investment are exactly equal. More importantly though, the expression depends only on q, μ and exogenous parameters. This being the case, the optimal adjustment of quality investment in response to changes in mortality can be derived from the total differential of (iii)

$$\left(\frac{(f_{qq} \cdot q + f_q) \cdot f - f_q q \cdot f_q}{(f)^2}\right) dq + \left(\frac{q \cdot \partial \varepsilon_q}{\partial \mu}\right) \cdot d\mu = 0 \text{ (iv)}$$

$$\text{Using that } q^* = \frac{f}{f_q} \leftrightarrow \frac{dq}{d\mu} = -\frac{\frac{q \cdot \partial \varepsilon_q}{\partial \mu}}{f_{qq} \cdot q} < 0^{12}$$

[12] Note that the sigmoid form of the fitness function allows for f'' > 0 at low levels of q. Yet, choosing such a small q violates the optimality condition $\varepsilon_{q*} = 1$. For since f(N,0) = 0, qf'(q) is strictly larger than f(q) for all q on the convex segment of f.

The interpretation of this expression is straightforward: As in the qualitative case of the "rabbits" and the "kangaroos", increased extrinsic mortality reduces the optimal investment in child quality. Yet, if this is the "quality" side of the parental tradeoff, what about "quantity"? To answer this question, we solve (ii) for N and compute the total derivative of the expression:

$$N = \frac{1}{pq} - \frac{\alpha}{\beta(1-\mu)f(\mu,q)}$$

$$dN = \left(-\frac{1}{pq^2} - \frac{\alpha f_q}{\beta(1-\mu)f^2}\right) dq - \frac{\alpha}{\beta}\left(\frac{f-(1-\mu)f_\mu}{(1-\mu)^2 f^2}\right) d\mu$$

Relying on the optimality condition $\varepsilon_{q*} = 1$, the relationship can be restated as follows:

$$\frac{dN}{d\mu} = \left(-\frac{N}{q}\right)\frac{dq}{d\mu} - \frac{\alpha}{\beta}\left(\frac{f-(1-\mu)f_\mu}{(1-\mu)^2 f^2}\right) \text{ (v)}$$

The first product is strictly positive[13] and captures the fact that parents respond to harsh environments by substituting child quality by quantity. The second one is – however – negative because increased mortality reduces offspring fitness for any initial level of q^*. As a result, the tradeoff between reproduction and consumption is recast and adults face incentives to increase their own consumption while reducing both q *and* N. Contrary to quality investment, the total effect on optimal family size is therefore ambiguous and depends on the parameters of the fitness and utility functions.

Inspecting (v) is nevertheless instructing because we observe that high mortality environments produce large N but in two exceptional cases: In the first, the weight of consumption α is extremely large compared to the altruism or "child-preference" parameter β. In the second, mortality forces fitness to such low levels that parents balance the opportunity costs of reducing family size by switching from reproduction to increased consumption (i.e. f is small and f_μ large in absolute terms).

Why is it justified to consider the above scenarios "exceptional"? The answer is, because both biological and socio-economic arguments cast doubt on their applicability to real populations – in particular to those in a demographic transition. For one, demographic studies of ancestral populations have shown that very few successful reproductive events suffice to raise human fitness above average (Jones JH, Bird RB. (2014), pp. 65-71.). Unless the probability of survival virtually drops to zero, the biological opportunity costs of investing neither in child quality nor quantity will therefore be too high to be offset by higher parental consumption. Adding to that, it is frequently observed that parents in traditional societies compensate the lack of modern financial markets and social security by relying on children as a form of old-age insurance. Consequently, raising a certain minimum number of offspring to adulthood seems to be a socio-economic necessity in pre-demographic transition economies which in turn ensures that the propensity to consume (α) is not too large relative to the propensity to reproduce (β).

[13] Recall that $\frac{dq}{d\mu} < 0$!

Seeing this evidence, even critical spirits should agree that under realistic assumptions, high levels of μ reduce the relative price of fertility vs. child quality and translate into a shift towards higher reproduction and lower parenting effort. Conversely, improved survival boosts the benefits of offspring quality and serves to suppress the optimal level of fertility.

The central message of these results is that biology buffers Malthusian dynamics in that predictable, low-mortality environments produce an endogenous fertility decline whereas under inimical circumstances, this pattern is reversed and both birth and death rates remain on high levels. Even more intriguing for theorists, the model also predicts that a preceding fall in *absolute* mortality is no necessary condition for the onset of the fertility decline. Rather, parents will start to increase quality at the expense of offspring quantity as soon as the perceived effectiveness of parental effort rises. In this case, the elasticity of the fitness function with regard to q (ε_{q*}) will be larger and the optimal investment in offspring quality must rise in order to fulfill the fitness maximizing condition $\varepsilon_{q*} = 1$. Well in line with a number of historical studies, model therefore concludes that reductions in family size can accompany or precede the mortality decline provided that parents obtain better means to influence their children's fitness. Besides for its fit with historical evidence, this scenario is also relevant to current development policies because many transition societies experience improvements in medical care and public education as byproducts of social and economic progress. Finally, the suggested Darwinian fertility model provides theoretical support to sewer the empirically questionable link between rising per capita income and the demographic transition. For, since the optimal level of parenting effort – and hence the resources devoted to each child – are determined by environmental factors, higher real wages are neither a necessary nor a sufficient condition for incentives to increase offspring competitiveness and limit family size.

Closing the theoretical analysis, we note that the combination of Darwinian principles with economic modelling provides a remarkably simple explanation for otherwise puzzling demographic patterns. It is accordingly tempting to incorporate fitness maximization into a general equilibrium model in order to assess its role in human capital accumulation and endogenous economic growth. However, this analysis would be a mere academic exercise unless the evolutionary disposition to adopt faster or slower life history traits was evident in observed human fertility strategies. Let us therefore defray the calibration of a full-fledged endogenous growth model to future research and contend with the discussion of two fundamental preliminary questions: First, notwithstanding its theoretical appeal, is the incorporation of evolutionary fitness into parents' objective function equally plausible from an empirical point of view? Or, more concretely, are there biological processes which cause *actual* humans – as opposed to the hypothetical *homines oeconomici* of the model – to behave like fitness maximizers? Moreover, if individual humans turn out to maximize their fitness, what about entire societies? And finally, in case evolutionary algorithms permeate the social sphere, under which conditions do these adaptations trigger or impede the onset of a demographic transition in real societies?

4. Linking the Model to Reality – Are Humans Fitness-Maximizers?

4.1. Physiological Adaptations

The ability to switch from slower to faster life history strategies belongs to our evolutionary heritage. In fact, the mechanism is rooted so deeply that it operates partly through the endocrine – hormonal – system without surfacing to the level of conscious decision. In a sense, humans are *coded* to maximize fitness – be it despite themselves in some cases. This is the case because the human foetus reacts to environmental signals from the first months of gestation onwards. If environmental conditions (e.g. nutrition) are stable and favorable, the endocrine system "judges" the prospects for growth and successful reproduction good and facilitates rapid growth and early maturation.[14] Conversely, extreme levels of disease prevalence, physical stress and malnutrition indicate inimical future conditions and therefore impair maturation.[15] Finally, in cases between these prototypes, life history traits exhibit great scope for evolutionary plasticity and follow Darwinian optimization algorithms. For example, young organisms respond to environments where episodes of stability and affluence alternate with insecurity and famine by seizing the few - probably short - spells of favorable conditions and precipitate maturation to facilitate fast and frequent reproduction.(Worthman (1999) pp. 135-165) Unlike the permanently elevated levels of stress caused by exceedingly harsh environments, the moderate and pulsating hormone responses to unpredictable environments do not inhibit maturation, but affect the endocrine system's hypothalamic-pituitary-gonadal and hypothalamic-pituitary-adrenal axes. These two hormonal circuits in turn co-determine the development of stress-and steroid-responsive neural processes which are associated with reproductive timing.(Cameron (2008), pp. 795-801). When it comes to the feasibility of fast or low LHT strategies, evolutionary forces therefore apparently to nudge humans towards fitness-maximizing behavior regardless of individual preferences or socio-cultural conditions.

If this was the only evidence to support the integration of Darwinian fitness into the economic theory of fertility, critical readers would rush to ask: What is all the fuss about minor biological curiosities? And indeed, their critique would not go entirely astray. For one, physiological processes like the onset of puberty or the length of the reproductive period are ultimately subject to genetic constraints. Their demographic effect is therefore strictly limited. Moreover, these adaptations merely determine an individual's *ability* to reproduce. They say nothing about parents' *willingness* to have additional children or to invest heavier in existing ones. Thus, in order to be consistent with the idea that humans choose their reproductive strategy (more or less) consciously to maximize their *utility*, there ought to be a link between the gains in fitness and the gains in perceived utility. Put differently although we have modelled humans as fitness maximizers, few people will regard themselves as such. They are more likely to strive for "happiness", "love and mutual affection", "harmony with traditions and social norms" and many more. Given that physiological processes have just been showed to determine only a small fraction of the overall variation in human life history strategies other evidence is therefore needed to make fitness maximization a viable approximation of actual behavior. In particular,

[14] In actual population data, this reaction corresponds to the empirical observation of a sustained secular decline in the age of menarche in virtually all developed countries. See: Wood (2009), p. 417

[15] Corresponding evidence is discussed by Ellis (2004), pp. 920-958; Chisholm,and Quinlan (2005), pp. 233-265

high environmental risk should not only reduce the marginal impact of quality investment in the fitness function. It should also lower the expected utility gains – e.g. the emotional rewards – parents associate with caring for their children. Addressing these matters involves two empirical questions, namely first, whether parents adjust their *post-natal* investment – as opposed to the physiologically determined pre-natal ability to reproduce – according to the model's predictions and second, whether the underlying psychological processes really respond to mortality risks or to some other unobserved variable.

