JFEA

Managing Refugee Protection Crises: Policy Lessons from Economics and Political Science

Dominik Hangartner
Center for Comparative and International Studies, ETH Zurich
Immigration Policy Lab, Stanford University and ETH Zurich
dominik.hangartner@gess.ethz.ch

Matti Sarvimäki
Dept. of Economics, Aalto University School of Business
VATT Institute for Economic Research
Helsinki Graduate School of Economics
matti.sarvimaki@aalto.fi

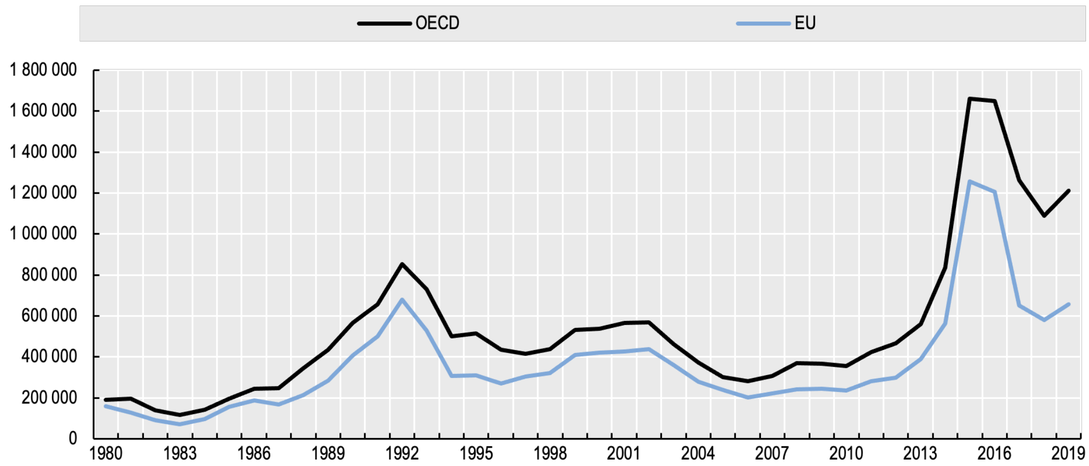
Judith Spirig
Dept. of Political Science, University College London
Immigration Policy Lab, Stanford University and ETH Zurich
j.spirig@ucl.ac.uk

Abstract

We review and interpret research on the economic and political effects of receiving asylum seekers and refugees in developed countries, with a particular focus on the 2015 European refugee protection crisis and its aftermath. In the first part of the paper, we examine the consequences of receiving asylum seekers and refugees and identify two main findings. First, the reception of refugees is unlikely to generate large direct economic effects. Both labor market and fiscal consequences for host countries are likely to be relatively modest. Second, however, the broader political processes accompanying the reception and integration of refugees may give rise to indirect yet larger economic effects. Specifically, a growing body of work suggests that the arrival of asylum seekers and refugees can fuel the rise of anti-immigrant populist parties, which may lead to the adoption of economically and politically isolationist policies. Yet, these political effects are not inevitable and occur only under certain conditions. In the second part of the paper, we discuss the conditions under which these effects are less likely to occur. We argue that refugees' effective integration along relevant linguistic, economic, and legal dimensions, an allocation of asylum seekers that is perceived as 'fair' by the host society, and meaningful contact between locals and newly arrived refugees have the potential to mitigate the political and indirect economic risks.

This paper is based on a background report commissioned by Finland's Economic Policy Council and circulated as "Dealing with the Refugee Crisis: Policy Lessons from Economics and Political Science" in 2017. We thank Janne Tukiainen (the editor), an anonymous referee, Roope Uusitalo, and seminar participants at Bank of Finland, Ministry of Finance, and Prime Minister's Office for insightful comments. Sary Alasiri supported the compilation of the literature review.

Figure 1: New asylum applications since 1980 in the OECD and the EU



Source: OECD (2020)

1. Introduction

In 2015, the number of individuals seeking asylum in Europe increased rapidly. For many governments, the situation posed a major challenge, because they quickly had to find ways to host asylum seekers and process their asylum claims. The "European Refugee Protection Crisis" of 2015 also pushed many countries to introduce more restrictive asylum and new integration policies.[1] By and large, these decisions were made on an ad-hoc basis, with limited foresight on their likely impacts and a small evidence base.

While 2015 was certainly exceptional in comparison to previous years, relatively sudden increases in asylum applications and refugee reception are quite common, also in Europe. As shown in Figure 1, Europe experienced the arrival of comparatively large numbers of asylum seekers also during the 1990s. Earlier refugee crises not covered in this figure include, among others, the millions of displaced Europeans in the aftermath of World War II.

Given that demand-side drivers, most notably wars and violent conflict (see Hatton 2004), account for much of the variation in asylum application numbers, it is likely that the number of people seeking asylum will remain outside host countries' complete control and keep fluctuating. Accordingly, while governments' reception and integration challenges might have been exceptionally large in the aftermath of the 2015 refugee protection crisis, similar situations are likely to arise also in the future. Thus, it is a good time to take stock on lessons learned from the past crisis and prepare for future challenges.

[1] *Initially, some policies made seeking asylum easier. Most notably, for a short period, Germany stopped sending Syrian refugees back to the country of first entry and started invoking the sovereignty clause of the Dublin regulation, i.e. to evaluate Syrian asylum seekers' claims in Germany. Subsequent changes in asylum policies had been almost exclusively restrictive and aimed to reduce the number of (staying) asylum seekers – including Germany retracting from the invocation of the sovereignty clause and Sweden no longer offering permanent residency permits upon arrival. Most countries have also introduced new integration policies focusing on asylum seekers with a high probability of staying or already accepted refugees with the main goal to facilitate their labor market integration. This mainly included language trainings and integration courses. For example, Germany started obliging accepted refugees and asylum seekers with a high probability of staying to participate in integration classes. Sweden has increased the availability of vocational introduction jobs and work experience placements for asylum seekers and expanded the range of services in household work that allow for tax deductions (see Swedish Government 2015).*

This paper aims to help both researchers and policy makers in this task. We present a selective review and an interpretation of the research examining the economic and political effects of receiving refugees in Europe and North America. We focus predominantly on *published* studies of the political and economic consequences of the arrival of *forced migrants* (i.e., asylum seekers, refugees, people with subsidiary protection)[2] in *Europe using credible research designs* to identify causal effects. However, two notes are in order. First, where necessary, we also include studies looking at the effects of immigration more broadly, working papers, and expand our scope to other countries (e.g., the U.S.). Second, given that this is a rapidly developing literature, we have no claim to completeness.

We start with a short review of research on labor market and fiscal impacts of refugee immigration. Based on this literature, two main findings emerge. First, the average impact of refugees on native wages or employment is likely to be small. Second, while natives may be affected through public finances, the fiscal effects of receiving refugees are likely to remain quite limited even under pessimistic assumptions. We substantiate this argument by discussing the sources and the extent of uncertainties embedded in any estimate of the long-term fiscal impacts.

We then proceed with a review of research on the political consequences of asylum immigration. Our main argument is that receiving refugees may have more important economic effects through the broader political process. That is, increasing arrivals of asylum seekers often fuel anti-immigrant attitudes and the rise of authoritarian, populist, anti-immigrant parties. Often, these parties not only promote anti-immigrant policies, but tend to support isolationist policies more broadly (e.g., withdrawal from trade agreements or the European Union). In addition, in response to the increasing popularity of such parties, also mainstream parties tend to shift their policy positions towards isolationism. While it is hard to quantify how these political changes will affect policy – or what the economic effects of these policy changes are – we argue that this channel is likely to pose a larger risk for economic effects than any conceivable labor market or fiscal effects of receiving refugees.

The risk of rising right-wing populism leaves non-isolationist policy makers with the dilemma of how to respond. In order to inform policy decisions, we present a review of the relevant economics and political science research on the effects of asylum and integration policy responses. We argue that effective reception and integration policies both promote refugees' integration outcomes and decrease the potential of an isolationist backlash.

We first survey the literature on voter preferences about refugee policy. This research suggests that across European countries, the majority of residents have a skeptical view of current asylum policies and prefer to curb future refugee arrivals. However, they also tend to support the acceptance of refugees deserving of asylum (according to the Refugee Convention) and are willing to accept more asylum seekers as long as the allocation is 'fair' and proportional to the country's capacity.

We next discuss lessons from the existing literature on how to organize the reception of asylum seekers in a way that minimizes the potential for anti-immigrant backlash. Our suggestions include providing timely information to locals; facilitating repeated and meaningful contact between locals and newly arrived refugees; ensuring that the allocation of asylum seekers within countries is perceived as 'fair'; and taking into account that rural and more conservative constituencies tend to exhibit stronger backlash against refugees.

Finally, we review impact evaluations of (broadly defined) integration policies. This body of research suggests four lessons: interventions improving the match quality between active labor market policy measures and the individual characteristics of each refugee have large effects; long waiting periods, such as those arising from lengthy asylum process, reduce later employment prospects; temporary employment bans for newly arrived asylum-seekers, such as those currently enacted by most European countries, have a detrimental impact on the long-term economic integration of refugees; and proficiency in local language is strongly associated with labor market success and there are good reasons to think that at least part of this association reflects a causal effect of language skills on labor market success.

[2] *Please note that we mainly use the terms "asylum seekers" and "refugees" to refer to forced migrants who have arrived in European countries to seek asylum. The terms are used interchangeably.*

In sum, while aggregate labor market consequences appear small, both potential fiscal and political effects of a sudden increase in the number of asylum seekers are not inevitable. Our overview of the insights accumulated from earlier experiences suggest that a swift integration of refugees, an allocation of asylum seekers perceived to be fair, and policies that encourage meaningful interaction between refugees and host communities have the potential to mitigate the fiscal and political risks associated with large-scale refugee arrival – in addition to their direct effects.

2. Labor market and fiscal effects

This section discusses two channels through which the reception of refugees could have direct effects on the host country's economy. First, refugees could compete with natives in the labor market and thus drive down wages or employment. Second, refugees pay taxes, receive transfers and use public services, and may thus create fiscal effects. We present a short review on the literature examining these effects and conclude that both effects are likely to be relatively small.

A. Impact on native wages and employment

A simple economics textbook model provides a useful starting point for thinking about the labor market consequences of immigration (see, e.g., Borjas 2015 for discussion). In these models, immigration affects native wages only if the skill-mix of immigrants differs from that of natives. If the distribution of skills is identical among immigrants and natives, the arrival of immigrants should increase the size of the economy, but have no impact on long-run native wages or employment. Sudden and large immigration does affect the capital-labor ratio and temporarily lowers wages, but over time capital adjusts and wages revert back to their pre-immigration level. However, if the skill-mix of immigrants and natives differs, some natives will win, and others will lose. More precisely, those who have *complementary* skills in comparison to immigrants will become more productive and can thus demand higher wages.[3] Symmetrically, natives who are *substitutes* to immigrants in terms of their skills will become less productive. Thus, the labor market effects crucially depend on the extent to which immigrants and natives are substitutes vs. complements in the labor market.

A large empirical literature has examined the impact of immigration on natives' wages and employment by comparing natives working in labor markets differentially exposed to immigration. The key challenge of this research is that immigrants are not randomly allocated into labor markets.[4] Thus natives working in labor markets with many immigrants are likely to differ from natives working in labor market with few immigrants also in ways that have nothing to do with immigration.

This identification challenge has led part of the literature to focus on quasi-experimental research designs where some labor markets experience the arrival of immigrants for reasons that are plausibly exogenous to the (unobserved) characteristics of these labor markets. Given that a large number of high-quality reviews on this topic already exists, we do not attempt to provide another one here, but rather refer the reader to Borjas (1999), Hanson (2009), Blau and Mackie (2016) and Dustmann, Schönberg and Stuhler (2016).

Our reading of this literature is that the impact of typical refugee flows on native wages and employment is likely to be modest. This conclusion is partly driven by the relatively small estimates for labor market effects even in situations where the number of immigrants is large (see the reviews cited above for details). On the other hand, in most cases of refugee inflows that create a substantial political reaction, the number of refugees tends to be relatively small compared to the

[3] *The discussion on the expected impact of immigration on native wages is almost exclusively conducted using models of competitive labor markets. In recent work, Amior and Manning (2021) point out that if employers enjoy greater market power over migrant than native labor, immigration will allow them to extract greater rents. As a consequence, immigration could simultaneously increase natives' productivity and reduce their wages.*

[4] *When labor markets are defined as geographical units, the resulting estimates are likely to be biased upwards, because immigrants tend to move to booming areas. On the other hand, when labor markets are defined by occupations, estimates are likely to be downward biased, because immigrants often work in low-wage jobs.*

size of the labor market.[5] Finally, refugees typically struggle to find employment in the host country's labor market, which further reduces the competitive pressure they exert on natives. Thus, it seems unlikely that labor market effects would be the primary channel for refugees to affect natives' economic well-being.

B. Impact on public finances

The most important direct economic effect of receiving refugees is likely to occur through public finances. Refugees tend to have lower employment rates than other immigrant groups or natives throughout Europe (Brell, Dustmann, and Preston 2020, Fasani, Frattini, and Minale 2021). As a consequence, they tend to receive substantially more social transfers and to pay less taxes than natives or other immigrants. These differences give rise to dramatic differences in some forms of transfers. Ruist (2015) documents the situation in Sweden in 2017, where refugees accounted for 5.1% of the population in 2007 (the largest refugee population share in Europe) and 55% of social assistance spending. On the other hand, public spending on refugees' pensions, health and education was much lower than that of natives. As a consequence, refugees accounted for 5.6% of total public spending, i.e., quite close to their population share. Due to their low employment rate, however, refugees paid less taxes than they received benefits and used public services and thus created a net fiscal cost. In total, one percent of Swedish GDP was redistributed to the accumulated refugee population (including refugees' family members). For comparison, Sweden's foreign aid budget is roughly of the same size (OECD 2013).

While such cross-sectional observations clearly contain information, it is unclear how well they capture the long-term fiscal impacts of immigration. The reason is that the net cost or surplus that an individual creates for the public sector varies dramatically over her lifecycle. Everyone is a net burden during childhood and adolescence. Furthermore, most people create large net costs to the public sector during their last years of life. In order to truly capture the fiscal impacts of immigration, these dynamics would have to be taken into account. That is, the appropriate measure for fiscal impacts is the discounted sum of all future taxes, transfers and costs due to public service consumption (Lee and Miller 2000, Storesletten 2000, Ruist 2020).

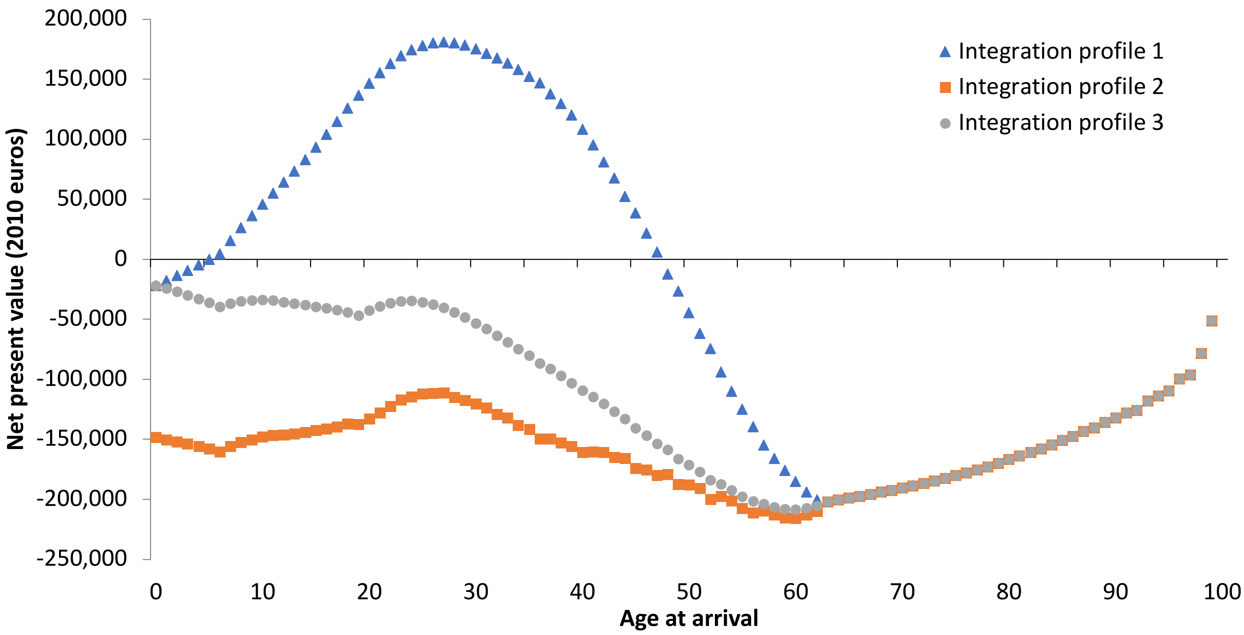
The challenge in incorporating lifecycle dynamics into the estimates of the fiscal impact of immigration is that much of the costs and benefits will take place in the future. Thus, researchers have to make strong assumptions about the future labor market performance of immigrants, the cost of providing public services, the structure of the tax and benefit system, overall economic growth and so forth. Clearly, our ability to forecast these factors for the next decades is very limited. Accordingly, all estimates of the long-term fiscal effects of immigration are best understood as scenario exercises.

Figure 2 illustrates this issue using data from Finland. Each point in the figure refers to the net present value of the fiscal impact of an additional immigrant arriving in Finland as a function of age at arrival (x-axis) and future labor market performance (marker style).[6] The top series correspond to a scenario where the immigrant immediately starts to follow the average profile of natives. That is, in this scenario, a person arriving at age 30 is assigned the average taxes, transfers and cost of public services of current 30-year-old natives. In the next year, she is assigned the averages of current 31-year-old natives and so forth. Furthermore, her offspring is also assigned the age profiles corresponding to current natives. Fertility and mortality are assumed to follow the age profiles observed in current data and pensions are estimated within the model using the earnings profiles of current natives. The future costs and taxes are then discounted to net present value using 3% discount rate and assuming 1% annual earnings growth.

[5] *Of course, there have been cases of extraordinarily large refugee flows that may have had meaningful labor market effects (see, e.g., Borjas and Monras, 2017). We also recognize that the literature on the labor market effects of immigration is evolving and remains somewhat contested.*

[6] *The details are presented in Sarvimäki et al. (2014) and its online appendix available at www.vatt.fi/maahanmuutto.*

Figure 2: Scenarios of the long-term fiscal impact of immigration from Sarvimäki et al. (2014). See Section 3.2 for discussion.



The top scenario highlights the importance of age at arrival in the (unrealistic) situation where immigrants enter the labor markets immediately at the level of natives. Those arriving after age 50, and those arriving before age six constitute a net burden to the taxpayer. Newborn natives, i.e., those "arriving" at age zero, have a negative net present value in these scenarios. (This observation alone highlights the fact that these scenarios do not aim to provide credible predictions.) In contrast, immigrants arriving as young adults spend a long period working and paying taxes, but the expenses due to their education and health care during childhood is paid somewhere else. Thus, they make a large positive contribution to public finances if they integrate immediately into the labor markets.

Thus far, however, immigrants – and refugees in particular – have experienced difficulties in finding stable employment in Finland (Sarvimäki 2011, 2017). The bottom scenarios illustrate this fact. Immigrants and their offspring are now assigned the observed age profile of immigrants (including non-refugees) in 1995–2012, who arrived in Finland during the 1990s. The net present value on public finances is now negative regardless of the age at arrival and varies between €110,00 and €150,000.

The difference between the top and bottom scenarios highlights the importance of labor market integration in generating the fiscal impact of immigration. Among those arriving to Finland between the ages of 20 and 40, the difference between the two scenarios is almost €300,000 in net present value. One way to interpret these results is that if one could design integration training that would help moving immigrants from the bottom to the top scenario, such program would be cost-efficient even with a cost of €300,000 per participant. In comparison, the average investment in training for immigrants who participate in integration programs in Finland has been around €15,000 (Sarvimäki and Hämäläinen 2016). Of course, designing integration programs that actually move immigrants from the bottom to the top scenarios may not be feasible. Nevertheless, a comparison of these alternative scenarios illustrates that effective integration programs can have substantial fiscal impacts.

The third set of scenarios, presented in the middle of Figure 2, correspond to an integration profile, where immigrants follow the profiles of the 1990s immigrants, but their children follow the profiles of current Finnish natives. The difference between these scenarios and bottom scenarios – roughly €75,000 for immigrants arriving in their mid-20s – illustrates the importance of the integration of the children of immigrants to the host country's labor markets.

We stress that none of the scenarios presented in Figure 2 should be interpreted as "price-tags" of immigration. Clearly, Finland's tax and benefit systems as well as the way public services are provided will change in the future. Other assumptions embedded in these scenarios, such as a steady 1% annual earnings growth, are also unlikely to be accurate predictions of the future. However, scenarios such as those presented in Figure 2 can help to clarify which factors matter the most and thus assist policy makers in focusing on these factors. In particular, Figure 2 shows that increasing employment through efficient integration policies has potential to create large benefits purely from a public finances viewpoint.

3. Political effects

We now turn to the literature on refugee arrivals' political effects in the receiving countries. We first present a relatively detailed review of the recent literature on the impact on votes and attitudes. At the end of the section, we also briefly discuss the existing evidence on the tendency of populist parties to support other isolationist policies and their influence on the policy stances of mainstream parties.

A. First-order effects: attitudes and votes

There are two main theories as to how individuals respond to the presence of a 'foreign' group of people. One the one hand, the 'contact hypothesis theory' (Allport 1954) suggests that interaction between natives and immigrants (majority and minority group) can reduce negative attitudes of majority group members toward the minority group and – under some circumstances – tackle xenophobic fears. On the other hand, 'group threat theory' (see, e.g., Blumer 1998; Quillian 1995) suggests that it is actually the presence of immigrants that causes and exacerbates such animus.

Empirically, it is challenging to credibly document the impact of immigrants on their host societies, because immigrants typically make residential choices based on private information about their destinations, such as local labor market conditions or xenophobic attitudes. Since these local conditions or attitudes are likely to also influence support for right-wing parties, refugee self-selection poses a serious challenge to causal inferences about the effect of immigration (see Dustmann and Preston 2001). In other words, it is inherently difficult to distinguish between co-occurrence of refugees and local attitudes towards immigration among natives and the isolated effect of an increase in refugee migration on native citizens' attitudes and political behavior. We believe that the endogeneity issues are potentially severe, which is one reason why we largely focus on studies that employ research designs that can credibly claim to identify causal effects.

Another challenge has to do with the fact that citizens experience immigration through different channels. Immigration, and newly arriving asylum seekers especially, are one of the most salient topics in the media. Almost on a daily basis, newspapers report on asylum-related issues (see, e.g., Spirig 2021). At the same time, experiences with refugees can also take place on the micro-level. Citizens who live in a community that hosts asylum seekers might see and/or meet asylum seekers on a daily basis. As Hopkins' (2010) study on anti-immigrant attitudes in the U.S. shows, these channels are not independent from each other. He presents evidence that "at times when rhetoric related to immigrants is highly salient nationally, those witnessing influxes of immigrants locally will find it easier to draw political conclusions from their experiences" (Hopkins 2010, 44).

This finding points to several contextual factors that could moderate the influence of refugee arrival on political outcomes (via its impact on voters' attitudes). We discuss this in more detail below. In addition, while the macro-level channel might be more impactful on individuals' attitudes this is particularly challenging to empirically substantiate – among other reasons, because of the lack of exogenous variation. This is why we emphasize the importance of factors determining citizens' macro-level experience of the arrival of asylum seekers – in particular the salience and framing of the issue in the media and political debates – but focus our review mainly on studies exploring regional differences in right-wing, anti-immigrant party support within states, rather than between states. This focus on regional variation in all likelihood underestimates the total (regional and national) effect of immigration on right-wing, anti-immigrant parties (see, e.g., Steinmayr 2021).

Empirical studies that look at the micro-level and are able to tackle these inferential challenges do not provide a uniform answer as to what the direct effect of immigration on political outcomes is. The results spread from a significant negative effect of accommodating asylum seekers on right-wing party support (Steinmayr 2021), to a strong positive effect of experiencing asylum seeker arrivals on extreme-right party support (Dinas et al. 2019; Steinmayr 2021). The following section summarizes this range of studies.

The recent study by Steinmayr (2021) is one of the first to find an overall negative effect of accommodating asylum seekers on (an increase in) far-right party support: Upper Austrian municipalities that accommodate asylum seekers display an about 3.9 percentage point smaller increase in the 2015 state election vote share for the anti-immigrant FPÖ than municipalities that did not host refugees. According to the study, qualitative interviews imply that "in almost all cases, the level of anxiety declined after the asylum seekers had been there for some time since most of the feared consequences did not materialize" (Steinmayr 2021, 321). In new working papers on the effect of refugee reception centers in Italy and France, Gamalerio et al. (2021) and Vertier, Viskanic, and Gamalerio (2020), respectively, provide additional evidence that when meanigful contact between locals and refugees is promoted and possible (i.e., the number of refugees is small), the presence of refugee centers leads to a reduction in right-wing party support.

Two studies, one focusing on the effects of asylum seeker arrivals on support for the right-wing AfD in Eastern Germany (Schaub, Gereke and Baldassarri 2021) and one focusing on refugee arrival in Danish municipalities (Jensen 2020), document null effects on party vote shares. In addition, Savolainen (2016) examines electoral effects of opening asylum centers in Finnish municipalities in 1990-2011, and finds neither evidence for an impact on anti-immigration nor pro-immigration parties' vote share.[7]

A relatively large set of studies documents positive effects of asylum seeker arrivals on support for anti-immigrant parties. Bratti et al. (2020) show that in Italy, proximity to refugee centers that do not provide integration services increases support for populist parties. Dustmann, Vasiljeva, and Damm (2019) focus on refugees and find that not only votes for the right-wing party, but also the center-right party increase with larger shares of refugees being allocated to a municipality. An increase in the municipality's refugee share by one percentage point leads to a 1.2 percentage point increase in the anti-immigration party's vote share in parliamentary elections, and also the center-right party benefits. Furthermore, they document that anti-immigrant parties are more likely to run in municipal elections when refugee shares are higher and that this relationship does not hold in big urban areas.[8] Hangartner et al. (2019) and Dinas et al. (2019) examine the effect of asylum seekers passing through Greek islands close to Turkey. These two papers find that the passing-through of asylum seekers, who do not stay but continue onwards with their journeys, leads to "lasting increases in natives' hostility toward refugees, immigrants, and Muslim minorities; support for restrictive asylum and immigration policies; and political engagement to effect such exclusionary policies" (Hangartner et al. 2019: 442) and, in the short run, an increase of two percentage points (more than 40% at the mean) in the vote share of *Golden Dawn*, arguably the most extreme right-wing and anti-immigrant party holding office in Europe (Dinas et al. 2019). Steinmayr (2021) relatedly finds that support for the far-right Austrian FPÖ increased about 1.5 percentage points more in municipalities that asylum seekers passed through and did not stay as opposed to comparable municipalities that did not see asylum seekers pass through, and Gessler, Tóth and Wachs (2021) document a similar effect for Hungary.

---

[7] *However, relatively few asylum centers were established during the study period and thereby raising concerns that the study might be underpowered.*

[8] *Barone et al. (2016) find the same result with regard to cities in Italy.*

Taken together, these studies indicate that the (short-term) effect of the presence of newly arrived asylum seekers on electoral outcomes and support for far-right parties is not only theoretically, but also empirically unclear. The most natural interpretation given the different contexts under study is that the effects depend strongly on moderating factors, such as the facilitation of inter-group contact (see, e.g., Steinmayr 2021), the size of the refugee arrivals (see, e.g., Dinas et al. 2019), pre-existing political attitudes (see, e.g., Dustmann, Vasiljeva, and Damm 2019), and macro-level determinants, such as issue salience (see., e.g., Hopkins 2010). Yet, there is a lack of research as to which moderating factors are most crucial when it comes to political consequences of refugee immigration. To gauge which moderating factors are the most relevant, however, we briefly consider a larger set of studies investigating the effect of immigration (in different forms) on political outcomes in the following.

While there are many studies looking at the effect of (refugee) migration on electoral outcomes, there is very limited attention to other outcomes such as native citizens' preferences for redistribution or trust in political institutions. McLaren (2012, 2015) is one of the few exceptions, however. Her research focuses on the impact of anti-immigrant attitudes on political trust and argues that some voters feel that immigration threatens a sense of national identity that lies at the heart of the liberal state and/or want to hold the state accountable for 'failing' to control immigration adequately. Accordingly, immigration not only fuels anti-immigrant attitudes,[9] but might also lead to a decrease in political trust: "politicians and institutions are likely to be blamed for failing to control immigration adequately" (2012, 171). In a similar vein, a few studies focus on the effect of ethnic diversity on social trust (for an overview, see Dinesen, Schaeffer, and Sønderskov, 2020) and attitudes regarding welfare state spending (Dahlberg, Edmark, and Lundqvist 2012). Both of these papers find a negative effect of ethnic diversity on mentioned outcomes. Dahlberg, Edmark, and Lundqvist (2012) investigate the effect of the arrival of refugees in assigned Swedish municipalities on voters' responses to a survey on welfare state spending and find evidence for the so-called 'in-group bias': Individuals display lower preferences for redistribution if the share of refugees placed in their municipality is larger.[10]

The literature dealing with the impact of immigration – studies focusing on asylum sekers are rare – on more extreme forms of political behavior, such as political violence, does not provide a uniform answer, either. Dancygier (2010) investigates the impact of immigration on violent attacks and documents a positive relationship between the two in Greater London. Braun and Koopmans (2010) find similar effects in Germany, as do Krueger and Pischke (1997) in the German Democratic Republic, but not in Western Germany. Accordingly, Krueger and Pischke (1997) note that local political processes – how local conflicts are handled – play a large role: Immigration and violent outbursts are positively correlated, if local political processes facilitate mobilization. A qualitative study by Karapin (2002) also suggests that whether immigration by ethnic minorities led to violent, anti-immigrant riots in 1990s Germany depended on local political processes, such as, among others, facilitation of non-violent political participation. Analyzing more recent data from Germany, Marbach and Ropers (2018) find that in times when immigration is salient on the national level, increases in the number of asylum seekers at the local level are associated with more anti-asylum seeker violence on the local level. Falk, Kuhn and Zweimüller (2011), however, find very little, or no impact of the size of the immigrant community on political violence.

In sum, there is a range of studies employing credible research designs that suggest that immigration leads to political shifts: More immigration appears to lead to increases in votes for anti-immigration, and typically right-wing, parties, more political violence directed at immigrants, and potentially lower levels of political trust and preferences for redistribution. Accordingly, this research sheds light on potential political repercussions of receiving asylum seekers. Rises in right-wing authoritarian attitudes, erosion of trust in political institutions and democratic governance, and electoral success of extreme-right parties (such as, for example, in Greece) have the potential to fundamentally affect the civic fabric of a society and undermine the credibility of its political system. This democratic backsliding could prove to be much more substantial than the short-term negative economic and fiscal consequences of an increase in the number of arriving asylum

[9] *Like, for example, in the UK, where immigration was not dominating elections in the past because major parties did not found their campaigns on it.*

[10] *Nekby and Pettersson-Lidbom (2016) question the validity of Dahlberg, Edmark, and Lundqvist's (2012) identification strategy and the representativeness of their sample (see also Dahlberg, Edmark, and Lundqvist's (2016) reply).*

seekers. However, it is crucial to note that taken together, the studies discussed above also imply that the effect of asylum immigration on political outcomes is not deterministic. The consequences appear to be highly dependent on the political context and policy choices. While we gauge the potentially substantial second-order effects below, Section 4 will discuss which policies are most likely to reduce the political effects of refugee protection crises.

B. Second-order effects

In contrast to asylum immigration's first-order effects on (far-)right party support, there is much less research on its second order-effects: How does the immigration-fueled rise of anti-immigrant parties affect isolationist policies and the political platform of mainstream parties? The existing literature suggests a twofold answer. First, anti-immigrant parties (e.g., the French Rassemblement National or the U.K.'s Brexit Party) or candidates (e.g., President Trump) tend to favor isolationist policies such as exiting the European Union, reinstating border controls, and – more recently – taking anti-globalization stances more broadly (see, e.g., Walter 2021), such as curbing international trade. Second, the rise of populist parties may affect policy even if these parties do not enter the government. This view is supported by a related literature in political science examining how successes of far-right parties exert electoral pressures on mainstream parties, and are thereby able to shift government parties' policy positions closer to their ideal point (see, e.g., Abou-Chadi and Krause (2020), Spoon and Klüver (2020), as well as the review article by Golder (2016) and the references therein).

A prominent example of such second-order effects of asylum salience is the Brexit referendum. According to Moore and Ramsey's (2017) analysis of all articles published by leading U.K. news outlets during the 2016 EU referendum campaign, immigration, and in particular asylum migration, was the most prominent campaign issue (based on the number of times it led newspaper print front pages), with almost 80% of them appearing in Leave-supporting newspapers. In contrast, economic issues and, in particular, the vexing question of the impact of leaving the Single Market, received significantly less frontpage attention. While assessing the impact of the UK leaving the EU is extremely challenging, it seems likely that the significance of refugees for the U.K. economy is miniscule, and the connection to the question of EU membership tenuous, especially in comparison to the importance of accessing the European Single Market. This focus on the issue of asylum migration threatened to crowd out attention on the many other, and arguably more consequential, legal and economic issues tied to Brexit.

4. How to mitigate the fiscal and political risks?

The research reviewed above illustrates the gravity of challenges and constraints policy makers face when their countries receive refugees and asylum seekers. Thus far, we have provided little guidance on how to respond to these challenges. In this section, we review research that we hope will help decision makers to design appropriate policy responses. We first provide an overview of voter preferences and proceed with a relatively thorough review of the available evidence on the impacts of alternative reception and allocation of asylum seekers, integration policies and language training as well as eventually giving refugees permanent residence permits and citizenship.

A. Voter preferences

We first survey the literature on voter preferences about refugee policy. This research suggests that across all Western countries, the majority of residents have a skeptical view of current asylum policies and prefer to curb the number of future asylum seekers. However, they also tend to support the acceptance of refugees deserving of asylum and are willing to accept more asylum-seekers as long as the allocation is 'fair' and proportional to the country's capacity.

The increase in asylum applications and the differences across European countries have not gone unnoted. Asylum law experts and policy-makers alike have repeatedly expressed concern about the unequal distribution of asylum seekers across Europe, the Dublin regulation, and how the EU is handling the increasing pressure at its borders (see, e.g., Thielemann 2010; Angenendt, Engler, and Schneider 2013; Malmström 2014). Furthermore, citizens across Europe share the impression that the Dublin system is unfair. Employing an online survey experiment involving 18,000 voters across fifteen European countries, Bansak, Hainmueller and Hangartner (2017) find that only 18 percent of respondents support the current Dublin regulations, which state that asylum seekers usually have to submit their claim in the European country of first entry. Interestingly, the support is very low even in countries that benefit from the current status quo in the sense that they receive relatively few asylum claims. In stark contrast, 70 percent of respondents prefer proportional allocation of asylum seekers based on the country's capacity (a function of population size, GDP, unemployment rate, and number of past applications). When voters are randomly prompted with the actual numbers of asylum seekers their country would receive under each allocation, they are somewhat more likely to support the allocation that yields the lowest number of asylum seekers for their own country. However, even under this treatment condition, in all but three countries (Czech Republic, Poland, and the UK) a majority of voters prefers proportional allocation over the status quo. (Note that ten out of the fifteen countries would have to host more asylum seekers under proportional allocation.) These findings indicate that a majority of citizens is willing to provide refuge to additional asylum seekers as long as they know that the overall allocation across 'Dublin countries' is proportional to a country's capacity.