4.2. Psychological Adaptations:

Being used to think of the canonical quantity-quality tradeoff, economists might be surprised that deliberate discrimination in the allocation of scarce resources to individual offspring is not self-evident but constitutes a key difference between humans and other primates. For, no matter how pervasive and culturally enshrined the ideal of selfless and sacrificing mother love might be, it does not belong to our genetic blueprint. One important explanation for this is forwarded by anthropologists like Hrdy (1999) who point out that human infants differ from young primates by their inability to cling immediately after birth. As the neonate cannot cling and signal to its mother "care for me, I am worth the effort", the degree of parental investment hinges on the expected "value" parents associate with their newborn. The facultative nature of postpartum investment constituted a major selective advantage for early hominids because – following the rationale of fitness maximization – it ensured that the high costs of brain development and childhood dependency were only incurred if the survival of the child in question was sufficiently certain to justify the effort.

Traces of this evolutionary heritage can be detected in the infant mortality rates observed in pre-modern human societies. Advanced medical care being unavailable, the survival of newborns was critically dependent on costly parental behaviors like close supervision, protection and constant provision of high-quality nutrients (e.g. by regular breastfeeding). Consequently, if humans adapted psychologically to selective pressure, parental investment and its observable outcome – childhood survival – ought to exhibit a high degree of variation in response to environmental conditions. Following this intuition, Volk and Atkinson (2013) compared infant mortality rates in pre-modern Europe, present-day hunter and gatherer societies with those observed among primates and non-human monkeys. As figure 2 illustrates, their findings strongly support the hypothesis of facultative post-natal investment, because in spite of their parents' superior intellectual capacities, the depicted survival rates of human infants are extremely heterogenous and not necessarily higher than those of newborn monkeys.

Figure 2: Infant Mortality Rates in Human Societies and among Primates

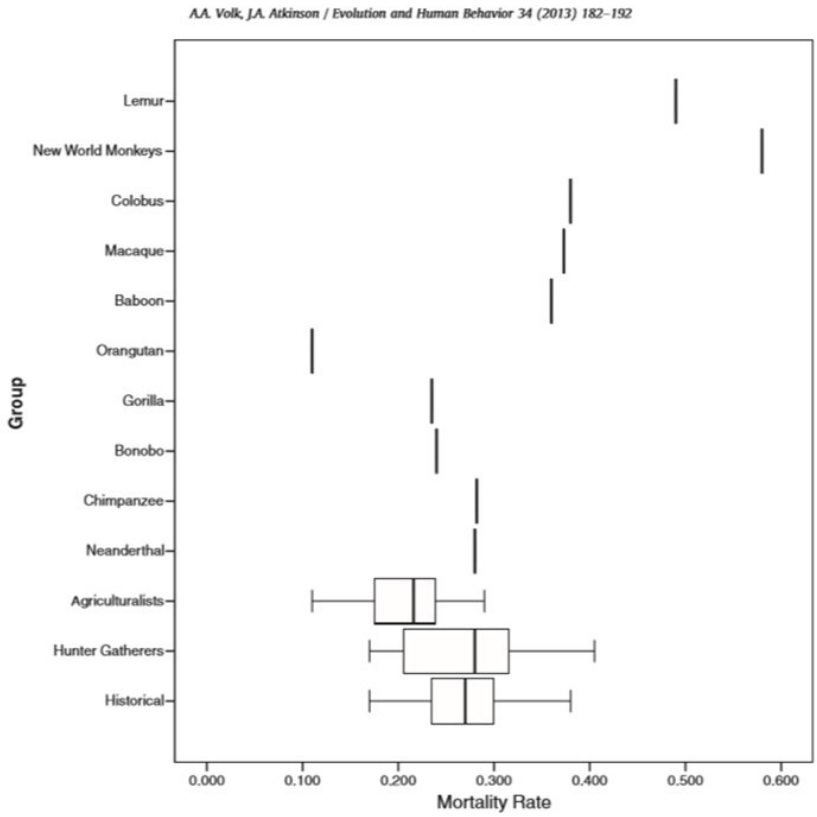


Fig. 1. Boxplot of infant mortality rate by group.

The above observation is a powerful antidote against excessive human self-esteem. More importantly though, it suggests that the human mind is probably as sensitive to extrinsic mortality as the endocrine system. Hence, as predicted by the Darwinian fertility model, human parents seem to follow psychological reaction norms which channel them towards a "high fertility and low investment strategy" whenever environmental conditions render this strategy fitness maximizing. The existence of this "Darwinian" development trap is indicated by a growing body of literature which suggests that the negative relationship between mortality and parenting effort and the positive correlation between mortality risk and fertility are universal patterns of human populations.

To give but a few examples of this work, Bereczkei (2001) found highly discriminative maternal effort based on child characteristics in a longitudinal study of 600 Hungarian women. In his sample, mothers of high-risk infants reduced their quality investment q by shortening the duration of breast-feeding and inter-birth intervals, compared to women whose infants stood better survival chances.(Bereczkei (2001), pp. 197-212). Similar results were forwarded by Nettle (2009) who used data from the Millennium Cohort Study (n = 8660 families) to show that in deprived neighborhoods in Britain, life expectancy was reduced while age at first birth was younger, breastfeeding duration shorter and reproductive rates higher. Remarkably enough, roughly the same pattern has been detected in historical European populations where – as contemporary medical textbooks put it – urban children raised in "precarious, immoral

conditions" tended to reproduce earlier than their peers in rural areas or in higher strata of the society (Krieger (1869), pp. 115). Again, the historical assessment is supported by present-day researchers including Anderson, Kermyt G. (currently developing countries), Caroline Uggla and Ruth Mace (Ireland), Charlotte Strömer and Virpi Lumaa (Pre-demographic Transition Europe) (2014), all of whom detect a positive relationship between fertility and mortality and negative effects on measures of child quality.

Seeing these corroborative correlations, it is sufficiently safe to consider the first question settled: Post-natal parenting effort is subject to evolutionary plasticity and it responds to signals of environmental harshness and unpredictability in the way predicted by the Darwinian model. Let us therefore turn to the components of the fitness function and proceed to the second question: Were we right in considering extrinsic mortality and the controllability of childhood survival (i.e. the effectiveness ε_{q*} of parenting effort) are the main causes of the observed variation in reproductive strategies? And if so, which psychological mechanisms translate high values of μ and low values of ε_{q*} into lower (perceived) utility gains from parental solicitude?

Scholars in the field of evolutionary psychology answer this question by pointing to the fact that high mortality, be it due to recurring food shortages, bad sanitary conditions, pollution or social conflict, exposes (prospective) parents to chronic stress. Prolonged exposure to harsh and unpredictable environments results in permanently elevated levels of glucocorticoids, the "stress hormones" which in turn influence the core hardware of human cognition and behavior. In particular, sustained stress affects two portions of the brain – the hippocampus and the prefrontal cortex – and impairs learning, emotional regulation and the control of aggressive outbursts. What is more, under stress complex choice problems are processed within the dorsolateral striatum of the prefrontal cortex which results in rapid, habit-based and increasingly myopic decision making. At the same time, glucocorticoids buffer the impact of neuropeptids like oxytocin and prolactin which promote bonding, trust and nurturing behavior.[16] This implies that higher extrinsic risk depresses the impact of parental effort q in the fitness *and* the individual utility function because the stress responses limit altruism, increase impatience and reduce emotional responsiveness. Under these conditions the value assigned to the (uncertain) future survival of children is low and the ability to enjoy stable and mutually affectionate relationships impaired which will discourage parents from investing too heavily in their offspring. The notion that quality investment necessitates a minimum level of security is corroborated by research in evolutionary psychology and economics alike. Thus, in a variety of different cultural and social contexts, it has been shown that parents' willingness to provide their children with formal education increased significantly after public health interventions that improved childhood survival rates.[17] Moreover, recent work by Quinlan reveals that parenting effort declines as aggressive pathogens render premature mortality an ubiquitous and uncontrollable phenomenon.[18] Finally, the influence of evolutionary algorithms is evident in defective parental behavior, particularly in postpartum depression PPD. With an

[16] Schwabe and Wolf (2009), pp. 7191-7198; Graybiel (2008), pp. 359-387; Porcelli, Lewis and Delgado (2012), p. 157

[17] Examples of this literature include: Bleakley (2007), pp. 73–117; Gibson et al. (2011), pp. 97 - 105

[18] These issues are discussed in Quinlan (2007), pp. 1606-1636

incidence rate of 5 to 25 percent (subclinical forms of the disorder not being accounted for), PPD occurs frequently enough to be of demographic and economic relevance. More importantly though, PPD is no random byproduct of the hormonal changes associated with puerperium. Rather, besides by certain genetic predispositions, post-natal depression is most accurately predicted by factors which inform mothers that either the costs of raising offspring are high (stress, lack of social and kinship support, unwanted/illegitimate pregnancy) or that the fitness gains from supporting a particular neonate are low (low birth weight, high birth order, wrong sex). Under such inimical circumstances, symptoms like fatigue, sadness, listlessness or aggression motivate mothers to withdraw investment from offspring with low reproductive value. As a result, scarce energy resources can be redirected to a new pregnancy (thereby increasing fertility and spreading the risk of reproductive failure among many offspring) or towards older evolutionary more valuable children.[19]

Let us briefly summarize the message of the above findings: Human reproductive strategies depend crucially on evolved psychological responses to environmental mortality cues. Moreover, the influence of these adaptations is highly persistent because parenting styles are "heritable" to a substantial extent. In other words, children whose parents have followed a fast life history strategy are more likely to exhibit opportunistic mating and limited nurturing effort. One reason for this is the fact that what constitutes "right" parental behavior belongs to the individual's assumptive worldview which is formed during childhood. Moreover, children who grow up under stressful conditions and insensitive or insufficient parental solicitude, experience permanently increased levels of glucocorticoids which impair their ability to form close and lasting interpersonal bonds in their adult life. (Kaplan and Lancaster (2003), pp. 170-223)

The heritability of parenting behavior is so remarkable because it leads to the perpetuation and diffusion of life history traits which maximize the fitness of their carriers. If this was not the case, individual adaptations would fail to translate into society-wide institutions and their impact would be too weak to play a perceptible role in the progression of the demographic transition. Seeing this concern, the following section will assess how high mortality stress affects the set of formal and informal institutions adopted by a population.

4.3. From the Individual to the Social Sphere:

Although losing a child is first and foremost a personal tragedy, societies with elevated levels of premature mortality tend to erect collective barriers against the traumatic impact of bereavement. The evolutionary purpose of these institutional shields is obvious: guaranteeing that parents maintain mental stability and continue to fulfill their social and familial responsibilities. Given this inherent logic, the workings of collective coping are best deciphered on the basis of their "targets", that is, based on the situational factors which determine the level of psychological stress associated with negative life events.

[19] An introduction to the topic is provided by Hahn et al. (2013), pp. 253-264

Concerning the intensity of psychological distress, trauma and resilience research posit that experiences are perceived more unsettling, the more disturbing and unforeseen the underlying event. The same holds, if the situation contradicts a person's learned world view or involves feelings of helplessness and loss of control.[20] Present day parents will hardly hesitate to associate all of the above characteristics with the death of a child. By contrast, ancestral populations adopted norms and practices which affected both adults' perception of and their expectations about childhood death.

To begin with, at times where the state of medical art did not allow to improve childhood survival significantly, the death of young children was widely considered a part of God's judgment and beyond the scope of human control. The idea that those who died young were blessed to "shuffle off their mortal coil" had an unbroken tradition from Greco-Roman times and was further reinforced by the Christian notion of a heavenly afterlife. Providing a rescue from feelings of guilt or from the emotional void of otherwise senseless loss, these beliefs permeate historical sources to the point leaving their imprint in every-day language. For example, in the 18th and 19th centuries, it was a common euphemism to say a child had "himmelt" (i.e. ascended to heaven) when referring to a premature death.[21] Moreover, religious rituals and popular traditions structured the transition from grief to the resumption of the family's normal life – a function which has been found to reduce the risk of complicated grief in both Western and non-Christian societies like e.g. China.[22]

Mostly stripped off religious connotations, a second set of psychological mechanisms targeted the unexpectedness and "injustice" of childhood death. In particular, unlike today, the first years were not regarded as a carefree prelude but as the most perilous period in life. Accordingly, high rates of infant mortality were viewed as a sad but natural condition. Like religious coping, resignation of this kind increased psychological resilience because prospective parents could emotionally prepare for the possibility to bury some of their children. At the same time, however, this process of *anticipatory mourning* produced a measure of ambivalence towards young children and caused parents to limit or postpone heavy emotional commitment.

The notion that parental attachment responds sensitively to mortality cues is supported by evidence from a wide range of past and present societies. Thus, in her seminal study on the social value of children, Zelitzer (1985) shows that until the 19th century, the death of young children was met with resignation rather than despair. Her findings are supported by the work of Stone on preindustrial England and Ariès' study on early modern France (Ariès (1982), pp. 82-83) which point to the fact that parents rarely attended the funerals of their infants or buried children "who had died to soon" in the backyard as we do with pets today.[23] With few exceptions in upper-class families, elaborate burial rites and mourning were equally uncommon in Roman Italy if a child died younger than five.(Rawson, Benjamin (2003), ch. 2)

[20] For an overview see: Zeidner and Endler (1996); Martz and Livneh (2007)
[21] Hazzi, (1802), p. 182, Wolff (1998), p. 267; Briese (2003), p. 263
[22] Lalande and Bonanno (2006), pp. 303-324.

[23] For an overview see: Stone (1977)

Of course, this does not imply that parents did not care for their children at all, it simply means that they followed a strategy, evolutionary theorists would refer to as "emotional bet-hedging". Rather than devoting all their love, physical energy and financial resources to one or two children, they spread their risk by allowing for more births and maintaining a certain distance until their offspring had survived the critical first years of life. In the words of Montaigne: "I have lost two or three children in infancy, not without regret but without great sorrow."[24] Ethnologists have documented similar attitudes in present-day traditional societies. In particular, many cultures consider children as "incomplete" and deny them full personhood until they have passed a certain age. Thus, for newborns of the Ayoreo in Bolivia it is "not unusual to remain unnamed for several weeks or months, particularly if the infant is sickly. The reason given is that – should the child die – the loss will not be so deeply felt." Similarly in Zambia, "when a Tonga child died before it was named, there was no mourning…the old women will tell the mother to hush her wailing, saying this is only a ghost."[25]

Curiously, it was as late as in the 19th century that doctors in Europe complained parents treated young infants with more or less apparent indifference. For example, a German doctor in Prussia noted that medical care was seldom asked for in case of small children, because under the age of 2 to 5 years they died too easily anyways. His confreres in Württemberg observed with barely suppressed disgust: "There are districts in our country, where parents watch 10 little children die till they call the doctor for the 11th ." And finally, country doctors in Bavaria reported that in times of economic hardship, it was always the youngest children of poor families who suffered most from indolence and therefore experienced dramatic excess mortality.[26]

Since variants of these practices have developed in uncounted cultural and historical contexts, there is little doubt that high mortality societies do transform individual adaptations into institutions at the population level. What is more, in historical settings, both the emergence and the decline of these institutions conformed closely with the predictions of the Darwinian fertility model. To appreciate this point, let us consider a set of social norms which contributed directly to elevated mortality and high birth rates. This category equally includes drastic behaviors like shortening the inter-birth interval by infanticide or abandonment of "defective" newborns and milder reactions like allo-parenting or cooperative child-rearing. While their impact on mortality is self-evident, these institutions also tended to raise fertility because "outsourcing" some childcare activities allowed parents to support more dependent children and because the spell of postnatal amenorrhea is shorter if mothers do not breastfeed or wean their babies early. A prominent example for this fragmentation of parental responsibilities is the practice of wet-nursing. In fact, the use of wet-nurses was so wide-spread in preindustrial Europe, that it had become a highly organized profession by the 17th century. Still, although some states – especially France – developed laws to regulate their employment, the majority of wet nurses were destitute peasant women and it was well known that children who were breastfed by their mothers stood much better chances at survival than nurslings.[27] Yet, since mortality was high

[24] Quoted by Zelitzer (1985), p. 18
[25] Quoted by Lancy (2014), p. 84
[26] Ärztlicher Verein Württemberg (ed.) (1868), pp. 26-28; Böing (1898), p. 64; Landeskundliche Gesellschaft Bayern (ed) (1865), pp. 399-400 , Ärztlicher Ausschuss (ed.), (1882), p. 165
[27] A survey of different countries has been undertaken by Stevens et al. (2009), pp. 32-39

anyways and – medical science being in its infancy – largely out of parental control, a couple could nevertheless increase their fitness by trying their luck at a higher number of pregnancies.

Things changed, when medical innovations like vaccination, bacteriology, disinfection and pasteurization rendered childhood survival less random. These advances not only boosted the impact of parenting effort but also tilted selective pressure from r- towards K-selective characteristics. As a result, wet-nursing became increasingly contested and vanished around the turn of the 20th century.[28] The salient feature of this process is that the decline of wet-nursing was paralleled by the rise of "modern" institutions which evidenced a trend towards higher valuation of children and "slower" life history strategies. Thus, confirming the predictions of the theoretical model, increased effectiveness of parental solicitude ($\varepsilon_q\uparrow$) caused society to redirect resources from fertility promoting to child quality promoting institutions. One example for these newly formed institutions is the establishment of pediatrics as a separate branch of medicine in late 19th century Europe.[29] Given that the differentiation occurred only *after* scientific progress had provided parents with the means to control offspring survival more effectively, the supply of specialized medical services can be interpreted as a sign of increased willingness to invest in children from a very young age onwards.

A second example relates to early childhood education: Around the middle of the century, reformers like Fröbel, Pestalozzi and – somewhat later – Montessori pioneered the idea that children were neither miniature nor incomplete or defective adults. Instead, infancy came to be considered a central developmental stage which required specific, age-appropriate stimulation of cognitive and social skills. As the new educational paradigm spread, more and more parents felt the pressure to commit precious time and resources to the education of their children. Refusing to breastfeed or failing to take basic hygienic precautions began to be publicly frowned upon just like swaddling infants, leaving them unattended, rocking them excessively or feeding them alcohol to keep them quiet.[30] This said, the fact that parental effort became fashionable while practices which had long served to facilitate frequent reproduction turned unacceptable, suggests that society was *in toto* moving from fast to slow life history strategies. Institutional change assumed a critical role in this transition because it contributed to the preservation and rapid diffusion of successful adaptations to the changed selective pressures. Due their mediating function between society and evolutionary environment, many more institutions were affected by the scientific and social leaps of the 19th century. Consequently, there is no doubt that much more could be said about labor laws and compulsory schooling or about maternity leave and the codification of child rights. However, to avoid going off on tangents, let us content with our glimpse on institutional development and summarize the major implications of the section: First, institutions mimic individual adaptations in that increased effectiveness of parental care and lower extrinsic risk trigger a self-reinforcing trend towards smaller families, higher quality investment and improved survival. Notwithstanding the potential to promote socioeconomic modernization, this result points also to development risks and suggests that improvements in survival rates alone are insufficient to promote the shift from

[28] A useful introduction is provided by: Hrdy (1994)
[29] In Germany, the first chair of pediatric medicine was established in Berlin in 1895, when the demographic transition was already on its way.
[30] Pollock (1983), pp. 2-18; Larass (2000), p. 176, Arnd (1835), p. 150

r- to K-strategies. This is the case because families will only modify fertility and parenting patterns if traditional institutions (e.g. castes, limits to social mobility) do not reduce the expected payoff of quality investments and if parents recognize the increased effectiveness of their efforts ($\varepsilon_{q\uparrow}$).