In a companion paper based on the same survey, Bansak, Hainmueller and Hangartner (2016) employ a conjoint analysis asking the 18,000 respondents to evaluate fictitious profiles of asylum seekers that randomly varied along personal attributes. They find that asylum seekers who are highly skilled, contribute to the host country's economy, have more consistent asylum testimonies and severe vulnerabilities, and are Christian rather than Muslim receive the greatest public support. Bansak, Hainmueller and Hangartner (2016) argue that these results point to tough challenges for policy makers who are struggling to meet their legal responsibilities to protect refugees in line with the 1951 Refugee Convention. The public's strong anti-Muslim bias and preference for highly skilled asylum seekers who can speak the language of the host country hinder the acceptance and integration of asylum seekers given that most currently originate from Muslim-majority countries and may lack the desired professional and language skills. At the same time, Bansak, Hainmueller and Hangartner (2016) argue that the findings also point to opportunities for policy makers: the fact that citizens across Europe share common humanitarian concerns for refugees with consistent asylum claims suggests that large segments of the public have at least partially internalized the central pillars of international refugee law.

B. Reception and allocation of asylum seekers

We next discuss lessons from the existing literature on how to organize the reception of asylum seekers in a way that minimizes the potential for anti-immigrant backlash. Our suggestions include providing timely information to locals; facilitating repeated and meaningful contact between locals and newly arrived refugees; ensuring that the allocation of asylum seekers across and within countries is perceived as 'fair'; and taking into account that rural and more conservative constituencies tend to exhibit stronger backlash against refugees.

First, a distribution of asylum seekers across receiving communities that is perceived as fair by voters increases support for reception. As discussed in detail above, Bansak, Hainmueller and Hangartner's (2017) cross-country survey that examines Europeans' attitudes towards different asylum seeker allocation mechanisms shows that most citizens in Europe prefer a proportional allocation of asylum seekers over the status quo of country of first entry.

Second, large, liberal municipalities (big urban areas, cities) appear to exhibit smaller (attitudinal and) electoral responses to increases in asylum immigration. Right-wing, anti-immigrant parties usually do not receive their largest support from cities (see Dustmann, Vasiljeva, and Damm 2019). In addition, both Barone et al. (2016) and Dustmann, Vasiljeva, and Damm (2019) show that in large urban areas, immigration does not increase right-wing party support in Italy and Denmark, respectively. Accordingly, it seems worth to take this strong treatment effect heterogeneity into consideration when deciding where to host asylum seekers.

Third, macro-level salience matters. Therefore, information campaigns and narratives prior to the arrival of asylum seekers could help mitigate initial backlashes. Hopkins' (2010) study on attitudinal changes of Americans in response to immigration highlights the importance of the macro-level narrative on immigration and how that might influence and structure citizens' perception of face-to-face encounters with refugees in their community.

C. Length of the asylum process

Once the asylum seekers have been received and allocated to asylum centers, policy makers need to decide how much resources to use for their applications to be processed. The available resources, in turn, largely determine how fast the applications can be processed. Hainmueller, Hangartner and Lawrence (2016) show that the processing time affects subsequent integration of refugees into the host society. More specifically, they provide evidence as to how the length of time that refugees 'wait in limbo' for a decision on their asylum claim impacts on their subsequent economic integration. Exploiting exogenous variation in wait times and using registry panel data covering asylum seekers who had applied for asylum in Switzerland between 1994–2004, they find that one additional year of waiting reduces the subsequent employment rate by 4 to 5 percentage points, a 16% to 23% drop compared to the average rate. This deleterious effect is remarkably stable across different subgroups of refugees stratified by gender, origin, and age at arrival.

The findings of Hainmueller, Hangartner and Lawrence (2016) are consistent with previous cross-sectional and qualitative evidence (Stepick and Portes 1986; Waxman 2001; Bakker, Dagevos and Engbersen 2014) suggesting that waiting in limbo dampens refugee employment through psychological discouragement, rather than a skill atrophy mechanism. In other words, whereas recent reductions in refugees' labor market access waiting times point to the importance of early labor market access for economic integration, Hainmueller, Hangartner and Lawrence (2016) highlight an additional factor that affects asylum seekers' ability to integrate: the degree of, and time period, in uncertainty about the future. Their partial equilibrium cost benefit analysis suggests that even policy reforms marginally reducing the waiting period for asylum seekers would help refugees to navigate the difficult transition from a life in legal limbo to a successful integration into the host community better. Moreover, from a host country perspective, such reforms would reduce public expenditures for welfare benefits significantly due to the increase in employment and the resulting increase in tax contributions of employed refugees.

D. Integration programs and active labor market policies

Many countries provide integration programs for refugees and asylum seekers if their applications are approved. These integration programs are often arranged as part of active labor market policies and are offered also to other unemployed immigrants. We next review research based on plausible research designs that aim to evaluate the impacts of these programs. Overall, these studies suggest that integration programs can be remarkably efficient in increasing employment and earnings of refugees (as well as other immigrants struggling to find their way into the host country's labor market).

Andersson Joona and Nekby (2012) study a Swedish program, where newly arrived immigrants were provided extensive counseling and coaching on employment prospects. A trial of this intervention was conducted through Public Employment Service (PES) in 2006–2008. The caseworkers participating in intensive coaching were trained to work exclusively with newly arrived immigrants and would handle less than 20% of the caseload in comparison to regular caseworkers. The intervention aimed to facilitate direct contacts with employers and to improve the match quality between the immigrants and ALPM measures. The evaluation was conducted as a randomized controlled trial (RCT), where PES officers were asked to randomly assign newly arrived immigrants into treatment (intensive coaching) and control (regular introduction programs) groups.

The results suggest that intensive coaching increased the share of immigrants in regular employment by 6 percentage points two years after the start of the intervention. Given that the employment rate of the control group was only 14 percent, this corresponds to 43% increase in employment rate. The effect was sufficient to cover the cost of the program in 2–3 years. Furthermore, the overall effect appeared to be due to men being responsive to the treatment, while no impact was found for women.

An unfortunate feature of this trial is that the treatment and control groups were not fully balanced on pre-assignment characteristics. A possible reason for the unbalance is that the PES officers conducted the randomization and may not have followed the randomization protocol. In particular, they may have attempted to select better participants into the program. If this were the case, the estimated impact of the intervention would overstate the true impact. This hypothesis is supported by the fact that controlling for observed characteristics reduces the estimated employment impact from seven to six percentage points. While this reduction is not particularly dramatic, it suggests that the estimates should be interpreted as upper bounds for the true impact. Furthermore, the results highlight the importance of avoiding situations where parties who might have a stake in the results are responsible for the randomization process.

Åslund and Johanson (2011) examine an earlier Swedish intervention called Special Introduction Programs (SIN), piloted in 2003. It was based on methods originally used for helping workers with disabilities to find employment and focused on immigrants and refugees who were considered to be at risk of becoming long-term unemployed. PES offices executed the intervention and the participating offices were given additional funding to hire caseworkers. These caseworkers had only 10% of the caseload of regular caseworkers, which allowed them to work intensively with each of their clients. The intervention consisted of caseworkers finding suitable jobs for their clients and running an introductory session in these jobs together with the employer, colleagues, and union representatives. This was followed by an internship period lasting up to six months after which the caseworker organized a follow-up session with the aim of turning the internship into a regular job.

The program was piloted in 20 Swedish municipalities in 2003. Åslund and Johanson (2011) evaluate its impacts using a difference-in-differences strategy, where they compare changes in treatment municipalities with changes in non-participating locations in the same local labor market. They find that the intervention increased transitions from unemployment to work experience schemes and improved future employment probabilities for those who entered these schemes.

In a recent contribution, Dahlberg et al. (2020) present evidence from a randomized control trial implemented in Gothenburg, Sweden in 2016–2020. The program targeted newly arrived, low-educated refugees and included language training, supervised work practice, job search assistance, and extended cooperation between the local public sector and firms. The baseline services provided to the control group include all of these elements, but the experimental intervention was much more intensive. For example, the baseline language training provided 15 hours of teacher per week, while those participating in the experimental program received 40 hours per week. The results suggest that these additional investments doubled employment rates during the first year following the program's end (from 15% among to control group to roughly 30% among the treatment group).

In comparison to these rather intensive Swedish programs, the "integration plans" introduced in Finland in 1999 were very light and cheap. These integration plans are prepared in a joint meeting of a caseworker, the immigrant, and an interpreter with the aim to find a sequence of training and other measures that would be the most suitable for each immigrant. In principle, similar meetings took place with all unemployed immigrants already before the reform. However, the integration plans aimed to improve the communication between caseworkers and immigrants. For example, the new guidelines stated that the caseworker had to make sure that the immigrant fully understood the content of her integration plan and knew how to follow it. In addition, the reform aimed to increase the caseworker's capacity to better take into account the specific skills and circumstances of each immigrant.

Sarvimäki and Hämäläinen (2016) evaluate the impact of the integration plans using a research design based on a phase-in rule dictating that participation was mandatory only for unemployed immigrants who had entered the population register after May 1st, 1997. This rule creates a discontinuity, where there is a 35 percentage point difference in the likelihood of receiving an integration plan between those who had arrived on May 1st, 1997 and those who had arrived slightly earlier. Comparing the two otherwise similar groups shows that those who had arrived on May 1st, 1997 and were thus much more likely to receive an integration plan earned cumulatively roughly €7,000 more in the ten-year follow up period than those arriving just before the specified date. Scaling this effect with the change in the likelihood of receiving an integration plan suggests a local average treatment effect of 47% increase in cumulative earnings. Using the same approach, they also find a 13% decrease in the reception of cumulative social benefits. These effects seem to be due to increased language training and other training courses specifically designed for immigrants, replacing more "traditional" active labor market training such as job-seeking courses. That is, there is no detectable impact on the overall amount of training. Furthermore, Pesola and Sarvimäki (2021) find that the integration plans also have large intergenerational effects on grades and educational attainment of the children of the affected immigrants.

Arendt et al. (2021) examine the impacts of another reform on integration policies using a similar empirical approach as Sarvimäki and Hämäläinen (2016). Refugees arriving to Denmark after January 1st, 1999, were required to take substantially more language training than those arriving before the cutoff date. The reform also reduced welfare benefits available for some refugees, altered the way refugees were allocated across municipalities and shifted the responsibility to provide integration training from the central government to municipalities. These changes increased refugees' employment and earnings and facilitated skill-upgrading. In addition, male children of refugees whose both parents arrived just after the threshold date were more likely to complete lower secondary school and committed fewer crimes.

The research discussed in this section shares two central themes. First, all studies examine interventions aimed at improving the match quality between immigrants and training programs. Second, all papers find much larger effects than what is typically documented in the literature on the impacts of active labor market policies on natives' labor market integration (see e.g. Card, Kluve and Weber 2010, 2018 for reviews). These observations are consistent with the hypothesis that refugees (as well as some other immigrant groups) may lack the type of skills that can be improved through training provided by the employment offices, and that they first of all need support navigating the system. Thus, even the very small interventions such as Finland's integration plans can have large effects. These findings also suggest that further policy experimentation on how to improve training and counseling could yield high returns on the public investment.

E. Proficiency of local language

There is little doubt that proficiency in the host country language is crucial for a successful integration. Yet, estimating the causal effect of language on immigrants' integration is challenging because the correlation of language proficiency and labor market outcomes raises a well-founded fear of endogeneity.[11] In addition, the general focus on language *learning* (almost) throughout this literature entails critical challenges for the estimation of the importance of language on labor market outcomes. Most importantly: The fact that language learning ability is correlated with many other, often unobservable, characteristics that could also influence immigrants' job search and earnings is often noted, but difficult to overcome. Yet, even though many of the studies discussed below might not be able to deliver causal estimates, they provide, at the very least, an upper bound on the benefits of language courses, as the following summary illustrates.

Labor market participation seems to increase with local language skills in various countries. Dustmann and Fabbri (2003) find that in the UK, English language acquisition (ELA) increases the chances for a male job-seeker to find gainful employment by 26 percent. For women, the estimates are not statistically significant. In addition, they also find a significant positive effect of English proficiency on earnings.[12] Grondin (2007) shows that the same positive relationship between English speaking ability and probability of employment also exists in Canada. Aldashev, Gernandt and Thomsen (2009) find that in Germany, language proficiency does not only affect immigrants' labor market participation, chances of employment, and earnings, but also their occupational choice.

A positive effect of local language skills has also been documented for earnings. In an early analysis, Tainer (1988) finds a statistically significant positive effect of English proficiency for foreign-born men in the U.S. The extent of the effect,

---

[11] See, for instance, Chiswick and Miller 1995. There are a few studies that are trying to address the potential endogeneity with an instrumental variables strategy (Bleakley and Chin 2004, 2010; Miranda and Zhu 2013; van Ours and Veenman 2006). The instruments that these studies employ depend on the language spoken in the immigrant's country of origin and often also on the age-at-arrival, since both factors influence person's language learning ability.

[12] Note that according to Miranda and Zhu (2013), however, these results were not significant once they controlled for potential endogeneity (which could also be due to the small sample size as they note).

however, varies across ethnicities: There is a larger effect for Hispanics and Asians than for European-born men. Chiswick and Miller (1995) analyze the impact of English language fluency on earnings in Australia, Canada, the U.S. and Israel. In all countries, they find a significant positive effect, varying between 5.3 percent higher earnings in Australia and 16.9 percent higher earnings in the U.S.. The results are also confirmed in later studies for Israel (Chiswick and Repetto 2000), the U.S. (Chiswick and Miller 2002) and Canada (Chiswick and Miller 2003).

Bleakley and Chin (2004) use an instrumental variables strategy based on age of child's arrival in the U.S. and her source country's language to determine the effect of English language skills in the U.S.. They find a significant positive effect of language ability on earnings, arguably mainly driven by years of schooling. Dustmann (1994) confirms this positive effect also for German language ability in former-West Germany.[13] Simliarly, Ispohrding, Otten and Sinning (2014) find a strong positive effect of language ability on wages, arguably mainly mediated through occupational choice. Similar results also exist for ELA in the UK. Shields and Wheatley Price (2002) and Miranda and Zhu (2013), both using an IV strategy, estimate a large positive effect of ELA on wages.[14] Finally, Budría and Swedberg (2015) find that in Spain as well, there is a general positive effect of language abilities on earnings, but it is more pronounced for high-skilled workers. They earn about 50 percent more if they speak Spanish.

Focusing on France, Lochmann, Rapoport and Speciale (2019) leverage a discontinuous assignment role to government-offered language training to document a significant effect of assignment to training on labor force participation of immigrants. This effect is increasing in the immigrants' education levels. Using administrative data from Switzerland, Hangartner and Schmid (2021) are also able to address above-mentioned concerns about endogeneity with a difference-in-differences design. They exploit the quasi-random placement of refugees to Swiss states (cantons) and the existence of a sharp language border dividing German and French-speaking areas and examine the size of the economic gains from proficiency of the host country's language. Compared to otherwise similar English-speaking African asylum seekers, French-speaking asylum seekers have an 80 percent higher probability of finding a job in the first year after arrival due their proficiency in French. This effect is persistent for at least the first five years upon arrival.

Despite some shortcomings, these studies leave little doubt about the importance of proficiency in the host country's language of immigrant integration. They suggest that for arriving asylum seekers' economic integration and the receiving country's public expenditures, providing extensive language training to asylum seekers (and future residents) could prove highly beneficial.

F. Permanent residency permits and citizenship

Another policy that has the potential to facilitate the integration of immigrants is to allow for faster access to permanent residency and citizenship. Faster access to a more permanent form of residence eliminates fears of deportation, and at the same time incentivizes immigrants to invest in a long-term future in the host country. However, that does not necessarily imply that residency permits and citizenship should be offered at the earliest stage. In theory, lowering the threshold for residency permits and citizenship could also have the opposite effect: Rather than incentivizing integration, issuing residency permits and citizenship too early might destroy immigrants' strive to integrate and learn the local language (Hainmueller, Hangartner and Pietrantuono 2017). While causal evidence on the impact of permanent residency and citizenship is fairly scant, there are a few studies that generally show that giving immigrants permanent residency and citizenship has i) a positive effect on political and social and, to a lesser degree, economic integration and ii) that these 'integration returns' are larger if immigrants receive these statuses earlier in the residency period.

---

[13] *See, also, Dustmann and Van Soest (2002).*

[14] *Shields and Wheatly Price (2002) estimate a positive effect of about 16.5% on immigrants' mean hourly occupational wages and Miranda and Zhu (2013) estimate that English deficiency leads to 23% lower wages in the UK.*

With regard to residency permits, Hainmueller, Hangartner and Lawrence (2016) provide panel data evidence based on Swiss registry data (see above) that asylum seekers probability of finding a job increases by 10 percentage points (a 50% percent increase over the average) if they receive subsidiary protection in the first year after arrival. The boost associated with subsidiary protection, which arguably captures both the increase in refugees' motivation of finding work and decrease in employer's uncertainty about the refugee being deported, fairly linearly decreases the longer the refugee has to wait for receiving protection status and is essentially zero after five years upon arrival.

In the domain of citizenship rights, there are several panel data studies that show a positive association between naturalization and labor market outcomes (see Bevelander and Veenman 2008 and OECD 2011 and the references therein). One common problem with these studies is that even when employing panel data, the coefficient for naturalization might not have a causal interpretation if an unobserved factor, such as the decision to stay in the host country for good, causes immigrants to simultaneously apply for citizenship and finding a (better) job.[15] However, Gathmann and Keller (2014) can exploit discontinuities in eligibility rules for immigration reforms in Germany that changed the residency requirements for naturalization. Based on an intention-to-treat analysis, they find only few economic returns for men, but significant, albeit substantively small, returns for women.

To circumnavigate the confounding issue associated with panel data, Hainmueller, Hangartner and Pietrantuono (2015; 2017; 2019) exploit the quasi-random assignment of citizenship in Swiss municipalities that used referendums to decide on naturalization applications of immigrants. Comparing otherwise similar immigrants who narrowly won or narrowly lost their naturalization referendums, they find that receiving Swiss citizenship strongly improved long-term economic, political and social integration. More specifically, they present evidence that barely naturalized – as compared to barely non-naturalized – immigrants have higher earnings, higher levels of political efficacy and knowledge, are more likely to read also Swiss and not exclusively foreign newspapers, are less likely to plan to return to their (or their parents') country of origin, and are less likely to feel discriminated against. Using an index of these outcomes, their studies show that naturalization increases both political and social integration by one standard deviation. They also find that the integration returns to naturalization are much larger for more marginalized immigrant groups[16] and somewhat larger when naturalization occurs earlier, rather than later, in the residency period.

Taken together, these studies support the policy paradigm arguing that naturalization is a catalyst for improving the economic, political and social integration of immigrants – rather than merely the crown on the completed integration process.

G. Fostering meaningful interaction between locals and refugees

In combination, recent studies suggest that it may be important that interactions between asylum seekers and locals are meaningful and sustained (as opposed to mere exposure). Several of the papers discussed in this review speak to the importance of how contact happens. Steinmayr (2021) shows that the clear trend towards more support for right-wing, anti-immigrant parties overall was less extreme in municipalities that were assigned to host asylum seekers. In these municipalities, the arrival and integration of asylum seekers was accompanied, encouraged and facilitated by professionals and volunteers (see also, Gamalerio et al. (2021) on integration-promoting refugee reception centers in Italy). Large positive effects of arriving asylum seekers on anti-immigration party support, however, were documented where refugees predominantly just passed through on their way to other European countries (Dinas et al. 2019, Steinmayr 2021). This implies that the problem-centered media coverage of asylum issues (see, e.g., Eberl et al. 2018) might be less effective at structuring locals' perception of asylum seekers if direct contact is meaningful and repeated (see also Allport 1954, Hopkins 2010).

[15] Consistent with this confounding pattern, Engdahl (2014) finds that immigrants' wages in Sweden actually increase before, not after, their citizenship application is decided.

[16] From (former) Yugoslavia and Turkey.

Recent papers focusing on individual interactions between members of different groups complement the observational evidence mentioned above. Mousa (2020) shows that repeated interactions on the soccer pitch are able to reduce exclusionary attitudes between Christians and Muslims even in a challenging post-conflict context. In addition, not only repeated interaction between asylum seekers and the local community, but also between the (local) government and the local community could prove important. Both Krueger and Pischke (1997) and Karapin (2002) indicate that local political participation seems an important moderator for political violence against immigrants. Despite a huge literature on the 'contact hypothesis' in social psychology and related fields, there is still limited experimental and actionable evidence for policymakers about how best to foster interaction (Paluck and Green 2009). This remains one of the most important avenues for further research.

5. Conclusions

This paper was inspired by our conversations with policy makers, journalists and fellow researchers in and outside of academia. Our role has been to present summaries of the research on labor market and fiscal impacts of (asylum) immigration, political effects of asylum seeker arrival and presence, and the effectiveness of various integration policies. Sooner or later, these conversations inevitably gravitated towards the question: "OK, but what should *we do*?"

In this paper, we offer a twofold answer. First, we argue that an essential component of the response to increases in asylum seeker arrivals is to remain calm. We acknowledge that this may be a formidable task as media coverage of refugees tends to capture the public imagination and worries about arriving refugees may thus receive a disproportionate weight in public debate. We do not have to look far for anecdotes supporting this view. For example, fears about refugees loomed large in the debate preceding the vote for Britain's exit from the EU despite the UK receiving relatively few asylum seekers in 2015. While it is hard to predict the economic impact of the UK leaving the EU, it seems safe to assume that these effects are likely substantially larger than the direct labor market or fiscal effect of refugees living in the UK. More generally, heated public debate increases the risk that important decisions will be made without a sufficient analysis of their first-order effects.

We stress that we do not make a statement about policy *objectives*, but on the *quality* of decision-making. Regardless of the objectives, cool heads are needed to evaluate whether the proposed policies are likely to lead to the desired outcomes. We also recognize that just telling people to calm down is unlikely to be helpful. A substantial share of voters holds deeply skeptical views of current refugee policies, and a growing literature shows that higher numbers of arriving refugees fuel the rise of populist, anti-immigrant parties. We do not advice policy makers to neglect these facts. Rather, we argue that these findings highlight the importance of seeking ways to mitigate the impact of receiving refugees on the broader political process.

Accomplishing this goal likely requires a multifaceted approach, but we view efficient integration policies as a central part of a response. Existing work suggests that such polices can have surprisingly large effects. However, we still lack a sufficient body of research to determine which policies are the most efficient (and for whom).

The second part of our recommendation is thus to increase policy experimentation and evaluation. We would particularly like to see more work on integration policy and programs, and on designing the asylum process (length of the process, labor market access, welfare support) with an eye towards rapid integration of asylum seekers who have a high likelihood of obtaining some form of protection status. Piloting new policies in a way that they are amenable to evaluation is, in principle, relatively straightforward. Many interventions can be tested with RCTs (for example by randomizing the timing when a new policy is implemented) or, alternatively, by creating research designs through the staggered rolled out of policies or the use of discontinuities in eligibility criteria. Given the prominence of refugees in the policy debate, researcher may have a good chance to persuade governments to engage in such experimentation. For example, the Finnish government is already running a large RCT to test a new approach for improving employment of refugees. We believe that similar opportunities are available also elsewhere. □

6. References

Abou-Chadi, T., and Krause, W., 2020. The causal effect of radical right success on mainstream parties' policy positions: A regression discontinuity approach. *British Journal of Political Science*, 50(3), pp.829–847.

Aldashev, A., Gernandt, J. and Thomsen, S. L., 2009. Language usage, participation, employment and earnings. *Labour Economics*, 16(3), pp.330–341.

Allport, G., 1954. *The Nature of Prejudice.* Addison-Wesley.

Altonji, J. G. and Card, D., 1991. The effects of immigration on the labor market outcomes of less-skilled natives. In: *Immigration, Trade, and the Labor Market*, University of Chicago Press, pp.201–234.

Amori M. and Manning, A., 2021. Monopsony and Wage Effects of Migration. *Working paper*.

Andersson Joona, P. and Nekby, L., 2012. Intensive coaching of new immigrants: an evaluation based on random program assignment. *The Scandinavian Journal of Economics*, 114(2), pp.575–600.

Angenendt, S., Engler, M. and Schneider, J., 2013. European Refugee Policy: Pathways to Fairer Burden-Sharing. *SWP Comments, Berlin: German Institute for International and Security Affairs*.

Åslund, O. and Johansson, P., 2011. Virtues of SIN Can Intensified Public Efforts Help Disadvantaged Immigrants?. *Evaluation review*, 35(4), pp.399–427.

Arendt, J. N, Bolvig, I., Foged, M., L. Hasager and Peri, G., 2021. Language training and refugees' integration. *NBER Working Paper* 26834.

Bakker, L., Dagevos, J. and Engbersen, G., 2014. The importance of resources and security in the socio-economic integration of refugees. A study on the impact of length of stay in asylum accommodation and residence status on socio-economic integration for the four largest refugee groups in the Netherlands. *Journal of International Migration and Integration*, 15(3), pp.431–448.

Bansak, K., Hainmueller, J. and Hangartner, D., 2016. How economic, humanitarian, and religious concerns shape European attitudes toward asylum seekers. *Science, 354*(6309), pp.217–222.

Bansak, K., Hainmueller, J. and Hangartner, D., 2017. Europeans support a proportional allocation of asylum seekers. *Nature Human Behavior 1*, 0133.

Barone, G., D'Ignazio, A., de Blasio, G. and Naticchioni, P., 2016. Mr. Rossi, Mr. Hu and politics. The role of immigration in shaping natives' voting behavior. *Journal of Public Economics*, 136, pp.1–13.

Barslund, M., Busse, M., Lenaerts, K., Ludolph, L. and Renman, V., 2016. Labour market integration of refugees. A comparative survey of Bosnians in five EU countries, *Mercator Dialogue on Asylum and Migration*, Brussels: Centre for European Policy Studies.

Bevelander, P. and Veenman, J., 2008. Naturalisation and Socioeconomic Integration: The Case of the Netherlands, in *The Economics of Citizenship*, pp.63–88.

Blau, F. and Mackie, C. (editors), 2016. *The Economic and Fiscal Consequences of Immigration*. National Academies of Sciences, Panel on the Economic and Fiscal Consequences of Immigration.

Bleakley, H. and Chin, A., 2004. Language Skills and Earnings: Evidence from Childhood Immigrants. *Review of Economics and Statistics, 86*(2), pp.481–496.

Bleakley, H. and Chin, A., 2010. Age at arrival, English proficiency, and social assimilation among US immigrants. *American economic journal. Applied economics, 2*(1), pp.165–192.

Blumer, H., 1998. Race prejudice as a sense of group position. In *New Tribalisms*. UK: Palgrave Macmillan, pp.31–40.

Borjas, G.J., 1999. The Economic Analysis of Immigration, In: Orley C. Ashenfelter and David Card, Editor(s), *Handbook of Labor Economics 3A*, pp.1697–1760, Elsevier.

Borjas, G, J., 2015. *Labor Economics*. 7th edition. McGraw-Hill.

Borjas, G. J., 2016. The Wage Impact of the Marielitos: Additional Evidence, *NBER Working Paper* w21850.

Borjas, G. J., Monras, J., 2017. The labour market consequences of refugee supply shocks, *Economic Policy*, 32(91), pp.361–413.

Bratti, M., Deiana, C., Havari, E., Mazzarella, G. and Meroni, E. C., 2020. Geographical proximity to refugee reception centres and voting. *Journal of Urban Economics*, 120, pp. 103290

Braun, R. and Koopmans, R., 2010. The diffusion of xenophobic violence in Germany: The role of social similarity. *European Sociological Review*, 26(1), pp.1–25.

Brell, C., Dustmann, C. and Preston, I., 2020. The Labor Market Integration of Refugee Migrants in High-Income Countries. *Journal of Economic Perspectives*, *34*(1), pp.94–121.

Brunner, B. and Kuhn, A., 2014. Immigration, cultural distance and native' attitudes towards immigrants: evidence from Swiss voting results. *IZA Discussion Paper*, 8409.

Budría, S. and Swedberg, P., 2015. The impact of language proficiency on immigrants' earnings. *Revista de economía aplicada, 23*(67), pp.63–91.

Card, D., Kluve, J. and Weber, A., 2010. Active Labour Market Policy Evaluations: A Meta-Analysis. *Economic Journal, 120*(548), pp.F452–F477.

Card, D., Kluve, J. and Weber, A., 2018. What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations. *Journal of the European Economic Association, 16*(3), pp.894–931.

Chiswick, B. R. and Miller, P. W., 1995. The Endogeneity between Language and Earnings: International Analyses. *Journal of Labor Economics, 13*(2), pp.246–288.

Chiswick, B. R. and Miller, P. W., 2002. Immigrant Earnings: Language Skills, Linguistic Concentrations and the Business Cycle. *Journal of Population Economics, 15*, pp.31–57.

Chiswick, B. R. and Miller, P. W., 2003. The complementarity of language and other human capital: immigrant earnings in Canada. *Economics of Education Review, 22*(5), pp.469–480.

Chiswick, B. R. and Repetto, G., 2000. Immigrant Adjustment in Israel : Literacy and Fluency in Hebrew and Earnings Immigrant Adjustment in Israel : *IZA Discussion Paper, 177*.

Cools, S., Finseraas, H. and Rogeberg, O., 2021. Local Immigration and Support for Anti-Immigration Parties: A Meta-Analysis. *American Journal of Political Science*.

Dahlberg, M., Edmark, K. and Lundqvist, H., 2012. Ethnic diversity and preferences for redistribution. *Journal of Political Economy, 120*(1), pp.41–76.

Dahlberg, M., Edmark, K. and Lundqvist, H., 2016. Ethnic diversity and preferences for redistribution: Reply. Scandinavian Journal of Economics.

Dahlberg, M., Egebark, J., Vikman, U., Özcan, G., 2020. Labor Market Integration of Low-Educated Refugees: RCT Evidence from an Ambitious Integration Program in Sweden. IFN Working Paper No. 1372.

Dancygier, R.M., 2010. *Immigration and conflict in Europe*. Cambridge University Press.

Dancygier, R. M. and Laitin, D. D., 2014. Immigration into Europe: Economic discrimination, violence, and public policy. *Annual Review of Political Science, 17*, pp.43–64.

Dinas, E., Matakos, K., Xefteris, D. and Hangartner, D., 2019. Waking up the golden dawn: does exposure to the refugee crisis increase support for extreme-right parties? *Political analysis, 27*(2), pp.244–254.

Dinesen, P. T., Schaeffer, M. and Sønderskov, K. M., 2020. Ethnic Diversity and Social Trust: A Narrative and Meta-Analytical Review. *Annual Review of Political Science*, *23*(1), pp. 441–465

Dustmann, C., 1994. Speaking fluency, writing fluency and earnings of migrants. *Journal of Population Economics, 7*, pp.133–156.

Dustmann, C. and Fabbri, F., 2003. Language Proficiency axnd Labour Market Performance of Immigrants in the UK. *Economic Journal, 113*(489), pp.695–717.

Dustmann, C. and Preston, I., 2001. Attitudes to ethnic minorities, ethnic context and location decisions. *The Economic Journal, 111*(470), pp.353–373.

Dustmann, C., Schönberg, U. and Stuhler, J., 2016. The Impact of Immigration: Why Do Studies Reach Such Different Results? *Journal of Economic Perspectives, 30*(4), pp.31–56.

Dustmann, C. and Van Soest, A., 2002. Language and the Earnings of Immigrants. *Industrial and Labor Relations Review, 55*(3), pp.473–492.

Dustmann, C., Vasiljeva, K. and Piil Damm, A., 2019. Refugee migration and electoral outcomes. *The Review of Economic Studies, 86*(5), pp.2035–2091.

Eberl, J., Meltzer, C. E., Heidenreich, T., Herrero, B., Theorin, N., Lind, F., Berganza, R., Boomgaarden, H. G., Schemer, C. and Strömbäck, J., 2018. The European Media Discourse on Immigration and Its Effects: A Literature Review. *Annals of the International Communication Association, 42*(3), pp.207–23.

Engdahl, M., 2014. Naturalizations and the economic and social integration of immigrants. *Uppsala Center for Labor Studies Working paper, 6*.

Fasani, F., Frattini, T. and Minale, L., 2021 (The Struggle for) Refugee Integration into the Labour Market: Evidence from Europe. *Journal of Economic Geography*.

Falk, A., Kuhn, A. and Zweimüller, J., 2011. Unemployment and Right-Wing Extremist Crime: Unemployment and Right-Wing Extremist Crime. *Scandinavian Journal of Economics, 113*(2), pp.260–285.

Foged, M. and Peri, G., 2016. Immigrants' Effect on Native Workers: New Analysis on Longitudinal Data. *American Economic Journal: Applied Economics, 8*(2), pp.1–34.

Gamalerio, M., Luca, M., Romarri, A. and Viskanic, M., 2021. Refugee reception, extreme-right voting, and compositional amenities: evidence from Italian municipalities. *Working Paper*. https://www.matteogamalerio.com/content/uploads/2021/10/SPRAR_Far_Right_2021.pdf

Gathmann, C. and Keller, N., 2014. Returns to Citizenship? Evidence from Germany's Recent Immigration Reforms. *Working Paper*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2432911

Gessler, T., Tóth, G. and Wachs, J., 2021. No Country for Asylum Seekers? How Short-Term Exposure to Refugees Influences Attitudes and Voting Behavior in Hungary. *Political Behavior*.

Giuliano, P. and Tabellini, M., 2021, The Seeds of Ideology: Historical Immigration and Political Preferences in the United States. *Harvard Business School BGIE Unit Working Paper No. 20-118*. https://ssrn.com/abstract=3606506

Golder, M., 2016. Far right parties in Europe. *Annual Review of Political Science, 19*, pp.477–497.

Grondin, C., 2007. Knowledge of official languages among new immigrants: How important is it in the labour market. *Ottawa: Statistics Canada*.

Hainmueller, J., Hangartner, D. and Lawrence, D., 2016. When lives are put on hold: Lengthy asylum processes decrease employment among refugees. *Science Advances, 2*(8), p.e1600432.

Hainmueller, J., Hangartner, D. and Pietrantuono, G., 2015. Naturalization fosters the long-term political integration of immigrants. *Proceedings of the National Academy of Sciences, 112*(41), pp.12651–12656.

Hainmueller, J., Hangartner, D. and Pietrantuono, G., 2017. Catalyst or crown: Does naturalization promote the long-term social integration of immigrants?. *American Political Science Review, 111*(2), pp.256–276.

Halla, M., Wagner, A. F. and Zweimüller, J., 2017. Immigration and voting for the far right. *Journal of the European Economic Association, 15*(6), pp.1341–1385.