That problems of this sort are present in currently developing countries is highlighted by the notoriously poor record of birth control programs, many of which fail although contraceptives are widely available and survival rates often improved. Still, only if parents *know* that certain health interventions will prevent infant and childhood mortality, their subjective environmental risk is reduced and the tradeoff between child quantity and quality substantially altered. Consequently, societies which see for the provision of fundamental health care and the spread of medical knowledge ought to experience an earlier and more rapid fertility decline than populations where these conditions are not fulfilled. Assessing the empirical content of this proposition is of course possible, but it requires us to abandon the sphere of general psychology and cross cultural institutional patterns. For, to analyze the relative importance of biological factors rigorously, it is indispensable to control for countervailing influences like variation in geographic or socio-economic conditions.

5. Historical Case Study: Assessing the Darwinian Gift of Medicalization in 19th Century Germany

Historical evidence offers a unique possibility to test whether and to what extent medicalization promoted the transition from fast (r-selective) to slower (K-selective) life histories. The evidence in question is the experience of smallpox vaccination in 19th century Germany. The case is highly intriguing for several reasons. First, in the light of the previous sections, medical innovations must exert a strong and lasting impact on mortality if they are to alter parenting behavior perceptively. Moreover, the effect needs to be immediately apparent and the causality between the health intervention and improved survival must be sufficiently obvious to become common knowledge. Briefly summarizing the historical epidemiology of smallpox one finds that vaccination fulfills these criteria with ease.

5.1. A Brief History of Smallpox

The first unmistakable descriptions of smallpox date back to 4th century China. By the 13th century, the disease was endemic in the known world, extending its grip from Japan to Britain and from Iceland to northern Africa. From that time onwards, it followed a characteristic epidemiological pattern of regular explosive epidemics along the main trading routes of Eurasia and the Mediterranean. Since a single attack of smallpox conferred long-lasting immunity, epidemics mainly affected young children and spared communities until the pool of infectible individuals sustained a new outbreak. In spite of this periodicity, the number of children who became exposed was so overwhelming that 18th century scholars grimly observed that no-one would make a difference between the ravages of smallpox and the plague, were it not for the

habit to count children's lives only once they had survived the former.[31] Similarly stressing the omnipresence of smallpox, 19th century historians referred to them as a "communist disease" which struck rich and poor with blind disregard for class, sex or mode of living. (Bohn (1875), p.1).

For early periods, such assessments have to be taken at face value but from the 18th century onwards it is possible to quantify smallpox mortality in a number of European cities and regions. Swedish and French census data for example indicate that in the late 18th century, smallpox deaths on average accounted for 11% of overall mortality. These numbers are confirmed by some German states which feature shares between 11% (Berlin, 1754-1800), 12.5% (Eastern Prussia, 1765-1785) and 11% (Wurttemberg, 1750-1800).[32] Concerning smallpox prevalence, the picture drawn by the historical sources is even gloomier as 18th century scholars from France and Germany estimated that approximately 80% of each cohort would catch the disease – a number which their present-day confreres only slightly corrected to 66-75%. (Wolff (1998), p. 101). This situation changed dramatically with the discovery of the cowpox vaccination in 1796. In a mere 75-pages volume, English country doctor Edward Jenner revolutionized smallpox prevention by describing how immunity could be acquired by inoculating individuals with cowpox, a flu-like disease with local skin eruptions that caused no severe danger to the life of the infected person.(Jenner [1796])

Seeing the enormous death tolls of smallpox, the impact of vaccination on extrinsic mortality was substantial and it must have been sufficiently obvious to recast the tradeoff between the number and quality of children. This strong "treatment effect" alone is nevertheless insufficient to make vaccination a viable natural experiment for the demographic impact of medicalization. For any meaningful comparison it also needs exogenous variation in the spread of medical knowledge (to obtain untreated controls) and detailed data on vaccination rates, population characteristics and vital statistics. For research purposes it is therefore a fortunate fact that the South Western German states of Bavaria, Baden and Wurttemberg, passed mandatory vaccination laws in the early 19th century and carefully documented the effect of these laws ever afterwards. Statistical and anecdotal evidence are thereby particularly detailed in the case of the Grand Duchy of Baden, where the data are – for many years at least – available down to the district level. This fact bears central importance for identification because certain characteristics of Baden's vaccination system caused the impact of medicalization to be stronger in some districts than elsewhere.

5.2. Using Vaccination as an Indicator for the Perceived Effectiveness of Parenting Effort: Does Medical Knowledge Contribute to the Fertility Decline?

The reason for the differential spread of medicalization stemmed from the way inoculation was practiced during the 19th century. In practical terms, the success of the vaccination system hinged on the availability of fresh cowpox-lymph because preservation and cool storage of the vaccine remained on rudimentary standards throughout most of the 19th century. The lymph

[31] This is discussed by contemporaries and current research, see: Struve (1802), pp. 1-2; Baldwin (1999), p. 244
[32] All Data from Kübler (1901), chapter 4

used in the process was either derived from the dried material of some previous vaccination or was directly transferred from the pustules of one vaccinee to the open cuts on the forearms of another (arm-to-arm vaccination). In the first case, chances were high that the lymph had degenerated during storage and failed to cause the immunizing outbreak of cowpox. In the second case, doctors had to make sure that there was an unbroken "supply" of individuals seeking vaccination in order to maintain the process of vaccine transmission. Timing was all the more of essence as the lymph could be collected only within a limited time-frame: If the pustules were opened too early, they leaked only small amounts of fluid with little infectious material. If on the other hand the operation took place at an advanced stadium, the lymph would already be mixed with pus and debris which reduced the power of the vaccine and increased the risk of transmitting other diseases when inoculating new vaccinees.[33] To address this challenge, Baden was fast in backing up its immunization policies with an institutional infrastructure. In particular, the country established public vaccination institutes (*Impfinstitute*), that is, institutions where parents could have their infants vaccinated for free whenever they desired to do so. In exchange, they were obliged to allow the vaccinators to collect fresh lymph from the children once the pustules had reached the proper stadium. The doctors who worked for the vaccination institutes therefore needed to maintain close contact with the families in the proximity of the institute in order to make sure they would always find an unvaccinated child to produce fresh vaccine. This interaction required a great deal of patience and persuasion as a Mannheim doctor confessed in 1808:

> "...whenever vaccinations are to be performed, some children are reported ill, others do not show up because the mother has suddenly made up her mind or because the grandmother or the neighbors have argued against the operation. Or again, there is no recently vaccinated child available because the overly-protective mother refuses to wake her toddler from his sleep, because she wants to spare him the pain or because she fears to expose him to cold weather. And then probably after having waited for hours, the doctors have to restart the whole business and rash to find some other [recently vaccinated] child - a vexation that unnerves even the most patient personality."[34]

But what was arguably a nuisance for the physicians might have had life-saving consequences for the babies they "milked" for vaccine. This was the case because public doctors and publicly salaried midwives were exhorted to frequently check on the sanitary conditions in their constituencies, to inform parents of proper childcare and promote health preserving habits like breastfeeding. But of course, the zeal with which these duties were fulfilled were a matter of the respective health practitioners' discretion. Yet, these matters were more likely addressed if the doctor had to visit the families anyways.[35] What was more, while vaccination was mandatory for all children, the provision of lymph for future vaccinations was not. The vaccinators would therefore have to explain the benefits of the measure to young parents and convince them of its safety. By doing so they contributed to a profound change in the popular

[33] This particularity is discussed both by Cless (1875), chapter 1 and Heim (1838)
[34] GLA Karlsruhe 236-16034 (Sanitary Commission), Report of the *Physicus* of Mannheim, 1st of November 1808
[35] Laws and ordinances in: (Baden) Philipp Carl Baur von Eiseneck, Sammlung sämtlicher Gesetze, Verordnungen, Instructionen, Belehrungen und Entscheidungen, welche in dem Gross-Herzogthume Baden über Gegenstände der Gesundheitspolizey erscheinen sind, Karlsruhe 1830

conception of infant and childhood mortality. For, prior to vaccination, deeply rooted folk wisdom had perpetuated almost all of the adaptations predicted by the theory of anticipatory mourning, Thus, it held that diseases like smallpox were an inevitable ailment, a process of purification rather than a disease and were necessary to eliminate evil substances from the body before it reached adulthood. Sometimes, the idea of physical purification was moreover coupled with the religious belief that epidemics were manifestations of God's judgment and punishment.[36] As a result, early sources like the following 1813 report are littered with complaints of "popular ignorance and prejudice"

> "...Thus, the following prejudice is deeply rooted [in popular opinion], that the natural smallpox poison is a priori contained in the body (...), that this poison cannot be emitted from the body through the few pustules [after vaccination] and that this will eventually give rise to other diseases even if it might prevent the outbreak of smallpox."[37]

In spite of the vaccination's early success, these attitudes did not vanish in the course of time. If anything, traditional conceptions of smallpox took on an elitist and pseudo-scientific inflection from midcentury onwards. This tilt of balances resulted from the fact that increasing professionalization and medical progress went along with a heightened interest in what would today be referred to as "alternative medicine." Among other centers, the university towns and spas in the German South attracted all sorts of naturopathic practitioners – trained doctors as well as laymen – who produced an impressive body of literature and succeeded in conquering their share of the mass media including newspapers and popular journals. A motley group they were, the adherents of homeopathy, herbal cures or hydrotherapy were nevertheless united in their contempt for "unnatural" methods like vaccination and promised corporeal balance through "natural" cures and diets. Again, this attitude suggested that disease was an integral part of life, more a manifestation of physical disequilibrium than a harm to be combated. And since alternative medicine received an uncommonly intensive coverage in contemporary media, the naturopathic fashion fostered anti-vaccionism just the same way as traditional prejudices had before.