Hangartner, D., Dinas, E., Marbach, M., Matakos, K. and Xefteris, D., 2019. Does exposure to the refugee crisis make natives more hostile?. *American Political Science Review, 113*(2), pp.442–455.

Hangartner, D. and Schmid, L., 2021. Migration, Language and Employment. *Working Paper*.

Hanson. G., 2009. The Economic Consequences of the International Migration of Labor. *Annual Review of Economics, 1*, pp.179–208

Harmon, N.A., 2018. Immigration, ethnic diversity, and political outcomes: Evidence from Denmark. *The Scandinavian Journal of Economics, 120*(4), pp.1043–1074.

Hatton, T. J., 2004. Seeking asylum in Europe. *Economic policy, 19*(38), pp.6-62.

Hopkins, D. J., 2010. Politicized places: Explaining where and when immigrants provoke local opposition. *American Political Science Review, 104*(01), pp.40–60.

Isphording, I. E., Otten, S. and Sinning, M., 2014. The Effect of Language Deficiency on Immigrant Labor Market Outcomes in Germany. *Mimeo*.

Jensen, K., 2020. The Political Consequences of Immigration: Evidence from Refugee Shocks in Denmark. *Working Paper*. https://www.cemfi.es/ftp/pdf/papers/Seminar/Jensen_JMP_Jan11.pdf

Karapin, R., 2002. Antiminority Riots in Unified Germany: Cultural Conflicts and Mischanneled Political Participation. *Comparative Politics, 34*(2), pp.147–167.

Krueger, A. B. and Pischke, J. S., 1997. A statistical analysis of crime against foreigners in unified Germany. *Journal of Human Resources, 32*(1), pp.182–209.

Kymlicka, W., 1995. *Multicultural citizenship: A liberal theory of minority rights*. Clarendon Press.

Lee, R. and Miller, T., 2000. Immigration, social security, and broader fiscal impacts. *American Economic Review, 90*m pp.350–354.

Lonsky, J., 2021. Does immigration decrease far-right popularity? Evidence from Finnish municipalities. *Journal of Population Economics, 34*(1), pp.97–139.

Lochmann, A., H. Rapoport and B. Speciale. 2019. The effect of language training on immigrants' economic integration: Empirical evidence from France. *European Economic Review, 113*, pp.265–296.

Malmström, C. 2014. Commissioner Cecilia Malmström commemorates the Lampedusa tragedy. *European Commission Statement, 14/296* (2 October).

Marbach, M. and Ropers, G., 2018, Not in My Backyard: Do Increases in Immigration Cause Political Violence? *IPL Working Paper*, 18–02, https://ssrn.com/abstract=3310352

Martikainen, T., Saari, M. and Korkiasaari J., 2013. Kansainväliset muuttoliikkeet ja Suomi. In Martikainen, T., Saukkonen, P. and Säävälä, M. (toim.) *Muuttajat. Kansainvälinen muuttoliike ja suomalainen yhteiskunta*. Helsinki: Gaudeamus.

McLaren, L. M., 2012. Immigration and trust in politics in Britain. *British Journal of Political Science, 42*(01), pp.163–185.

McLaren, L. M., 2015. I*mmigration and Perceptions of National Political Systems in Europe*. USA: Oxford University Press.

Mendez, I. and Cutillas, I. M., 2014. Has immigration affected Spanish presidential elections results? *Journal of Population Economics, 27*(1), pp.135–171.

Miranda, A. and Zhu, Y., 2013. English deficiency and the native–immigrant wage gap. *Economics Letters, 118*(1), pp.38–41.

Moore, M. and Ramsay, G., 2017. UK media coverage of the 2016 EU referendum campaign. *Report by the Policy Institute at King's College, London*.

Mousa, S., 2020. Building social cohesion between Christians and Muslims through soccer in post-ISIS Iraq. *Science, 369*(6505), pp.866–870.

Nekby, L. and Pettersson-Lidbom, P., 2016. Revisiting the Relationship between Ethnic Diversity and Preferences for Redistribution. *Scandinavian Journal of Economics*.

OECD, 2011. *Naturalisation: A Passport for the Better Integration of Immigrants?* Paris: OECD Publishing.

OECD, 2013. Sweden – *DAC Peer Review of Development Co-operation*. Paris: OECD Publishing.

OECD, 2020. New asylum applications since 1980 in the OECD and the European Union, *International Migration Outlook 2020*, Paris: OECD Publishing, https://doi.org/10.1787/395bab9d-en

Otto, A. H. and Steinhardt, M. F., 2014. Immigration and election outcomes – Evidence from city districts in Hamburg. *Regional Science and Urban Economics, 45*, pp.67–79.

Paluck, E. L. and Green, D. P., 2009. Prejudice reduction: What works? A review and assessment of research and practice. *Annual review of psychology, 60*, pp.339-367.

Pesola, H. and Sarvimäki, M., 2021. Intergenerational Spillovers of Integration Policies: Evidence from Finland's Integration Plans. *Working paper*.

Putnam, R. D., 2007. E pluribus unum: Diversity and community in the twenty first century the 2006 Johan Skytte Prize Lecture. *Scandinavian political studies, 30*(2), pp.137–174.

Peri, G. and Yasenov, V., 2015. *The labor market effects of a refugee wave: Applying the synthetic control method to the Mariel boatlift*. National Bureau of Economic Research, No. w21801.

Quillian, L., 1995. Prejudice as a response to perceived group threat: Population composition and anti-immigrant and racial prejudice in Europe. *American Sociological Review, 60*(4), pp.586–611.

Ruist, J. 2015. The fiscal cost of refugee immigration: the example of Sweden. *Population and Development Review*, 41, pp.567–581

Ruist, J. 2020. The fiscal aspect of the refugee crisis. *International Tax and Public Finance*, 2, pp.478–492.

Sarvimäki, M., 2011. Assimilation to a Welfare State: Labor Market Performance and Use of Social Benefits by Immigrants to Finland. *Scandinavian Journal of Economics, 113*(3), pp.665–688.

Sarvimäki, M., 2017. Labor market integration of refugees in Finland. *Nordic Economic Policy Review*, 7(1): pp.91–114.

Sarvimäki, M. and Hämäläinen, K., 2016. Integrating Immigrants: The Impact of Restructuring Active Labor Market Policies. *Journal of Labor Economics, 34*(2), pp.479–508.

Sarvimäki, M., Ansala, L., Eerola, E., Hämäläinen, K., Hämäläinen, U., Pesola, H. and Riihelä, M., 2014. Maahanmuuttajien integroituminen Suomeen. *VATT Julkaisut 67*.

Savolainen, R., 2016. How Policies Influence Voter Sentiments: Evidence from Asylum Center Placements. *Working Paper*.

Schaub, M., Gereke, J. and Baldassarri, D., 2021. Strangers in Hostile Lands: Exposure to Refugees and Right-Wing Support in Germany's Eastern Regions. *Comparative Political Studies, 54*(3-4), pp.686–717.

Shields, M. A. and Wheatley Price, S., 2002. The English language fluency and occupational success of ethnic minority immigrant men living in English metropolitan areas. *Journal of Population Economics, 15*, pp.137–160.

Spirig, J. 2021. When issue salience affects adjudication: evidence from Swiss asylum appeal decisions. *American Journal of Political Science*.

Spoon, J. J. and Klüver, H., 2020. Responding to far right challengers: does accommodation pay off? *Journal of European Public Policy, 27*(2), pp.273–291.

Steinmayr, A., 2021. Contact versus Exposure: Refugee Presence and Voting for the Far-Right. *The Review of Economics and Statistics, 103*(2): pp.310–327.

Stepick, A. and Portes, A., 1986. Flight into despair: a profile of recent Haitian refugees in South Florida. *International Migration Review, 20*(2), pp.329–350.

Storesletten, K., 2000. Sustaining fiscal policy through immigration. *Journal of Political Economy, 108*, pp.300–323.

Swedish Government, 2015. Measures to tackle the refugee crisis. http://www.government.se/contentassets/f8effa03946941c5987f7ae76b356a02/agreement-measures-to-tackle-the-refugee-crisis.pdf

Tainer, E., 1988. English Language Proficiency and the Determination of Earnings among Foreign-Born Men. *The Journal of Human Resources, 23*(1), pp.108–122.

Thielemann, E., 2010. The common European asylum system: In need of a more comprehensive burden-sharing approach. In Migrants and Minorities: *The European Response*, pp.82–97.

Van Ours, J. C. and Veenman, J., 2006. Age at immigration and educational attainment of young immigrants. *Economics Letters, 90*(3), pp.310–316.

Vertier, P., Viskanic, M. and Gamalerio, M., 2020. Dismantling the 'jungle': Migrant relocation and extreme voting in france. *Working paper*.

Walter, S., 2021. The Backlash Against Globalization. *Annual Review of Political Science, 24*(1), pp.421–442.

Waxman, P., 2001. The Economic Adjustment of Recently Arrived Bosnian, Afghan and Iraqi Refugees in Sydney, Australia. *The International Migration Review, 35*(2), pp.472–505.

Pricing the Pharmaceuticals when the Ability to Pay Differs*

Vesa Kanniainen
Discipline of Economics, University of Helsinki, Finland
and CESifo, Münich
vesa.kanniainen@helsinki.fi

Juha Laine
Department of Economics, InFLAMES Flagship,
University of Turku, Finland
juha.r.laine@utu.fi

Ismo Linnosmaa (Corresponding author)
Department of Health and Social Management,
University of Eastern Finland and
Welfare State Research and Renewal Unit,
Finnish Institute for Health and Welfare
ismo.linnosmaa@uef.fi

Abstract

A non-trivial fraction of people cannot afford to buy pharmaceutical products at unregulated market prices. This paper analyses the public insurance of a patent-protected pharmaceutical product in terms of price controls and socially optimal third-degree price discrimination. First, the paper characterizes the Ramsey pricing rule in the case where the producer price has to cover the *R&D* costs of the firm and patients' pharmaceutical expenditures are not covered by health insurance. Subsequently, conditions for a welfare increasing departure from the Ramsey pricing rule are stated in terms of price regulation and health insurance coverage. Unlike the earlier views expressed, the increased consumption of the pharmaceutical is shown to be welfare increasing. In the spirit of the Rawlsian view, a criterion for vertical equity is examined as an optimal means-tested health insurance. In this scheme, the regulator chooses a higher insurance coverage for individuals whose income is below an endogenously determined income threshold. The means-tested insurance scheme improves social welfare but also yields very equal market outcomes.

Keywords: *Pharmaceuticals, price regulation, public health insurance, third-degree price discrimination, equity criterion*
JEL codes: *L1, L5, I18*

**The technicalities of the paper have been developed over various stages. We gratefully acknowledge comments from the participants of the Health Economics Day 2013, Helsinki, and the Nordic Health Economics Study Group meeting 2013, Oslo, on a preliminary version of the article. The authors are also indebted to an anonymous referee and Mika Kortelainen, the editor, for helpful comments and suggestions. Linnosmaa acknowledges the financial support (grant N:o 138723) from Academy of Finland. A previous version of this article was published as a CESifo Working Paper No. 8031, January 13, 2020.*

1. Introduction

The ability to pay for pharmaceuticals varies among people. A non-trivial fraction of people cannot afford to buy pharmaceutical products at unregulated market prices. Those products are created through expensive and risky *R&D* programmes committing the pharmaceutical firms to rather high expenditures. Those expenditures should subsequently be covered through prices, which, however, may turn out to be too high to be socially acceptable. In the current paper, a question is raised concerning how to introduce means-tested subsidies to low-income citizens as part of optimal regulation and yet to maintain the incentives for a pharmaceutical company to invest in the *R&D*. Therefore, the conflict between efficiency and equity has to be resolved via optimal pricing.

Apart from the efficiency considerations, policy-makers typically emphasize equitable access to services due to the fact that in many countries, if not in most, low-income people are not able to buy the medication they need. Indeed, the health policy concerning the medical industry often expressed in the official documents states that "the purpose of the medical policy is to provide to citizens high-quality and cost-efficient pharmaceuticals at reasonable prices....". Moreover, the PPRI Report 2018 provides information about currently existing pharmaceutical pricing and reimbursement policies in the 47 PPRI member countries. It turned out that 42 PPRI network member countries have mechanisms in place to set medicine prices at the ex-factory (or sometimes wholesale) price level, mostly targeting reimbursable medicines or prescription-only medicines. 46 PPRI network member countries have at least one reimbursement list for outpatient medicines in place, and in 31 PPRI countries the reimbursement lists relate to both outpatient and inpatient sectors. In addition, hospital pharmaceutical formularies are managed at the level of hospitals in most PPRI countries. At least 43 countries charge co-payments for outpatient reimbursable medicines (frequently percentage copayments, but also a prescription fee and/or a deductible). All these 43 countries apply exemptions from or reductions of co-payments for vulnerable and other defined population groups (Vogler et al., 2019).

In Finland, pharmaceuticals are delivered and financed by different channels although the system is tax-based. Reimbursed drugs are delivered from pharmacies and costs are covered by Social Insurance Institution and patient copayments. Reimbursement categories (40%, 65% and 100%) are based on disease severity. Medication during hospital visits is covered by municipalities and costs are incorporated into the hospital payment. Decision-making bodies and the criteria used in health technology assessment and regulation vary in different channels. Patient income is not a decision-making criterion, but people with extremely low incomes may get pharmaceuticals for free from the Social Insurance Institution's income support.

Previous work based on the efficient price regulation of pharmaceutical products and health insurance has produced a number of important contributions. The basic idea has been cast in terms of the optimal product taxation in a one-person or many-people economy with Ramsey's (1927) idea of equal percentage reductions in (compensated) demands for all commodities (Diamond, 1975). Based on such foundations, Besley (1988) explored the trade-off between risk sharing and the incentives to consume medical care inherent in health insurance. Earlier, Feldstein (1973) had expressed concerns about the welfare cost of excess health insurance induced by the adverse incentives of the consumption of health care. The interaction of pricing and insurance coverage in the pharmaceutical market was addressed by Barros and Martinez-Giralt (2008), who considered the normative allocation of *R&D* costs across different markets served by a pharmaceutical firm. They showed that a higher insurance coverage calls for higher prices not only because of a lower demand elasticity but also due to a larger moral hazard effect in the consumption of the pharmaceuticals. The equilibrium pricing rule appeared to deviate from the standard Ramsey pricing rule: for equal demand elasticities, and given the distortion cost of funds, a country with a higher coverage rate will have higher-priced pharmaceuticals as well.

Gaynor et al. (2000) also focussed on the excessive consumption of medical products caused by insurance, that is, the moral hazard. In a related area, Grassi and Ma (2011; 2012) studied the provision of public supply of health care services but with non-price rationing when the income levels of people are different. When the rationing is based on wealth information (as is the case in the USA), the optimal policy in their analysis rations public services to low-income people, while leaving the high-income people to buy services from the private market. If also the cost is observed, the optimal rationing turns out to be based on cost-effectiveness (as in most European countries and Canada). Baicker, Mullainathan and Schwartzstein (2015) suggest that "behavioural hazard" can make people misuse health care. They suggest that health insurance can do more than just provide financial protection – it can also improve health care efficiency. A comprehensive survey of the literature on pricing pharmaceuticals has been documented by Borges dos Santos et al. (2019).

Abbot and Vernon (2005) have demonstrated how pharmaceutical price controls will significantly diminish the incentives to undertake early-stage *R&D* investment. In the current paper, and in contrast to the existing work in the area, the question is raised of how to introduce means-tested subsidies to low-income citizens as part of the optimal regulation and yet to maintain the incentives of a pharmaceutical company in investing in *R&D*. To fix the ideas of the paper, a market for a pharmaceutical product with one firm having innovated a new product is considered. The firm is the sole producer of the product, say through the patent protection. The cost of innovation is sunk at the time the product is sold on the market, and it causes a decrease in the average cost of producing the pharmaceutical product. Three policies will be analysed: Ramsey pricing without insurance, price regulation with insurance (henceforth, price-insurance policy), and a means-tested price and insurance policy (henceforth, means-tested price-insurance policy). Throughout the analysis, we will allow the patient population to be heterogeneous in terms of the ability to pay (income) and analyse questions related to the access to pharmaceutical treatment.

Initially, equity issues are ignored. As the equality between price and the marginal cost of producing the pharmaceutical does not represent a feasible starting point for the price regulation, the Ramsey pricing rule is a natural candidate to be studied in the absence of health insurance. Next, conditions for a welfare increasing departure from Ramsey pricing in terms of price regulation and optimal insurance coverage are derived, taking the social cost of public funds into account. Our results provide insights in to why both price regulation and social insurance are desirable. Subsequently, the paper also addresses the fact that a non-trivial fraction of patients cannot afford to buy the pharmaceutical product even at regulated and subsidized market prices and poses a question of whether means-tested insurance coverage rates have the potential to improve welfare. We thus arrive at the socially optimal, third-degree out-of-pocket price discrimination. Our analysis complements that of Grassi and Ma (2011; 2012) who analysed efficient non-price rationing schemes. Moreover, while Gaynor et al. (2000) worked with the case of a private health insurance market, the focus in the current paper is instead on the public (or social) health insurance.

When Ramsey pricing is compared with the price-insurance policy, our findings indicate that the moral hazard in terms of increased consumption of pharmaceutical products is welfare increasing. Without the health insurance, the prices would be excessively high as the firm's *R&D* costs have to be recovered. The result that the introduction of health insurance improves social welfare is due to our focus on designing a socially optimal health insurance coverage in a price-regulated pharmaceutical market characterized by increasing returns to scale.