Traditional and seemingly "modern" opposition against vaccination are central for our purposes because they influenced attitudes towards environmental risk and thereby co-determined the expected impact of parenting effort (ε_q). Whether or not a household perceived childhood mortality as something inevitable or as a danger to be actively prevented depended critically on the frequent interaction and constant communication with academically trained doctors. Parents who participated in the vaccination system were therefore arguably better able to prevent infant mortality and might therefore have switched to a slower life history strategy than other households.

[36] This was particularly true during the first decades of the 19th century when widespread ignorance in medical matters fatally coupled with the symptoms and age-specific incidence of smallpox. Well before the bacteriological revolution of the 1860s, it was simply *too* plausible to assume an infection inevitable that used to affect more than two thirds of all children. In addition, symptoms like high fevers, intense pain, rashes and pustules leaking a fetid mixture of lymph and tissue debris were easily interpreted as signs of the body struggling to free itself from evil fluids.

[37] Werfer (1813), p. 146

5.3. Estimating the Impact of Medicalization:

To test this conjecture, I collected socio-economic indicators and population data of Baden's 67 districts (*Amtsbezirke*). Regarding the measure of medicalization, archival records allowed me to locate the vaccination institutes and their years of establishment. In addition, it was possible to retrieve information on economic characteristics of the districts in the sample including urbanization and poverty rates. Using these data[38], I estimated the following econometric model:

$$TotalFert_{i,t} = x_1 Totalmort_{i,t} + x_2 VaccIns_{i,t} + x_3 Poverty_{i,t} + x_4 City + Tim\ Controls + Regional\ Controls + \varepsilon_{i,t},$$

The dependent variable *TotalFert* corresponds to the number of births per 100 inhabitants in district i and year j. Whether local fertility levels responded to evolutionary pressures or not is assessed by the coefficient of *Totalmort,* that is, by the number of infant deaths per 100 inhabitants.[39] Positive values of x_1 thereby indicate that high mortality was associated with high fertility, whereas negative x_1 suggest an inverse relationship between the two. Since the effectiveness of parenting effort depends on an interplay of scientific and social institutions, ε_q is approximated by two indicators of medicalization which capture key determinants of parents' ability to improve the fitness of their offspring. The first, of these indicators is the binary variable "*VaccIns*" which takes the value of 1 if a vaccination institute was operating in district i during year j and zero otherwise. Given the institutes' reliance on popular support, *VaccIns* can reasonably be assumed to reflect the spread of medical knowledge – and hence parents' expected capacity to exert control on childhood survival. Knowledge was – however – of limited use to improve childhood survival, if medical services were not available. The second variable, *"DOC",* therefore assesses whether households had access to fundamental healthcare at all. The intuition behind the measure is as follows: In addition to local mortality data, the Badenese mortaliy records note whether the official certificate of death was signed by a doctor or a local official. The value of this information stems from the fact that fatal diseases or accidents are commonly considered the core expertise of a doctor. If a death was signaled to some petty official, it is natural to suspect that the bereaved family was either unable to reach a physician or did not trust in modern medicine. Hence, by computing the fraction of deaths signaled by a doctor, it was possible to construct a measure for the availability and the popular reliance on medical services. In order to control for the potential influence of socio-economic characteristics, I further introduced the variables "City" and "Poverty", which respectively

[38] The sources are: Generallandesarchiv Karlsruhe:) 236-16019-16045 and 234-6041, Staatsarchiv Freiburg, 908/1-135 (Leichenschau), Statistik der Inneren Verwaltung des Großherzogthums Baden, various issues 1851-1893

[39] Since the data contain only information on mortality within certain age brackets, I labelled all deaths age < 1 as "infant deaths"

account for urbanization and the prevailing poverty rate.[40] Finally, district dummies and time controls were employed in order to capture local and time fixed effects. The results of the estimation are depicted in Table 1. In addition to the simple OLS model, specification 2 uses district and time controls. Moreover, to reduce the risk of extreme observations biasing the estimates, specification 3 performs a quantile regression of the median.

Turning to the results, one finds that the estimated coefficient reflect the predictions of life history theory almost perfectly. For one, high overall mortality – i.e. high environmental risk – comes along with statistically significant increases in fertility. Next, arguably reflecting Malthusian dynamics, poverty reduces fertility, whereas the impact of urbanization remains – somewhat surprisingly – insignificant. More importantly though, medicalization did promote the expected shift towards reduced family size because the coefficients of *VaccIns* and *DOC* are statistically significant and negative in all (respectively in two out of three) specifications.

Table 1: Estimated Effects of Medicalization on Fertility

Model	***OLS***		***OLS, Panel***		***Quantile Regression***	
Variable	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat
TotalMort	**1.0000306**	20.02	**1.175.945**	24.16	**1.089.841**	23.37
VaccIns	**-0.4718797**	-4.96	**-0.4563893**	-5.37	**-0.5315497**	-5.98
DOC	**-0.1364501**	-1.37	-0.1667520	-1.34	**-0.241978**	-2.54
Poverty	**-0.0453486**	-7.91	**-0.0393697**	-7.54	**-0.0443673**	-8.08
City	**-0.0636008**	-1.12	**-0.1561412**	-2.96	-0.0377227	-0.7
Time Controls	Yes		Yes		Yes	
Regional Controls	No		Yes (Amtsbezirke)		Yes (Amtsbezirke)	
Adj-R^2	0.51		0.62		Pseudo R^2 = 0.39	
Number of Observations	801		801		801	

Having recognized the congruence between the theoretical model and the historical dataset, let us dare a tentative assessment of the respective size of the mortality and medicalization effects: Addressing first things first, consider the coefficient of *Totalmort.* Across all specifications, x_1

[40] Poverty captures the fraction of individuals who were entitled to poor relief among the deceased in the corresponding year. Assuming that the poverty rate among the dead was approximately equal to that among the survivors, poverty should be an accurate measure of the actual poverty rate in a district.

is larger or equal to one which implies that each additional death was offset by a corresponding increase in fertility. This result is not trivial since it provides one of the central hypotheses of the Darwinian fertility model with empirical backbone: The effects of Malthusian checks are strictly limited because elevated levels of mortality channel the population towards an r-strategy with high levels of "replacement" fertility.

Reductions in environmental risk and increased effectiveness of parental effort produce exactly opposite results. The impact of improved information – as measured by *VaccIns* - is particularly strong with districts where a vaccination institute was located recording estimates of about 0.4-0.5 births less than elsewhere. Qualitatively, the impact of doctors' availability is similar, although the estimated coefficients do not attain statistical significance in all specifications. However, to the extent the estimates turn out significant, their quantitative impact is not negligible. In particular, having one percent more death certificates signed by a doctor comes along with approximately 0.13 - 0.24 fewer births per year.

There is yet another interesting link between the historical data and the Darwinian fertility model: Based on LHT, parents respond not only to cues of extrinsic mortality but also to variation in the effectiveness of their efforts. In particular, we expected fertility to be smaller, the greater the effectiveness of parental investment (ε_q). Using quantile regression and analyzing the estimated coefficients at different points of the distribution allows us to obtain some indication about the accuracy of this prediction. The intuition behind this procedure is as follows: If parents desired larger families when facing environments which reduced the impact of fitness promoting investment, the coefficients of the medicalization variables ought to be smaller for higher quantiles of the fertility distribution. This being the case, high fertility could be considered an indicator of ineffective or dysfunctional health institutions and – correspondingly – reduced parental inclination to adjust fertility plans downwards. The information to test this conjecture can be obtained from an interquantile range regression of the difference between quantiles. The coefficients in the difference regression are thereby interpreted as follows: If a coefficient x_i takes the value of zero, the sensitivity of fertility to variable i is the same at all points of the distribution. If x_i is significant in the interquantile regression and carries the same sign as in the original estimation, the expected effect of variable i is larger at higher levels of fertility. Finally, if x_i reverses its sign, the fertility response to changes in i is smaller at high levels than at lower ones. As shown in table 2, the coefficients of the medicalization variables seem to obey to the logic of LHT outlined above. Thus, although the difference is only significant for *VaccIns,* the absolute impact of medicalization is damped at the upper (90% quantile) than at the lower end (10% quantile) of the fertility distribution. Put differently, if fertility was high, the response to medical infrastructure was comparatively weak and parents apparently based their family planning on other factors. By contrast, at low levels of fertility, the response to medicalization is reinforced which is suggestive – albeit no conclusive proof – of a heightened sensitivity to improvements in ε_q and the presence of a quantity-quality tradeoff.

Table 2: Interquantile Regression

Model	***Interquantile Regression***	
Difference Variable	Coefficient	t-stat
TotalMort	-0.1644442	-0.66
VaccIns	**0.4907308**	**1.75**
DOC	0.0011318	0.32
Poverty	0.0165889	1.28
City	**-0.365399**	**-2.23**
Pseudo-R^2 0.1-quantile	0.43	
Pseudo-R^2 0.9-quantile	0.47	
Number of Observations	801	

5.4. Robustness against Endogenous Mortality:

At this point, critical readers might be troubled by the possibility of reversed causation in the relationship between fertility and mortality. This problem would occur if high mortality was a direct cause of high fertility. Given the facultative nature of parenting behavior, the idea is not too implausible. For example, parents might have sought to "correct" an unexpectedly large family size by withdrawing investment from children of higher birth orders or of the "wrong" sex.

In order to account for this problem and to obtain unbiased estimates for the effect of *extrinsic* mortality μ, I therefore performed an instrumental variable estimation for the coefficient of *Totalmort*. To be a viable instrument for exogenous mortality a variable should be strongly correlated with mortality while being independent of fertility. Fortunately, the scrutiny of Baden's statistical officers provides us with such a measure. For, besides information on fertility, mortality and core socio-economic characteristics, the records also feature annual statistics on the frequency of specific causes of death. Some of these, like diarrhea or frailty arguably reflect harsh environments (i.e. extrinsic mortality) *as well as* parental negligence. Others like accidents or coronary diseases had probably little impact on infant and childhood mortality. Yet, there is one disease which breaks with this pattern: pertussis. The reason for this is first that the disease follows a specific course which starts with mild cold-like symptoms before gradually escalating to the characteristic whooping cough. In addition, pertussis is an airborne disease which spreads when an infected person spills tiny droplets containing the

bacillus *Brodetella Pertussis*. Yet, although the latter causes the infection, the most severe symptoms are not the workings of the bacillus itself but stem from a toxic it produces to override the body's immune system. Thus, the disease is most contagious during its early stages when the concentration of *Brodetella Pertussis* is high whereas the symptoms are either mild and unspecified or missing altogether. In consequence, there was little to do for parents who sought to protect their children from the disease nor was it possible to expose unwanted children directly to contagion. At the same time, pertussis occurred epidemically and in a largely random manner. In other words, it is hard to conceive a channel through which high fertility might have influenced the prevailing level of pertussis mortality. Following this rationale and using the number of deaths from pertussis as an instrument for *Totalmort*, I found the estimation results depicted in Table 3.[41] As it turns out, the positive effect of mortality is robust and remains statistically highly significant. As predicted by life history theory, mortality therefore exerts a systematic influence on fertility.

Table 3: Instrumental Variables Estimation

Model	***IV***	
Variable	Coefficient	t-statistic
TotalMortIV	**0.7221714**	**2.15**
VaccIns	**-0.51874**	**-5.14**
DOC	-0.0006542	-1.39
Poverty	**-0.051114**	**-5.01**
City	**-0.1771814**	**-3.07**
Time Controls	Yes	
Regional Controls	Yes	
Adj-R^2	0.26	First Stage F_Statistic: 40.83 , Prob>F =0
Number of Observations	798	

[41] The data were drawn from the same sources as in Tables 1-2

5.5. Medicalization and the Quantity-Quality Tradeoff; Effects on the "Value" of Children

The previous results clearly suggest that the empirical patterns of human fertility respond to signals of elevated mortality. Still, with regard to theory, this is merely half the battle. After all, the Darwinian fertility model further predicts that the decline in fertility ought to initiate an increase in child quality investment. Thus, if the theory fits with socio-biological facts, we would expect parents in the better medicalized districts of the sample to have formed early emotional bonds and to have invested more heavily in the wellbeing of their children. Assessing this prediction is empirically more involved than identifying the effect on fertility because it requires to quantify the – almost immeasurable – degree of parental attachment and the propensity to invest in child quality.

At first glance, educational indicators like schooling or literacy rates seem suitable candidates for the desired measure of parental quality investment. Yet, there at least with respect to the Baden dataset, there are reasons to doubt the accuracy of these "natural" indicators. To begin with, literacy was almost universal in South Western Germany during the observation period. For example, a survey by the ministry of justice showed that 97% of the inmates in the prisons of the state of Baden had fundamental reading and writing skills in the 1850s.[42] In other words, being able to read and write was too common a characteristic to qualify as a measure of exceptionally high parental solicitude. Much the same holds for schooling because primary education was mandatory and free for all children under the age of 14. Secondary education by contrast was offered by many different institutions, ranging from the *Gymnasium* and the *Oberrealschule* to different types of vocational courses. As a result, data on the attendance of these – not necessarily public – schools is scattered and does probably not even capture the true extent of parents' quantity-quality tradeoff. For, while youngsters in an academic secondary school like the Gymnasium did not contribute to the family budget, students who enrolled in vocational schools like the *Höhere Bürgerschule* might also have pursued an apprenticeship and made a living on their own. Finally, schooling rates fail to reflect the influence of early childhood mortality because investments in formal education were made well after the critical first years of the child's life when mortality was highest.

Consequently there is little doubt that a different measure for the value placed in child quality is called for. In fact, the historical record does provide such an indicator. For, doctors, local officials, teachers and clergymen frequently observed that the mortality rate of legitimate children was far lower than that of illegitimate children. The same observers did not hesitate to spell out the reasons for this fact. As the prestigious Journal of Public Health (Zeitschrift für die Staatsarzneikunde) put it in its 1837 issue:

> "…having no other choice, mothers of illegitimate children leave their newborns with a nurse. In most cases, nurses are poor women who receive a mere pittance for the sad business of taking these starved, neglected children and letting them die as soon as possible. Only the friends of the poor and the doctors know in what sheds and cages these wretched little creatures are kept during their short lives, in what dirt they dwell, what disgusting stodge they are fed and what appalling maltreatment they experience every day…"[43]

[42] Diez (1863), p 102
[43] Zeitschrift für die Staatsarzneikunde (1837), p 448

To the extent that excess mortality of illegitimate children was chiefly due to neglect or maltreatment, it can be considered a sign of parents pursuing an exceedingly fast life history strategy and trying to get rid of the unwanted consequences. However, in districts where infant mortality rates were more or less the same for legitimate and illegitimate children, the value placed in each child – or the social sanctions for neglect - were probably high enough to set incentives for a slower "low fertility/high quality investment" strategy.

Relying on this intuition, I constructed the variable "*Value*" which measures the excess mortality of illegitimate children relative to overall infant mortality:

$$Value = \frac{Totalmort - Illegmort}{Totalmort}$$

The expression is easily interpreted: If there is no difference between the mortality rate of illegitimate children and the district average, *value* will be equal to zero. By contrast, if the risk of dying is higher for illegitimate than for legitimate children, *value* will be negative. Finally, in the unlikely case that the group specific mortality risks are reversed, value changes its sign and becomes positive.

As before, the underlying data were drawn from the medical statistics compiled by Baden's Ministry of Inner Affairs and comprise records of all births and neonatal deaths of legitimate and illegitimate children in the 67 districts (Amtsbezirke) during the period 1851 - 1863. Using this information, I estimated the following econometric model

$$Value_{i,t} = x_1 Totalmort_{i,t} + x_2 DOC_{i,t} + x_3 Poverty_{i,t} + x_4 VaccIns_{i,t} + x_5\, city + Tim\ Controls + Regional\ Controls + \varepsilon_{i,t}$$

The results of the estimation are depicted in table 4. As in the analysis of Baden's fertility patterns, the estimated coefficients fit well with evolutionary theory. In particular, high fertility and poverty seem to prevent early attachment and facilitate the neglect and abuse of unwanted illegitimate children (i.e. the coefficients of *Totalfert* and *Poverty* are significant and negative). By contrast, the indicators of advanced medical coverage, *DOC* and *VaccIns* come along with an increased valuation of children (although statistically not significant in the latter case). Thus, as predicted by LHT and the Darwinian model, parents who face low exogenous mortality ($\mu\downarrow$, small values of *Totalmort*) and who know that offspring survival is a matter of parenting effort and health investment ($\varepsilon_q\uparrow$, high values of *DOC* and/or *VaccIns* = 1) will be more willing to invest physically and emotionally in their children.

Table 4: The Determinants of Child "Value"

Model	***OLS***		***OLS, Panel***	
Variable	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**
TotalMort	**-0.0019081**	**-3.07**	**-0.0018655**	**-2.71**
DOC	**0.0000264**	**2.08**	**0.0000242**	**1.89**
Poverty	**-0.0001432**	**-1.96**	**-0.0001312**	**-1.78**
VaccIns	0.0011753	0.97	0.0013295	1.09
City	-0.0000692	-0.1	0.0001405	0.18
Time Controls	**Yes**		**Yes**	
Regional Controls	**No**		**Yes (Amtsbezirke)**	
Adj-R^2	**0.033**		**0.044**	
Number of Observations	**801**		**801**	

Summarizing the empirical evidence, the case of 19th century Baden comes down to two lessons, both of which support the earlier theoretical results: First, the transition from high to low levels of fertility is substantially accelerated if the extrinsic mortality risk is low and if fundamental medical knowledge is widely spread. Moreover, this beneficial effect on fertility carries over to the valuation of children and translates into reinforced parental effort and – arguably – higher investment in human capital.