Yet, the second-best equilibrium with public health insurance also has some undesirable properties: low-income people are left without the medication they need. As the optimal means-tested insurance, we explore an equity-based health insurance scheme in the spirit of the Rawlsian view. In this scheme, the regulator chooses a higher insurance coverage for individuals with an income below a threshold (low-income patients) and a lower insurance coverage for individuals above the income threshold (high-income patients). Under this scheme, the income threshold categorising patients into the low-income and high-income groups is determined endogenously.

Our results show that in the Rawlsian world with equity based on maximizing the aggregate consumer surplus and conditional on the improved access to pharmaceutical treatment by means-tested insurance coverages, the consumption of the pharmaceutical and the consumer surpluses are split equally between the low- and high-income patients. In this respect, the optimal means-tested policy yields very equal market outcomes. It is also shown that the optimal means-tested price-insurance policy provides a strictly higher social welfare than the optimal price-insurance policy with no equity concern.

Before presenting the model (Section 2) and its analysis (Sections 3–5), we comment on the potential information problems as follows. First, although the regulator is uninformed about the individual incomes of the patients in Sections 2–4, the income distribution is known. This is all the information needed in the Ramsey problem and in the optimal price-insurance policy analysis. Secondly, in Section 5, the regulator uses the means-tested approach to classify the patients into low-income and high-income groups and the regulator is assumed to have access to income information that is needed to construct optimal policies.

2. Model

We consider a market for a new pharmaceutical product. There is a single monopoly producer holding the patent and selling the product. The demand side of the market consists of patients in need of the pharmaceutical treatment that the firm produces. The current health state of each patient is h_0. The consumption of the pharmaceutical improves patients' health to $h_1 > h_0$. The effectiveness of the pharmaceutical treatment can be measured by the difference $\Delta = h_1 - h_0$.

Patients derive utility from health h and consumption goods x. Each patient has a utility function $u(x, h)$, which is assumed to be a strictly increasing function in both consumption goods and health. In the spirit of Grossman (1972), patients consume the pharmaceutical to produce health. The production function for health is $h = h(j) = h_0 + \Delta j$, where the indicator $j = 1, 0$ describes whether or not a patient consumes the pharmaceutical.

2.1 Ability to pay, willingness to pay and the demand for the pharmaceutical

Patients are heterogeneous in their ability to pay. To capture such heterogeneity formally, we introduce a randomly distributed income variable w, assumed to follow the $U\,[0, 1]$ distribution. The income variable measures disposable income and is adjusted for the patients' tax payments to the government.

We first show how the willingness to pay for the pharmaceutical product, denoted θ, is determined by the patient's ability to pay using the approach developed by Grassi and Ma (2011; 2012). Let the variable p denote the producer price of the pharmaceutical product. The budget constraint of the patient with income w can then be written as $w = x + (1 - r)pj$, where the binary indicator j describes the patient's consumption of the pharmaceutical, the variable r stands for the insurance coverage, and the price of consumption goods is normalized to one.

Assuming a separable utility function, the patient with income w obtains utility

$$u(w - (1 - r)pj) + v(h(j)) \tag{1}$$

from the consumption of the pharmaceutical. The first (second) part of the utility function $u(x)$ ($v(h)$) measures the patient's utility from consumption goods (health).

Using the Cobb-Douglas utility function with constant returns to scale, we show in the Appendix (Result 1) that the relationship between the willingness to pay and income is $\theta(w) = w\eta$, where

$$\eta = 1 - (1 + \frac{\Delta}{h_0})^{\frac{-\alpha}{1-\alpha}}, \tag{2}$$

and the fraction $\frac{\Delta}{h_0}$ measures the relative effectiveness of the pharmaceutical and $0 < \alpha < 1$ is the preference weight that patients give to health. Henceforth, the parameter $\eta > 0$ will be called quality weight as it is determined by the health effects of the pharmaceutical.

We will assume conditions that allow us to adopt the parametrization $\theta(w) = w\eta$, where η is given by (2). As a consequence, a patient with income w obtains consumer surplus

$$CS_w(j) = (w\eta - (1 - r)p)\, j \tag{3}$$

from consumption of the pharmaceutical. The consumer with income w_i is indifferent with regard to consuming or not consuming the pharmaceutical. Therefore, the condition $CS_w(1) = CS_w(0) = 0$ can be solved with respect to the income of the indifferent consumer:

(4) $$w_i = \frac{(1-r)p}{\eta}.$$

Given the producer price and the insurance coverage, the demand for the pharmaceutical is given by the number of buying, high-income patients:

(5) $$q(p,r) = 1 - w_i = 1 - \frac{(1-r)p}{\eta}.$$

The total consumer surplus from the consumption of pharmaceuticals is defined as follows:

(6) $$CS(p,r) = \int_{\frac{p(1-r)}{\eta}}^{1} (w\eta - (1-r)\,p)\,dw.$$

2.2 Producer

The profit of the pharmaceutical firm is

(7) $$\pi(p,r) = (p-c)q(p,r) - F,$$

where $c > 0$ is the marginal cost of production and $F > 0$ is a fixed (sunk) cost from *R&D* activities prior to the launch of the pharmaceutical product. In the following analysis, we assume that the quality weight exceeds the marginal cost of producing the pharmaceutical:

Assumption 1: $c < \eta$.

Assumption 1 guarantees the existence of an active market for the pharmaceutical. If Assumption 1 was not true, there would be no patients whose willingness to pay for the pharmaceutical exceeds the marginal cost of producing the pharmaceutical. This implies that there would be no possibilities for market exchange.

2.3 Regulator

The regulator is benevolent and chooses the producer price and the insurance coverage to maximize social welfare, which is defined as the sum of the consumer surplus and the firm's profit subtracted by the cost of financing health insurance:

(8) $$W = CS + \pi - (1+\lambda)T.$$

In (8), T is the tax revenue raised to finance the health insurance. We assume that each euro collected through taxation to finance the pharmaceutical expenditures costs $(1 + \lambda)$ for society and where $\lambda \geq 0$ measures the marginal cost of public funds. The regulator maximizes social welfare (8) subject to the budget constraint

(9) $$T \geq rpq(p,r) \equiv IE(p,r).$$

The right-hand side of the inequality (9) measures the public health insurance expenditures due to the consumption of the pharmaceutical.

Since the value of the social welfare function (8) decreases as the tax revenue T increases, the regulator is not willing to collect more tax revenue than the amount of the aggregate health insurance expenditure. This implies that the budget constraint (9) must be binding in any solution to the regulator's problem. The social welfare function can then be restated as follows:

(10) $$W = CS(p,r) + \pi(p,r) - (1+\lambda)IE(p,r).$$

The regulator's problem is to choose the price-insurance policy (p, r) which maximizes the value of social welfare (10) subject to the profit constraint

(11) $$\pi(p,r) \geq 0$$

and the constraints defining feasible price-insurance policies: $p \geq 0$ and $0 \leq r \leq 1$.

2.4 Timing

We will examine a strategic game between the regulator and the producer of the pharmaceutical. The sequence of moves in the game is as follows. The regulator first chooses the producer price p and the insurance coverage r, after which the firm either accepts or rejects the regulator's proposal. If the firm accepts the proposal, patients decide whether or not to consume the pharmaceutical and the firm produces the amount of the pharmaceutical demanded by the patients.[1] To concentrate on analysing equity consequences of various price-insurance policies, it is assumed throughout the article that the quality weight, marginal and fixed costs and the marginal cost of public funds are common knowledge.

2.5 First-best solution

An efficient benchmark to the regulator's problem is the first-best price and quantity of the pharmaceutical, which maximize social welfare that is not influenced by health insurance coverage:

(12) $$W_f = CS(p,0) + \pi(p,0).$$

The first-best welfare is achieved by setting the price of the pharmaceutical equal to the marginal cost, that is $p_f = c$. The amount of pharmaceuticals consumed in the first-best solution is $q(c, 0) = (\eta - c)/\eta$, and the corresponding social welfare value is

(13) $$\bar{W}_f = CS(p_f,0) + \pi(p_f,0) = \frac{(\eta-c)^2}{2\eta} - F.$$

It is also understood that the regulator cannot implement the marginal-cost pricing scheme because that would yield the profit $-F$, which the firm is not willing to accept.

3. Ramsey price

Understanding that the marginal-cost pricing cannot be implemented, we first consider the pricing that maximizes welfare and satisfies the firm's profit constraint as the benchmark case. Furthermore, and to leave the analysis of the optimal insurance coverage to the subsequent sections, we assume that the regulator does not subsidize the patients' pharmaceutical expenditures through health insurance, but selects $r = 0$. Under such policy, the consumption of the pharmaceutical has no effect on public health insurance expenditures.

[1] *The regulator acts as a Stackelberg leader relative to the producer and consumers.*

The problem of the regulator can be defined as finding the pharmaceutical price which maximizes social welfare

$$W = CS(p,0) + \pi(p,0) \tag{14}$$

subject to the profit constraint

$$\pi(p,0) \geq 0. \tag{15}$$

The solution of the above problem defines the Ramsey-Boiteux price (e.g. Armstrong and Sappington, 2007). With L denoting the value of the Lagrangian function, the necessary condition for the Ramsey price is defined as follows:

$$\frac{\partial L}{\partial p} = \frac{\partial CS(p,0)}{\partial p} + (1+\mu)\frac{\partial \pi(p,0)}{\partial p}$$

$$= -\left(1 - \frac{p}{\eta}\right) + (1+\mu)\left(\left(1-\frac{p}{\eta}\right) - \frac{(p-c)}{\eta}\right) = 0, \tag{16}$$

where μ is a positive-valued Lagrange multiplier of the profit constraint. In addition to the condition (16), the solution of the regulator's problem must satisfy $-\pi(p, 0) \leq 0$, $\mu \geq 0$ and $-\mu\pi(p, 0) = 0$.

Straightforward computation shows that social welfare (14) is decreasing with all pharmaceutical prices higher than the marginal cost[2]. Therefore, the regulator wants to reduce the price of the pharmaceutical until the excess profit of the pharmaceutical firm is exhausted. This implies that the firm must earn zero profit in the solution of the regulator's problem.

The first-order condition (16) can be solved together with the zero-profit condition $\pi(p, 0) = 0$ to obtain[3] the Ramsey price:

$$p_R = \frac{1}{2}\left(\eta + c - \sqrt{(\eta - c)^2 - 4\eta F}\right). \tag{17}$$

For the Ramsey price (17) to be well-defined, we must assume that

$$F < \frac{(\eta - c)^2}{4\eta}. \tag{18}$$

The assumption (18) is essential, because it guarantees that prices exist for which $\pi(p, 0) \geq 0$ and the regulator's strategy set is non-empty. Intuitively, the Ramsey price is sufficiently high so as to allow the firm to break even but it is lower than the monopoly price $(1/2)(\eta + c)$. The Ramsey price is related not only to the marginal or the fixed costs but also to the price elasticity of the demand (e.g. Armstrong and Sappington, 2007).

The solution is characterized by zero profits, which implies that social welfare equals the value of the consumer surplus. Therefore, the social welfare value in the Ramsey solution is

$$W_R = CS(p_R, 0) = \frac{1}{8\eta}\left(\eta - c + \sqrt{(\eta - c)^2 - 4\eta F}\right)^2. \tag{19}$$

---

[2] *The first derivative of social welfare with respect to price is* $-(p-c)\frac{1}{\eta}$ *and the statement follows from this.*

[3] *The solution of the first-order condition (16) and* $\pi(p, 0) = 0$ *defines the Ramsey price and the value of the Lagrange multiplier. The system of equations has two solutions,* $x_1 = (p_1, \mu_1)$ *and* $x_2 = (p_2, \mu_2)$*. The first (second) solution corresponds to the lower (higher) root of the zero profit condition. The value of social welfare is strictly decreasing at all price levels that exceed the marginal cost. Since the prices in the feasible set (i.e. prices which satisfy the profit constraint) are higher than the marginal cost, the lower root* x_1 *is the solution to the regulator's problem.*

We note from the Ramsey price that even if it eliminates excess profits, it forcefully limits the number of people who are able to buy the pharmaceutical.

4. Second-best efficient price and insurance policy

We next introduce public health insurance and ask whether adding a distortionary policy instrument to the regulator's strategy has the potential to improve social welfare. Intuitively, health insurance improves patients' welfare by lowering the out-of-pocket price that patients pay for the pharmaceutical, but the obvious social cost of health insurance is that it increases health insurance expenditures, which are financed through taxation. To examine whether the social benefits of public health insurance exceed social costs, we first derive the optimal price-insurance policy and thereafter assess its welfare properties.

The regulator's policy problem is to choose the price and insurance coverage (p, r) that maximize social welfare (10) subject to the profit constraint (11) and the feasibility constraints $p \geq 0$ and $0 \leq r \leq 1$. The solution of the regulator's problem is characterized in Proposition 1 below.

One of the features of the optimal price-insurance policy is that $\mu = \lambda$, where μ is the Lagrange multiplier of the profit constraint. To explain the logic of this result, we note that the multiplier μ measures the marginal social benefit of relaxing the firm's profit constraint, while λ is the marginal cost of tax funding. Since part of the firm's revenues are financed through the tax-funded health insurance expenditures $rpq(p, r)$, the regulator can use the price-insurance policy (p, r) to relax the firm's profit constraint. Proof of Proposition 1 in the Appendix shows that the optimal policy must satisfy the condition that the marginal social benefit of relaxing the profit constraint equals the marginal cost of tax funding.

Proposition 1. *If* $\lambda > 0$ *and*

$$\frac{(\eta - c)^2 \lambda (1 + \lambda)}{\eta (1 + 2\lambda)^2} < F, \tag{20}$$

the optimal price-insurance policy $(\tilde{p}, \tilde{r})$ *is*

$$\tilde{p} = c + \frac{\eta F (1 + 2\lambda)}{(\eta - c)(1 + \lambda)} \tag{21}$$

and

$$\tilde{r} = \frac{\eta F (1 + 2\lambda)^2 - (\eta - c)^2 \lambda (1 + \lambda)}{(1 + 2\lambda)\left[\eta F (1 + 2\lambda) + c(\eta - c)(1 + \lambda)\right]}. \tag{22}$$

Proof. See Appendix, Proof of Proposition 1.

The optimal price-insurance policy is designed so that it yields zero profit for the firm. The producer price $\tilde{p}$ exceeds the marginal cost of producing the pharmaceutical to cover the fixed *R&D* cost. The condition (20) guarantees that $\tilde{r} > 0$ and the optimal policy is an interior solution. If the condition was not satisfied, the necessary conditions of the regulator's problem (Appendix, Proof of Proposition 1) would support the Ramsey solution.

Proposition 2 below displays the effects of the fixed *R&D* cost and the quality weight on the optimal producer price and health insurance coverage. The results show that an increase in the fixed cost F leads to an increase in the optimal insurance coverage. Intuitively, this finding suggests that the regulator is more likely to introduce greater insurance coverage, the larger the fixed cost. Clearly, the insurance coverage allows the regulator to increase the consumer surplus by reducing the out-of-pocket price of the pharmaceutical. If health insurance was not available, an increase in the fixed cost would, on the contrary, increase the price of the pharmaceutical and decrease the demand for the pharmaceutical and consumer surplus.

Proposition 2. *Suppose that* $\lambda > 0$. *Then*

$$\frac{\partial \tilde{p}}{\partial F} = \frac{\eta(1+2\lambda)}{(\eta - c)(1+\lambda)} > 0; \; \frac{\partial \tilde{r}}{\partial F} = \frac{\eta(\eta - c)(1+\lambda)[\eta\lambda + c(1+\lambda)]}{[\eta F(1+2\lambda) + c(\eta - c)(1+\lambda)]^2} > 0$$

and

$$\frac{\partial \tilde{p}}{\partial \eta} = \frac{-cF(1+2\lambda)}{(\eta - c)^2(1+\lambda)} < 0; \; \frac{\partial \tilde{r}}{\partial \eta} = \frac{-(1+\lambda)\left[F(1+2\lambda)\left(\lambda\eta^2 + (1+\lambda)c^2\right) + c(\eta - c)^2\lambda(1+\lambda)\right]}{(1+2\lambda)\left[\eta F\left(1+2\lambda\right) + c(\eta - c)\left(1+\lambda\right)\right]^2} < 0.$$

Proof. See Appendix, Proof of Proposition 2.

The comparative statics results in Proposition 2 also show that a higher quality weight leads to reductions in both the optimal producer price and insurance coverage. An increase in the quality weight moves the inverse demand curve to the right and, in order to price the pharmaceutical according to average costs, the regulator responds by reducing the optimal producer price. At the same time, however, the regulator implements health insurance coverage that increases the patients' co-payment for the pharmaceutical. On the basis of these findings alone, the effect of a higher-quality weight on the out-of-pocket price remains inconclusive. However, our following analysis on the out-of-pocket price shows that a higher-quality weight leads to a higher consumer price for the pharmaceutical (Eq. 23).

The out-of-pocket price that patients pay in the optimal price-insurance policy is

(23) $$\tilde{p}(1-\tilde{r}) = c + \frac{(\eta - c)\lambda}{(1+2\lambda)}.$$

When taxation is distortionary and $\lambda > 0$, the consumer price exceeds the marginal cost of producing the pharmaceutical. This also implies that the demand for the pharmaceutical is below the first-best level, and

(24) $$q(\tilde{p}, \tilde{r}) = \frac{(\eta - c)(1+\lambda)}{\eta(1+2\lambda)} < \frac{\eta - c}{\eta} = q(c, 0).$$

In addition, one can prove that that the patient's out-of-pocket price (23) in the optimal price-insurance policy is lower than the Ramsey price (17), if the condition for the interior solution (20) holds true (Proof of Lemma 1, Online Appendix). Provided that the demand for the pharmaceutical (5) decreases as the out-of-pocket price increases, such a decrease in the out-of-pocket price also increases the consumption of the pharmaceutical beyond that in the Ramsey solution.

Next, we conduct the welfare analysis by evaluating the consumer surplus, the insurance expenditure and the level of social welfare in the optimal price-insurance policy. Table 1 displays these measures together with the corresponding measures in the first-best and Ramsey solutions. The consumer surplus associated with the optimal price-insurance policy is lower than the consumer surplus in the first-best solution with marginal cost pricing and no insurance coverage due to the positive marginal cost of taxation. On the contrary, it can be shown that the consumer surplus in the optimal price-insurance policy is higher than the consumer surplus in the Ramsey solution, if the condition for the interior solution (20) holds true (Proof of Lemma 2, Online Appendix). The underlying reason for this result is that the out-of-pocket price (23) is lower than the Ramsey price (17).

When the condition for the interior solution (20) is satisfied, the insurance expenditure in the optimal price-insurance policy is positive. Furthermore, we note that in the case of distortionary taxation, the expenditure is less than the fixed cost. On the other hand, when the marginal cost of taxation gets closer to zero, the insurance expenditure approaches the fixed cost. The intuition behind this relationship between the optimal insurance expenditure and the marginal cost of public funds is as follows: the higher (lower) is λ, the less (more) willing the regulator is to use taxation as a means to finance pharmaceutical expenditures via public health insurance.

Table 1: Consumer surplus, profit, insurance expenditure and social welfare

	First-best	Ramsey	Price-insurance policy
CS	$\frac{(\eta-c)^2}{2\eta}$	$\frac{1}{8\eta}\left(\eta-c+\sqrt{(\eta-c)^2-4\eta F}\right)$	$\frac{(\eta-c)^2}{2\eta}(\frac{1+\lambda}{1+2\lambda})^2$
π	0	0	0
IE	n.a.	n.a.	$F-\frac{(\eta-c)^2\lambda(1+\lambda)}{\eta(1+2\lambda)^2}$
W	$\frac{(\eta-c)^2}{2\eta}-F$	$\frac{1}{8\eta}\left(\eta-c+\sqrt{(\eta-c)^2-4\eta F}\right)$	$\frac{(\eta-c)^2}{2\eta}\frac{(1+\lambda)^2}{(1+2\lambda)}-F\left(1+\lambda\right)$

For the purpose of Proposition 3, we denote social welfare in the optimal price-insurance policy (Table 1) as follows:

(25) $$\tilde{W}=\frac{(\eta-c)^2}{2\eta}\frac{(1+\lambda)^2}{(1+2\lambda)}-F\left(1+\lambda\right).$$

The comparison of the social welfare in the first-best solution and in the optimal price-insurance policy does not to directly reveal that the first-best social welfare exceeds the social welfare in the optimal price-insurance policy (Table 1). However, Proposition 3 below demonstrates that – as expected – this indeed holds true.

Comparing the social welfare under the optimal price-insurance policy (25) with the social welfare in the Ramsey solution (19) leads to a striking observation. The introduction of public health insurance improves welfare because the resulting gain in the consumer surplus exceeds the increase in the publicly funded insurance expenditures (Proposition 3). This result is an illustration of the general theory of second best (Lipsey and Lancaster, 1956), where the introduction of a distortive policy instrument improves the welfare of an inefficient market.

The underlying reason for the finding that health insurance is welfare improving is the fact that the optimal health insurance in our model is combined with regulated producer prices. It is well known in health economics that if health insurance leads to higher prices in the health care market (Pauly, 1968; Feldstein, 1973), the introduction of health insurance is detrimental to welfare. In the context of our model, the introduction of health insurance will decrease the out-of-pocket price and increase the demand for the pharmaceutical but is also associated with a lower producer price due to economies of scale, hence leaving space for a possible welfare improvement (Gaynor et al., 2000).

Proposition 3. *The welfare ranking between the first-best solution, the Ramsey solution and the optimal price-insurance policy is the following:*

(26) $$\bar{W}_f>\tilde{W}>W_R.$$

Proof. See Appendix, Proof of Proposition 3.

Intuitively, the Ramsey solution produces a smaller welfare than the optimal price-insurance policy, because a great many people are not able to acquire the drug at Ramsey prices. The optimal policy $(\tilde{p},\tilde{r})$, however, does not reach an efficient solution because of the positive marginal cost of taxation.

5. Means-tested price-insurance policy

The previous analysis on the optimal price-insurance policy demonstrated how the introduction of health insurance can improve the efficiency of the pharmaceutical market. From the equity point of view, however, the optimal price-insurance policy has a serious limitation. Patients in the cohort of lowest incomes cannot afford to buy the pharmaceutical even in the presence of the health insurance. The number of such low-income patients is $1 - q(\tilde{p}, \tilde{r}) > 0$. Health is not like any other product, and equity considerations suggest that patients with low ability to pay should also have access to pharmaceutical treatment.

In this section, we examine an approach that adjusts the price-insurance policy to cope with vertical equity. In welfare economics, the idea of equity has been introduced in terms of the Rawlsian welfare criterion. Based on Rawls (1999), it is typically expressed as the maximin rule of the social choice[4]. Accordingly, the policy should aim at considering the utility of the individual who is worst off. In this section, the implications of the Rawlsian equity principle are examined in terms of a means-tested insurance policy that is implemented in the form of a third-degree price discrimination. In particular, we examine an optimal insurance policy that offers a higher insurance coverage for low-income patients who are not able to purchase the pharmaceutical at the out-of-pocket price paid by high-income patients. The advantage of the suggested approach is that it combines a solution for equity with an efficient insurance for those in higher income classes.

We analyse a model where people with high ability to pay and people with low ability to pay are entitled to different coverage rates, say $r_h \leq r_l$, where subscripts h and l refer to high ability to pay (high-income) and low ability to pay (low-income) patients, respectively. Hence, in this section, the focus will be on the price-insurance mechanism (p, r_l, r_h) with the feature $r_h \leq r_l$. Under this mechanism, the regulator offers the price p for the firm and selects the parameters of insurance coverage for high-income and low-income patients so that the out-of-pocket price of low-income patients is lower than that of high-income patients. Since the income variable is a continuous variable, we define low-income patients as the patient group who are not able to purchase a pharmaceutical at price p and insurance coverage r_h. This implies that the groups of low- and high-income patients are determined endogenously on the basis of the policy parameters (p, r_l, r_h) and raises particular questions about where to draw the demarcation lines between those who should have access to medication with price-insurance contract (p, r_l) and those with contract (p, r_h).

In what follows, we assume that the regulator has full information on patient incomes and hence is able to identify the low- and high-income patient groups and offer them different price-insurance contracts. If the regulator did not have full information on patient incomes and offered two price-insurance contracts (p, r_l) and (p, r_h), all patients in the market would prefer the contract offered to low-income patients because of the higher insurance coverage. As a result, the optimal contract to be derived next would not be incentive compatible. To make the contracts implementable, we assume that the regulator is fully informed about the patient incomes.

Given the price-insurance mechanism (p, r_l, r_h), the aggregate consumer surplus is given as follows:

$$CS(p, r_l, r_h) = \int_{\frac{p(1-r_l)}{\eta}}^{\frac{p(1-r_h)}{\eta}} (w\eta - (1 - r_l)p)\, dw + \int_{\frac{p(1-r_h)}{\eta}}^{1} (w\eta - (1 - r_h)\, p)\, dw. \tag{27}$$

[4] *The Rawlsian view has been widely discussed in welfare economics. For a recent analysis, one can refer to Stark, Jakubek, and Falniowski (2014), for example.*

Under this mechanism, the demand for the pharmaceutical is the sum of the demands of the buying high- and low-income patients:

$$q(p, r_l, r_h) = q_l(p, r_l, r_h) + q_h(p, r_l, r_h)$$

$$= \frac{p(1-r_h)}{\eta} - \frac{p(1-r_l)}{\eta} + 1 - \frac{p(1-r_h)}{\eta} = 1 - \frac{p(1-r_l)}{\eta}, \tag{28}$$

and the profit of the firm is given as follows:

$$\pi(p, r_l, r_h) = (p-c)q(p, r_1, r_h) - F. \tag{29}$$

Aggregate health insurance expenditures consist of the insurance reimbursements paid to subsidize the consumption of high- and low-income patients:

$$IE(p, r_l, r_h) = r_l p \left(\frac{p(1-r_h)}{\eta} - \frac{p(1-r_l)}{\eta} \right) + r_h p \left(1 - \frac{p(1-r_h)}{\eta} \right). \tag{30}$$

The regulator's policy problem is to choose the price and insurance policy (p, r_l, r_h) that maximizes social welfare (10) subject to the profit constraint $\pi(p, r_l, r_h) \geq 0$, the constraint on insurance coverage rates $r_h \leq r_l$, and the feasibility constraints $p \geq 0$ and $0 \leq r_t \leq 1$ for $t = l, h$. The consumer surplus, profit and insurance expenditures in the current problem are defined in expressions (27), (29) and (30), respectively. The following proposition characterizes the optimal means-tested price-insurance mechanism.

Proposition 4. *If* $\lambda > 0$ *and*

$$\frac{(\eta-c)^2 2(1+\lambda)(1+2\lambda)}{\eta(2+3\lambda)^2} < F, \tag{31}$$

the optimal means-tested price-insurance policy is

$$\hat{p} = c + \frac{\eta F\,(2+3\lambda)}{(\eta-c)\,2\,(1+\lambda)} \tag{32}$$

and

$$\hat{r}_l = \frac{\eta F(2+3\lambda)^2 - (\eta-c)^2\,2\lambda\,(1+\lambda)}{(2+3\lambda)\,[\eta F\,(2+3\lambda) + c(\eta-c)2(1+\lambda)]} \tag{33}$$

$$\hat{r}_h = \frac{\eta F\,(2+3\lambda)^2 - (\eta-c)^2\,2\,(1+\lambda)\,(1+2\lambda)}{(2+3\lambda)\,[\eta F\,(2+3\lambda) + c(\eta-c)2(1+\lambda)]}. \tag{34}$$

Proof. See Appendix, Proof of Proposition 4.

The insurance coverage of the low-income group (33) exceeds that of the high-income group (34) in the optimal means-tested price-insurance policy. This follows from the fact that $(1+\lambda)(1+2\lambda) > (1+\lambda)\lambda$. In addition, the comparison of the optimal prices $\hat{p}$ and $\tilde{p}$ demonstrates that $\hat{p} < \tilde{p}$, and the producer price in the means-tested price-insurance policy is lower than the price in the price-insurance policy (Section 4). Therefore, the introduction of the means-tested insurance coverage rates also have implications for the producer price of the pharmaceutical.

The above results will become explicit when we evaluate the welfare properties of the optimal means-tested price-insurance policy. The out-of-pocket price of high-income patients is

(35) $$\hat{p}(1-\hat{r}_h) = \frac{\eta(1+2\lambda)+c(1+\lambda)}{2+3\lambda},$$

and that of low-income patients is

(36) $$\hat{p}(1-\hat{r}_l) = \frac{\eta\lambda+2c(1+\lambda)}{2+3\lambda}.$$

Because $\hat{r}_l > \hat{r}_h$, the buying low-income patients pay less for the pharmaceutical than the buying high-income patients. Straightforward computation shows that the out-of-pocket price of high-income patients (and hence also the producer price) is higher than the monopoly price $(\eta + c)/2$. The out-of-pocket price of low-income patients exceeds marginal cost but is below the monopoly price. More strikingly, the optimal out-of-pocket payments ensure equal access to pharmaceutical treatment, and the low- and high-income patient groups consume the same amount of the pharmaceutical:

(37) $$q_l(\hat{p},\hat{r}_l,\hat{r}_h) = q_h(\hat{p},\hat{r}_l,\hat{r}_h) = \frac{(\eta-c)\,(1+\lambda)}{\eta(2+3\lambda)} \equiv x(\hat{p},\hat{r}_l,\hat{r}_h).$$

The aggregate consumption of the pharmaceutical is then $q(\hat{p}, \hat{r}_l, \hat{r}_h) = 2x(\hat{p}, \hat{r}_l, \hat{r}_h)$. The equal division of the market shows up also in the consumer surplus:

(38) $$CS_l(\hat{p},\hat{r}_l,\hat{r}_h) = CS_h(\hat{p},\hat{r}_l,\hat{r}_h) = \frac{(\eta-c)^2\,(1+\lambda)^2}{2\eta(2+3\lambda)^2} \equiv S^c(\hat{p},\hat{r}_l,\hat{r}_h).$$

The aggregate consumer surplus is $CS(\hat{p}, \hat{r}_l, \hat{r}_h) = 2S^c(\hat{p}, \hat{r}_l, \hat{r}_h)$. We state these findings as follows:

Proposition 5. *Under the Rawlsian principle of equity based on maximizing the aggregate consumer surplus and conditional on the better access to medication by low-income patients by means-tested insurance coverage, the final consumption of the pharmaceutical and the consumer surplus is split equally between low- and high-income patients.*

The result is sharp and it provides a yardstick when alternative equity principles are considered. Hence, and somewhat strikingly, although the patients with low ability to pay obtain the pharmaceutical at the lower out-of-pocket price, their surplus at the optimal solution is no higher than the surplus of the patients with high ability to pay.

By Proposition 5, the high- and low-income patient groups consume the same amount of the pharmaceutical. In addition, since the optimal insurance coverage of low-income patients is higher than that of high-income patients, insurance expenditures that the regulator pays to subsidize the consumption of the low-income group are higher than the corresponding expenditures of the high-income group:

(39) $$\hat{r}_l\hat{p}q_l(\hat{p},\hat{r}_l,\hat{r}_h) > \hat{r}_h\hat{p}q_h(\hat{p},\hat{r}_l,\hat{r}_h).$$

When evaluated in the optimal solution, the aggregate insurance expenditures amount to:

(40) $$IE(\hat{p},\hat{r}_l,\hat{r}_h) = F - \frac{(\eta-c)^2(1+\lambda)(1+3\lambda)}{\eta(2+3\lambda)^2}.$$

Straightforward comparison demonstrates that the insurance expenditure in the optimal means-tested price-insurance policy (40) is less than the insurance expenditure in the optimal price-insurance policy (Table 1). This is because different insurance coverage rates for low- and high-income patients allow the regulator to design financing schemes where the patients with high ability to pay pay a larger share of the pharmaceutical expenditures than the patients with low ability to pay.

Similarly as in the previous sections, the pharmaceutical firm earns zero profit (Proof of Proposition 4) in the means-tested price-insurance policy. The social welfare is then given as follows:

(41) $$\hat{W} = CS(\hat{p},\hat{r}_l,\hat{r}_h) - (1+\lambda)IE(\hat{p},\hat{r}_l,\hat{r}_h) = \frac{(\eta-c)^2\,(1+\lambda)^2}{\eta(2+3\lambda)} - (1+\lambda)F.$$

We next compare the social welfare obtained from the means-tested policy paying explicit attention to equity with the welfare obtained from the optimal price-insurance policy with no concern for low-income patients (Section 4).

Proposition 6. *The optimal means-tested price-insurance policy* $(\hat{p}, \hat{r}_l, \hat{r}_h)$ *yields a strictly higher welfare than the optimal price-insurance policy* $(\tilde{p}, \tilde{r})$ *and* $\hat{W} > \tilde{W}$.

Proof. That $\hat{W} > \tilde{W}$ follows directly from the fact $2 + 3\lambda < 2(1 + 2\lambda)$. □

Stated verbally, under the Rawlsian criterion, the social welfare exceeds the social welfare under the optimal price-insurance policy with a uniform coverage rate (Section 4). Third-degree out-of-pocket price discrimination in the form of means-tested insurance benefits increases the consumption possibilities of low-income patients. In addition, different insurance coverage rates for high- and low-income patients allow the regulator to design a price-insurance policy in which the aggregate insurance expenditure is lower than the insurance expenditure in the price-insurance policy with no means-testing. Both of these effects increase the social welfare in comparison to the situation in which no means-testing was available for the regulator.

6. Discussion and concluding remarks

The ability to pay for pharmaceuticals varies among people. A non-trivial fraction of people cannot afford to buy pharmaceutical products at unregulated market prices. Those products are created through expensive *R&D* programs. Expenses associated with *R&D* should subsequently be covered through prices, which, however, may turn out to be too high to be socially acceptable. In the current paper, the question has been raised of how to introduce means-tested subsidies to low-income citizens as part of an optimal regulation and also maintain incentives for pharmaceutical companies to invest in the *R&D*.

The paper has extended the previous work on the optimal price and health insurance regulation of pharmaceuticals in three ways. First, a market has been considered where the ability to pay for pharmaceuticals varies in the patient population. Second, the optimal price regulation together with insurance coverage has been derived and characterized. Third, price and insurance policies improving the access of low-income patients to pharmaceutical treatments has been explored.

The comparison of the social welfare under the optimal price-insurance policy with the social welfare obtained from the Ramsey solution with no health insurance led to a striking observation. The introduction of public health insurance improves welfare because the resulting gain in consumer surplus exceeds the increase in the publicly funded insurance expenditures. This result is an illustration of the general theory of second best where the introduction of a distortionary policy instrument improves welfare in an inefficient market.

The second-best policies explored do not, however, ensure full access to pharmaceutical products for all patients in the lowest income groups. To ensure full access, the regulator should choose full insurance for those low-income patients who are not able to personally finance the consumption of pharmaceuticals with the welfare maximizing price-insurance policy. Therefore, the implications of the Rawlsian equity principle were examined in terms of a means-tested price-insurance policy that is implemented in the form of a third-degree out-of-pocket price discrimination. In particular, we examined an optimal insurance policy that offers a higher insurance coverage for low-income patients. The advantage of the suggested approach is that it combines a solution for equity with efficient insurance for those in higher income classes.