4.6. Corroborative Evidence:

Seeing their theoretical and practical implications, it is indispensable to assess the general applicability of the above results. Fortunately, Baden's experience is no isolated case although it alone provides sufficiently detailed micro-level data to conduct a rigorous econometric analysis. In spite of these limitations, the notion that medicalization acts as a catalyst for the onset of a fertility decline is supported by the experiences of Baden's neighbors Württemberg and Alsatia. Starting with the former, one finds that medicalization in Württemberg was generally less comprehensive than in Baden. The country had not only failed to establish vaccination institutes but was also more lax in enforcing its vaccination laws. Moreover, in the course of the 19th century, Baden's universities turned into prestigious centers for medical research and provided the country with large numbers of academically trained doctors who – not the least for their own benefit – became important proponents of rapid comprehensive medicalization. By contrast, Württemberg was a hot-bet of vociferous anti-vaccination agitation

and a center of alternative medicine. Scholarly medicine was therefore met with more mistrust and information about new prevention and treatment methods spread much more slowly than elsewhere. There was – however – one district of Württemberg which broke with this pattern: In the Jaxtkreis, local schoolteachers and the clergy were convinced of the benefits from immunization and determined to promote the universal vaccination and revaccination of all young adults. Having no other means to accomplish this aim, they came up with the ingenious – albeit morally questionable – idea to officially announce a forged "amendment" to the Württemberg vaccination laws. The additional article was said to require the revaccination of all 14-year olds. Local officials saw no reason to prevent this staging and did not impede the regular revaccination of all students at the district's secondary schools.[44] The initiative of local officials, doctors and teachers therefore worked in a similar way as the Badenese *Impfinstitute* and spread fundamental medical knowledge. As a result, immunization in the Jaxtkreis was almost universal while smallpox mortality remained at negligible levels and hardly reached *half* the size of the Wurttemberg average. The remarkable feature of the story is that increased medicalization was not confined to successful prevention of smallpox. Rather it led to lower levels of general infant and childhood mortality and – as if to prove the predictions of life history theory true – this reduction of environmental risk translated into lower fertility rates. The available mortality and fertility rates of Württemberg and the Jaxtkreis are depicted in table 5. Both rates are significantly lower in the latter than the Württemberg average, although the dataset is too limited to construct the control variables for a multivariate regression as in the case of Baden.

Table 5: Mortality and Fertility in Wurttemberg and in the Jaxtkreis[45]

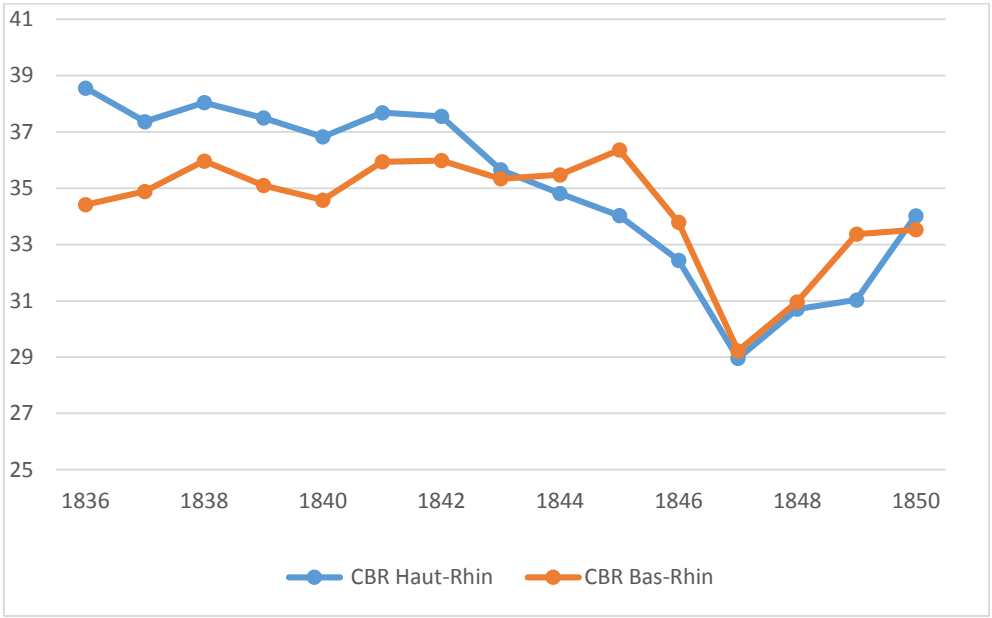
	1835-46	*1847-55*	*1856-67*	*1868-80*	*1881-95*
Infant Mortality Jaxtkreis (Deaths per 100 life births)	32.7	31.5	31.4	30.1	23.3
Infant Mortality Wurttemberg (Deaths per 100 life births)	33.9	34.8	35.4	31.9	24.8
Fertility Jaxtkreis (Births per 1000 inhabitants)	40.9	36.4	40.2	43.0	34.5
Fertility Wurttemberg (Births per 1000 inhabitants)	43.6	38.1	40.7	44.6	35.9

[45] Lange (1891), data appendix

Finally, similar findings can be established if one ventures a look across Baden's Western frontier to the French region of Alsatia. The adjacent two departments of the Bas-Rhin and the Haut-Rhin resembled their Eastern neighbor strongly in terms of socio-economic structure. Nevertheless, their health infrastructure deviated substantially from Baden's path In the Bas-Rhin, the prefectural administration had established so called *médecins cantonaux* at the beginning of the 19th century. Besides offering public vaccinations, these doctors were in charge of providing free health care for the poor and of supervising the sanitary conditions in their constituencies. A similar office was not established in the Haut-Rhin until a smallpox epidemic had ravaged the department in 1836/37 claiming more than 10000 lives. Medicalization was therefore less advanced and perceived extrinsic mortality arguably higher than in the Bas-Rhin. Figure 3 shows the effect of the institutional differences on fertility rates. Again, as predicted by LHT, the closer net of fundamental medical services apparently helped to lower fertility rates. Moreover, after the introduction of a local health-care system similar to that of the Bas-Rhin in the 1840s, the fertility difference between the two departments vanishes and eventually reverses.

Figure 3: Fertility in Two French Départements 1836-1851[46]

(Crude Birth Rates = Births per 1000 Inhabitants)



This impression is further corroborated by the available evidence on regional levels of school attendance. Since elementary schooling was less universal in rural France than in South-Western Germany, differences in the two departments' attendance rates are a more reliable indicator of variation in child quality investment than in Baden (or Wurttemberg for that matter). Here again, the Bas-Rhin outperformed its Southern neighbor in the 1830s and 1840s with an average attendance rate of 71% versus 64% in the Haut-Rhin. Yet, the difference disappears in the mid-1850s, where both departments recorded approximately equal schooling

[46] Data drawn from Archives Départementales du Bas-Rhin, 5M67 and Archives Départementales du Haut Rhin, 5M68

rates of 75.9% and 76.4%.[47] Still, with all due caution, the evidence seems to support life history theory in that lower fertility in the Bas-Rhin came along with higher human capital investment and more advanced medicalization.

6. Conclusion:

For the better part of human history, life was most fragile and death most imminent during infancy and early childhood. Today, the death of a child is considered a tragedy, hardly bearable and deeply unnatural. Yet, what seems abhorrent to present-day observers, was a sad but commonplace experience for our ancestors. While this is deeply revolting from a humanitarian perspective, certain currents in macroeconomics and economic history claim that the Malthusian check on population growth ultimately raised per capita income, facilitated capital accumulation and thereby contributed to slow but sustained growth in pre-industrial Europe. The present paper has questioned this convention. Its doubts are founded in recent results forwarded by the life sciences, including anthropology, evolutionary biology, medicine and psychology. Based on life history theory, researchers in these fields provide ample evidence that human reproductive behavior is subject to evolutionary plasticity. Mating and parenting strategies are apparently tailored to different ecological niches in order to maximize long term fitness. Fitness in turn has been shown to be a function of exogenous characteristics with unpredictable, high-mortality environments favoring high fertility low investment and stable, predictable environments giving a selective advantage to more investing low fertility high quality strategies.

Putting these observations into mathematical form and incorporating them into a conventional economic model of fertility the paper has established that declining extrinsic mortality or increased effectiveness of parental effort both lead to a switch from the pre-demographic transition pattern of high fertility and mortality to the post-transition pattern of reduced family size and heavy investment in child quality. That this result is more remarkable than it appears at first glance stems from the fact that evolution has created strong mechanisms which nudge parents towards faster or slower life history strategies. Thus, a number of physiological and psychological adaptations determine both the feasible number of births and parents expected costs and benefits from investing in the competitiveness of their offspring. Whether or not parents face a quantity quality tradeoff at all and how they value the different alternatives thereby depends on the prevailing mortality framework. Consequently, we expected fertility to decline earlier and child quality investment to increase faster the lower environmental mortality risk.

To test this proposition empirically, the paper has studied the effects of smallpox vaccination in 19th century Baden. Besides for its exceptionally good availability of statistical data, the case of smallpox was particularly instructive because prevalence was high and the decline in mortality after the introduction of universal immunization dramatic. Adding to that, the state of Baden was fast in establishing a highly institutionalized health system. This feature proved

[47] Unfortunately, the surviving statistics contain only averages and no annual data. Therefore a more precise analysis of schooling rates was impossible.

crucial for identification purposes because the centralized vaccination system required the establishment of so called vaccination institutes to ensure a constant supply of high quality vaccine. Seeing the technical constraints of the commonly practiced arm to arm vaccination this implied that the doctors of a vaccination institute needed to maintain permanent contact with the population in their district. Medical knowledge and fundamental health services were therefore more accessible in districts which had a vaccination institute than elsewhere. Using this difference allowed us to estimate the effect of reduced environmental risk through improved health conditions. As it turned out, progress in medicalization came along with lower fertility and higher valuation of the individual child even if socio-economic factors were controlled for. The results were moreover robust against potentially endogenous mortality and were backed by anecdotal evidence from adjacent regions.

The bottom line of these findings is quickly stated: Neither biology nor history obey neatly to Malthusian rules. If anything, the two of them have crafted human behavior in a way that mitigates the impact of positive checks on population and precludes the emergence of preventive ones under adverse conditions. Thus, where the provision of fundamental health care and educational progress are concerned, there is no contradiction between economic and humanitarian objectives. For, in essence the logic of evolution is simple: If death lurks around every corner, he can only be bet on numbers. Of course, this strategy ensures survival but it does so at the cost of reduced human capital investment, stagnation and miserable living conditions. Only if survival is sufficiently certain, the evolutionary dividend of parental effort rises and the population converges towards a new equilibrium in which families are smaller but – arguably – a great deal better off.

References:

Unprinted Sources:

Archival Resources:

1. Generallandesarchiv Karlsruhe (GLA Karlsruhe) 236-16019-16045, 234-6041

2. Staatsarchiv Freiburg, 908/1-135 (Leichenschau)

3. Archives Départementales du Bas-Rhin 5M67

4. Archives Départementales du Haut-Rhin 5M68

Historical Periodicals and Journals:

Annuaire Statistique du Département du Bas-Rhin

Annuaire Statistique du Département du Haut-Rhin

Hufelands Journal der practischen Heilkunde

Mitteilungen aus dem Reichsgesundheitsamt, various issues, 1881-1884

Rapport sur les vaccinations pratiquées en France pendant l'année

Statistik der Inneren Verwaltung des Großherzogthums Baden

Statistisches Jahrbuch für das Land Baden

Statistisches Jahrbuch für das Königreich Württemberg

Zeitschrift für Staatsarzneikunde

Literature:

Aries, Philippe, The Hour of Our Death, New York, 1982, pp. 82

Arnd, Karl, Die materiellen Grundlagen und sittlichen Forderungen der europäishen Cultur, Stuttgart 1835

Ärztlicher Ausschuss (ed.), Ärztliches Intelligenz-Blatt, vol. 29, München, 1882

Ärztlicher Verein Württemberg (ed.), Die Kindersterblichkeit in Württemberg, in: Medicinisches Correspondenzblatt, vol. 38 No. 3, 1868
Azarnert, Leonid, Child Mortality, Fertility and Human Capital Accumulation, in:Journal of Population Economics, vol. 19 no. 2, 2006, pp. 285-297

Barro, Robert, Becker, Garry, A Reformulation of the Economic Theory of Fertility, in: Quarterly Journal of Economics, vol. 103 no.1, 1988, pp 1-25

Baur von Eiseneck, Carl, Sammlung sämtlicher Gesetze, Verordnungen, Instructionen, Belehrungen und Entscheidungen, welche in dem Gross-Herzogthume Baden über Gegenstände der Gesundheitspolizey erscheinen sind, Karlsruhe 1830

Bleakley, Hoyt. "Disease and Development: Evidence from Hookworm Eradication in the American South." In: The quarterly journal of economics, vol. 122 No. 1, 2007, pp. 73–117

Böing, Heinrich, Neue Untersuchungen zur Pocken- und Impffrage, Freiburg 1898

Brainerd, Elizabeth, Siegler, Mark, The economic effects of the 1918 influenza epidemic, CEPR Discussion Paper, No. 3791, 2003

Briese, Olaf, Angst in Zeiten der Cholera, Berlin 2003

Caldwell, John, Demographic Transition Theory, Dordrecht, 2006

Cameron, N.M. et al.Epigenetic programming of phenotypic variations in reproductive strategies, in: Journal of Neuroendocrinology, vol. 20, 2008, pp. 795-801

Chisholm,**and Quinlan**, Early stress predicts age at menarche and first birth, in: Human Nature, vol. 16 No. 2, 2005, p. 233-265

Clark, G.,A farewell to alms: A short economic history of the world, Princeton University Press, 1st ed. Princeton 2007

Diez, Rudolf, Die Gewerbe im Großherzogthum Baden, Karlsruhe 1863

Ellis, Timing of pubertal maturation in girls: an intergenerational life history approach, Psychological Bull. Vol. 136 no. 6, 2004, p. 920-958

Gibson, Mhairi A. et al., Modernization increases parental investment and sibling resource competition: evidence from a rural development initiative in Ethiopia, in: Evolution and Human Behavior , Vol. 32 No. 2 , 2011, pp. 97 - 105

Graybiel, Ann, Habits, Rituals and the Evaluative Brain, in: Annual Review of Neuroscience, vol. 31, 2008, pp. 359-387

Hahn Holbrock et al (2013) Clin Psychol Sci., vol. 1 No. 3, 2013, pp. 253-264

Hazzi, Joseph, Statistische Aufschlüsselungen über das Herzogthum Baiern aus ächten Quellen geschöpft, Nürnberg 1802, p. 182

Heim, Franz Historisch-kritische Darstellung der Pockenseuchen, des gesamten Impf- und Revaccinationswesens im Königreiche Württemberg, Stuttgart 1838

Heinrich Bohn, Handbuch der Vaccination, Leipzig 1875

Hotz, Klerman, and Willis (1993)

Hrdy, Sarah, Fitness tradeoffs in the history and evolution of delegated mothering with special reference to wet-nursing, abandonment and infanticide, in: Parmigiani et al (ed.), Infanticide and Parental Care, Chur, 1994

Jenner, Edward, An inquiry into the causes and effects of the variola vaccine, a disease discovered in some of the western counties of England, particularly Gloucestershire, and known by the name of the cow pox, London 1796

Jones JH, Bird RB. The marginal valuation of fertility.in: Evolution and human behavior : official journal of the Human Behavior and Evolution Society, vol. 35 No. 1, 2014, pp. 65-71.

Kaplan, Hillard, Lancaster, Jane, An Evolutionary and Ecological Analysis of Human Fertility, Mating Patterns and Parental Investment, in: Wachter, Kenneth et al. (ed.), Offspring, Washington D.C., 2003, p. 170-223

Krieger, Eduard, Die Menstruation, eine gynäkologische Studie, Berlin 1869

Lagerlöf, Nils-Petter, From Malthus to Modern Growth: Can Epidemics Explain the Three Regimes? in: International Economic Review, vol. 44 no. 2, 2003, pp. 755-777.

Lalande, Kathleen M., Bonanno, George, Culture and continuing bonds: A prospective comparison of bereavement in the United States and the People's Republic of China, in: Death Studies vol. 30 No. 4, 2006, pp. 303-324.

Landeskundliche Gesellschaft Bayern (ed.), Bavaria. Landes und Volkskunde des Königreichs Bayern, vol.3, München 1865

Lancy, David, Babies Aren't Persons: A Survey of Delayed Personhood, in: Otto, Hiltrud et al., Different Faces of Attachment, Cultural Variations of a Universal Human Need, Cambridge 2014, p. 66-109

Lange, Wolfgang, die Bevölkerung Württembergs, Stuttgart 1891

Larass, Petra, Kindsein kein Kinderspiel, Halle, 2000

Martz, Erin, Livneh, Hanoch, Coping with Illness and Disability, New York 2007

Pollock, Linda, Forgotten Children: Parent-Child Relations from 1500 to 1900, Cambridge 1983

Porcelli, Anthony, Lewis, Andrea, Delgado, Mauricio, Acute Stress Influences Neural Circuits of Reward Processing, in: Frontiers in Neuroscience, vol.6, 2012, article 157

Quinlan, Robert, Human Parental Effort and Environmental Risk, Proceeding of the Royal Society, vol. 274, 2007, pp. 1606-1636

Rawson, Benjamin, Children and Childhood in Roman Italy, Oxford 2003

Schwabe, Lars, Wolf, Oliver, Stress Prompts Habit Behavior in Humans, in: The Journal of Neuroscience, vol. 29 No.3, 2009, pp. 7191-7198

Smith and Fretwell, The Optimal Balance between Size and Number of Offspring, in Proceedings of the Royal Society B, 1974, pp. 1152-1167

Stevens, Emily et al., A History of Infant Feeding, in: The Journal of Perinatal Education, vol. 18, No. 2, 2009, pp. 32-39

Stone, Lawrence, The Family, Sex and Marriage in England 1500-1800, New York, 1977

Struve, Christian August, Anleitung zur Kenntnis und Impfung der Kuhpocken, Breslau, 1802

Tamas Bereczkei, Maternal trade-off in treating high-risk children, Evolution and Human Behavior, Volume 22, Issue 3, May 2001, pp. 197-212

UNICEF, WHO, World Bank UN DESA/Population Division (ed.) Levels and Trends in Child Mortality, Report 2014, Data Appendix

Voigtländer, Nico, **Voth**, Hans Joachim, The three horsemen of riches: Plague, war, and urbanization in early modern Europe, in: The Review of Economic Studies, vol. 80 No.2 2013, pp. 774-811.

Wolff, Eberhard, Einschneidende Maßnahmen, Stuttgart 1998

Wood, James, Dynamics of Human Reproduction, New Brunswick, 2009

Worthman, Carol, Evolutionary Perspectives on the Onset of Puberty, in: Threvathan, Wenda, Evolutionary Medicine, Oxford 1999, p. 135-165

Young, Alwyn, The Gift of the Dying: The Tragedy of Aids and the Welfare of Future African Generations, in: The Quaterly journal of economics, vol. 120 no. 2, 2005, pp. 423-466.

Zeidner Moshe, Endler, Norman, Handbook of Coping: Theory and Applications, New York, 1996