Although it is known that full insurance may create inefficiency in the sense of excessive consumption (Pauly, 1968), the regulator is willing to implement such a policy if the improved access of low-income patients to pharmaceutical treatment is considered socially desirable. The improved access may have social value because of the resulting incremental health gains and the improved productivity of individuals in the labour market. Such social value may, however, not be based on patients' preferences (Brouwer et al., 2008).

In Kanniainen et al. (2020), we explored the welfare implications of price-insurance policies where low-income patients who cannot afford to purchase the pharmaceuticals at the current out-of-pocket price have access to free medication. The regulator assigns a social value to the pharmaceutical consumption of the low-income patients and evaluates the consumption of the pharmaceuticals of the high-income patients based on their consumer surplus. Our analysis suggested that the policy with free medication for low-income patients creates a conflict of interest between the high- and low-income groups. If the health gains in the low-income group are highly valued by the regulator, the out-of-pocket price can be increased leading to a decrease in the consumption and consumer surplus derived from the consumption of the pharmaceutical in the high-income group. The analysis indicates that the regulator wants to implement such a policy with free medication to the low-income group if the social value of health gains exceeds the social marginal cost of producing the pharmaceutical product.

There are other means of improving access to affordable medication besides policies relying on health insurance. One such tool is patent policy, defining the duration of the exclusive rights to sell the originator's drug. Potential entrants producing generic products have strong incentives to challenge such exclusive rights by entering the market before the expiry of the originator firm's patent, particularly when patents are long-lasting. Since the introduction of The Drug Price Competition and Patent Term Restoration Act in 1984, generic products in the US have been given a possibility to enter the pharmaceutical market before the expiry of the originator firm's patent (henceforth early entry). In such a case, a generic firm must file a paragraph IV certification claiming noninfringement or invalidity of the originator firm's patent (Branstetter et al., 2016.).

Izhak et al. (2020) explore the impact of patent length on the early entry of generic products using data from the US pharmaceutical market. They show that adding one year to patent length increases the early entry of generic products by five percentage points. Their findings are consistent with the literature on costly imitation (Gallini, 2002), suggesting that patents in the pharmaceutical sector should have a shorter duration and broader scope than in the current situation. When assessed from the perspective of patients in need of pharmaceutical treatments, a shorter patent duration implies earlier introduction of generic price competition, and also earlier access to affordable medication. The issue of patent length would serve as a fruitful topic for future research. Indeed, the early literature on patent length has suggested that rationally determined imitation makes socially optimal patents longer than what is suggested by models with non-strategic imitation (Kanniainen and Stenbacka, 2000).

Our modelling has some obvious limitations. We have worked with a model with linear demand for the pharmaceutical product and assumed uniform income distribution. In their analysis on public and private interaction, Laine and Ma (2017) illustrate what implications the assumption of a uniform distribution may have. Future work should therefore consider the possibility of generalizing our results. □

References

Abbott, T., and Vernon, J. (2005). *The Cost of U. S. Pharmaceutical Price Reductions: A Financial Simulation Model of R&D Decisions*. NBER Working Papers 11114.

Armstrong, M., and Sappington, D. (2007). "Recent Developments in the Theory of Regulation." In *Handbook of Industrial Organization.* Vol 3, 1557–1700. Eds. M. Armstrong and R. H. Porter.

Baicker, K., Mullainathan, S., and Schwartzstein, J. (2015). "Behavioral Hazard in Health Insurance." *The Quarterly Journal of Economics* 130(4), 1623–1667.

Barros, P. P., and Martinez-Giralt, X. (2008). "On International Cost Sharing of Pharmaceutical R&D." *International Journal of Health Care Finance and Economics* 8, 301–312.

Besley, T. (1988). "Optimal reimbursement health insurance and the theory of Ramsey taxation." *Journal of Health Economics* 7, 321–336.

Borges dos Santos, M.A., dos Santos Dias, L. L., Santos Pinto C. D. B., et al. (2019). "Factors influencing pharmaceutical pricing – a scoping review of academic literature in health science." *Journal of Pharmaceutical Policy and Practice* 2019, 12:24.

Branstetter, L., Chatterjee, C., and Higgins, M. J. (2016). "Regulation and welfare: evidence from paragraph IV generic entry in the pharmaceutical industry." *RAND Journal of Economics* 47(4), 857–890.

Brouwer, W. B. F., Culyer, A. J., van Excel N. Job A., and Rutten F. F. H. (2008). "Welfarism vs. extra-welfarism." *Journal of Health Economics* 28, 325–338.

Diamond, P. (1975). "A many-person Ramsey tax rule." *Journal of Public Economics* 4, 335–342.

Feldstein, M. S. (1973). "The welfare loss of excess health insurance." *Journal of Political Economy* 81, 251–280.

Gallini, N. T. (1992). "Patent policy and costly imitation." RAND Journal of Economics 23(1), 52–63.

Gaynor, M., Haas-Wilson, D., and Vogt, W. B. (2000). "Are invisible hands good hands? Moral hazard, competition, and the second-best in health care markets." *Journal of Political Economy* 108, 992–1005.

Grassi, S., and Ma, C.-t. A. (2011). "Optimal public rationing and price response." *Journal of Health Economics* 30, 1197–1206.

Grassi, S. and Ma, C.-t. A. (2012). "Public sector rationing and private sector selection." *Journal of Public Economic Theory* 14, 1–34.

Grossman, M. (1972). "On the Concept of Health Capital and the Demand for Health." *Journal of Political Economy* 80(2), 223–255.

Izhak, O., Saxell, T., and Takalo T. (2020). Optimal patent policy for Pharmaceutical Industry., VATT Working Papers 131.

Kanniainen, V., and Stenbacka, R. (2000). "Endogenous Imitation and Implications for Technology Policy." *Journal of Institutional and Theoretical Economics* 156(2), 360–381.

Kanniainen, V., Laine, J., and Linnosmaa I. (2020). *Pricing the Pharmaceuticals when the Ability to Pay Differs: Taking Vertical Equity Seriously*. CESifo Working Paper Series 8031, CESifo.

Laffont, J-J., and Martimort, D. (2002). *Theory of Incentives: The Principal-Agent Model*. New Jersey: Princeton University Press.

Laine, L., and Ma, C-t. (2017). "Quality and competition between public and private firms." Journal of Economic Behaviour and Organization 140, 336–353.

Lipsey, R. G., and Lancaster, K. (1956). "The General Theory of Second Best." *Review of Economic Studies* 24(1), 11–32.

Pauly, M. V. (1968). "The Economics of Moral Hazard: Comment." *The American Economic Review* 58(3), 531–537.

Ramsey, F. P. (1927). "A contribution to the theory of taxation." *Economic Journal* 37, 47–61.

Rawls, J. (1999). A Theory of Justice. Cambridge, MA: Harvard University Press.

Stark, O., Jakubek, M., and Falniowski, F. (2014). "Reconciling the Rawlsian and the utilitarian approaches to the maximization of social welfare." Economics Letters 122, 439–444.

Vogler, S., Zimmermann, N., and Haasis M. A. (2019). PPRI Report 2018 – Pharmaceutical pricing and reimbursement policies in 47 PPRI network member countries. WHO Collaborating Centre for Pricing and Reimbursement Policies, Gesundheit Österreich GmbH (GÖG / Austrian National Public Health Institute), Vienna. http://whocc.goeg.at/Publications/CountryReports

Appendix

Result 1: Derivation of the willingness to pay function $\theta(w)$.

Let us consider Cobb-Douglas utility function with constant returns to scale

$$u = x^{1-\alpha}h^{\alpha},$$

where $0 < \alpha < 1$ is the preference weight that patients give to health. The logarithmic transformation of the utility function enables one to present the patients' utility in a separable form:

$$\ln(u) \equiv \tilde{u} = (1-\alpha)\ln(x) + \alpha \ln(h).$$

Given the budget constraint $w = x + (1 - r)pj$, the utility that the patient with income w obtains from the consumption of the pharmaceutical is

$$(1-\alpha)\ln(w-(1-r)pj) + \alpha\ln(h_0 + \Delta j),$$

where $j = 1, 0$. The patient is indifferent between consuming and not consuming the pharmaceutical if

$$(1-\alpha)\ln(w-\theta) + \alpha\ln(h_1) = (1-\alpha)\ln(w) + \alpha\ln(h_0), \tag{42}$$

where θ is the patient's willingness to pay for the pharmaceutical. The equation (42) can be rearranged to obtain

$$\ln\left(\frac{w-\theta}{w}\right) = \ln\left(\left(\frac{h_0}{h_1}\right)^{\frac{\alpha}{1-\alpha}}\right)$$

or

$$1 - \frac{\theta}{w} = \left(\frac{h_0}{h_1}\right)^{\frac{\alpha}{1-\alpha}}. \tag{43}$$

The equation (43) can be solved with respect to θ to obtain

$$\theta = w\left[1 - \left(\frac{h_0}{h_1}\right)^{\frac{\alpha}{1-\alpha}}\right] = w\left[1 - \left(1 + \frac{\Delta}{h_0}\right)^{\frac{-\alpha}{1-\alpha}}\right], \tag{44}$$

where the last equality follows from the fact that $\frac{h_0}{h_1} = \frac{1}{1+\frac{\Delta}{h_0}}$. □

Proof of Proposition 1. The regulator's problem is to find the price-insurance policy (p, r) which maximizes social welfare

$$W = CS(p,r) + \pi(p,r) - (1+\lambda)IE(p,r)$$

subject to the profit constraint

$$-\pi(p,r) \leq 0$$

and the feasibility constraints

$$p \geq 0$$

$$0 \leq r \leq 1.$$

The above problem is called an original problem. In what follows, we analyse the solutions of the original problem without the feasibility constraints. Such a problem is called a relaxed problem. This approach to finding the solution to the regulator's problem through the relaxed problem rests on the intuition that, if solutions of the relaxed problem also satisfy the feasibility constraints, they must also solve the original problem. This approach has become a standard analytical tool in the principal-agent literature (e.g. Laffont and Martimort, 2002).

We assume throughout this proof that the conditions $\lambda > 0$ and

$$\frac{(\eta - c)^2\lambda(1+\lambda)}{\eta(1+2\lambda)^2} < F \tag{45}$$

hold true. Let $(\tilde{p}, \tilde{r})$ denote the price-insurance policy that solves the relaxed problem and μ be the Lagrange multiplier of the profit constraint. The Lagrangian function of the relaxed problem is given as follows:

$$L = CS(p,r) + (1+\mu)\,\pi(p,r) - (1+\lambda)IE(p,r).$$

The solution of the relaxed problem must satisfy the first-order conditions:

$$\begin{aligned} \frac{\partial L}{\partial p} = &-(1-r)\left[1 - \frac{p(1-r)}{\eta}\right] + (1+\mu)\left[1 - \frac{2p(1-r)}{\eta} + \frac{(1-r)c}{\eta}\right] \\ &-(1+\lambda)\,r\left[1 - \frac{2p(1-r)}{\eta}\right] = 0 \end{aligned} \tag{46}$$

$$\begin{aligned} \frac{\partial L}{\partial r} = &\;p\left[1 - \frac{p(1-r)}{\eta}\right] + (1+\mu)(p-c)\frac{p}{\eta} \\ &-(1+\lambda)\,p\left[1 - \frac{p(1-r)}{\eta} + \frac{pr}{\eta}\right] = 0. \end{aligned} \tag{47}$$

Moreover, the solution must satisfy the profit constraint and the complementary slackness conditions $-\pi(p, r) \leq 0$, $\mu \geq 0$ and

$$\mu\left[F - (p-c)\left(1 - \frac{p(1-r)}{\eta}\right)\right] = 0. \tag{48}$$

Lemma 1.1 If $(\tilde{p}, \tilde{r}, \tilde{\mu})$ solves the relaxed problem, then $\tilde{\mu} = \lambda$.

Proof. Contrary to the claim, suppose that $\mu \neq \lambda$ in the solution of the relaxed problem. Then the first-order conditions (46) and (47) have two solutions. The first solution is $\bar{p} = 0$ and $\bar{r} = [\eta\mu + c(1+\mu)]/[\eta\lambda + c(1+\mu)]$ and the second solution is $\check{p} = [\eta(1+\lambda) - c(1+\mu)]/[\lambda - \mu]$ and $\check{r} = [(\eta - c)(1+\mu)]/[\eta(1+\lambda) - c(1+\mu)]$. When evaluated at these two solutions, the profit of the firm is $-\pi(\bar{p}, \bar{r}) = c + F$ and $-\pi(\check{p}, \check{r}) = F$, respectively. Therefore, the solutions of the first-order conditions (46) and (47) never satisfy the profit constraint. This implies that, if $\mu \neq \lambda$, there is no price-insurance pair which would satisfy the necessary conditions of the relaxed problem. For solutions to exist, we must therefore have $\mu = \lambda$. □

Lemma 1.2 If $(\tilde{p}, \tilde{r}, \tilde{\mu})$ solves the relaxed problem, then any pair $(\tilde{p}, \tilde{r})$ satisfying

$$p = \frac{\eta\lambda + c(1+\lambda)}{(1-r)(1+2\lambda)} \tag{49}$$

satisfies both first-order conditions (46) and (47).

Proof. Suppose that $(\tilde{p}, \tilde{r}, \tilde{\mu})$ solves the relaxed problem. Then, the first-order condition (46) holds true for any pair (p, r) for which

$$p = \frac{\eta\tilde{\mu} + c(1+\tilde{\mu}) - r\left[\eta\lambda + c(1+\tilde{\mu})\right]}{(1-r)\left[1+2\tilde{\mu} - r(1+2\lambda)\right]} \tag{50}$$

and the first-order condition (47) is satisfied for any pair (p, r) for which

$$p = \frac{\eta\lambda + c(1+\tilde{\mu})}{1+\tilde{\mu}+\lambda - r(1+2\lambda)} \quad \text{or} \quad p = 0. \tag{51}$$

The solution $p = 0$ can be ruled out because it does not satisfy the profit constraint. By Lemma 1.1, the solution of the relaxed problem must satisfy $\tilde{\mu} = \lambda$. Evaluating the right-hand sides of the equations (50) and (51) at $\tilde{\mu} = \lambda$ yields the equation (49). □

Let us then characterize the solution of the problem. By Lemma 1.1 and the assumption $\lambda > 0$, we must have $\tilde{\mu} = \lambda > 0$. Then the complementary slackness conditions imply that the zero-profit condition $\pi(p, r) = 0$ must hold true at the solution of the regulator's problem. Solving the first-order condition (46) or (47) together with the zero-profit condition yields the optimal price and insurance coverage and the value of the Lagrange multiplier:

$$\tilde{p} = c + \frac{\eta F(1+2\lambda)}{(\eta - c)(1+\lambda)}$$

$$\tilde{r} = \frac{\eta F(1+2\lambda)^2 - (\eta-c)^2\lambda(1+\lambda)}{(1+2\lambda)\left[\eta F\left(1+2\lambda\right) + c(\eta-c)\left(1+\lambda\right)\right]}$$

$$\tilde{\mu} = \lambda.$$

When evaluated at the point $(\tilde{p}, \tilde{r}, \tilde{\mu})$, the determinant of the bordered Hessian matrix is

$$\left|\bar{H}\right| = \frac{\left[c(\eta-c)(1+\lambda) + \eta F\left(1+2\lambda\right)\right]^2}{\eta^3(1+2\lambda)} > 0,$$

which proves that the optimal policy is a local maximum.

Let us finally check that the solution of the relaxed problem satisfies the feasibility conditions. By straightforward calculation, one can demonstrate that the optimal price-insurance policy satisfies the condition $\tilde{r} < 1$. In addition, it holds true that $\tilde{r} > 0$, because the fixed cost satisfies the condition (45). Since the optimal price $\tilde{p}$ is strictly positive, the solution satisfies the feasibility conditions of the original problem. □

Proof of Proposition 2. Let us define

$$A(\eta, F) \equiv \eta F(1+2\lambda)^2 - (\eta-c)^2\lambda(1+\lambda) \text{ and}$$

$$B(\eta, F) \equiv (1+2\lambda)\left[\eta F\left(1+2\lambda\right) + c(\eta-c)\left(1+\lambda\right)\right].$$

and let A_η (B_η) and A_F (B_F) denote the partial derivatives of A (B) with respect to η and F.

First, we skip the proof for the expression $\frac{\partial \tilde{p}}{\partial F}$. Secondly, the partial derivative of optimal insurance coverage with respect to the fixed cost is

$$\frac{\partial \tilde{r}}{\partial F} = \frac{A_F B(\eta, F) - B_F A(\eta, F)}{B(\eta, F)^2} = \frac{\eta(1+2\lambda)^2 \left[B(\eta, F) - A(\eta, F)\right]}{(1+2\lambda)^2 \left[\eta F (1+2\lambda) + c(\eta - c)(1+\lambda)\right]^2}$$

$$= \frac{\eta(\eta - c)(1+\lambda)\left(\eta\lambda + c(1+\lambda)\right)}{\left[\eta F (1+2\lambda) + c(\eta - c)(1+\lambda)\right]^2} > 0.$$

Thirdly, the partial derivative of the optimal producer price with respect to the quality weight is

$$\frac{\partial \tilde{p}}{\partial \eta} = \frac{F(1+2\lambda)(\eta - c)(1+\lambda) - (1+\lambda)\left(\eta F(1+2\lambda)\right)}{(\eta - c)^2 (1+\lambda)^2} = \frac{-cF(1+2\lambda)}{(1+\lambda)(\eta - c)^2} < 0.$$

Finally, the partial derivative of the optimal insurance coverage with respect to the quality weight is

$$\frac{\partial \tilde{r}}{\partial \eta} = \frac{A_\eta B(\eta, F) - B_\eta A(\eta, F)}{B(\eta, F)^2} \tag{52}$$

After some lengthy calculations, the above partial derivative (52) simplifies to

$$\frac{\partial \tilde{r}}{\partial \eta} = \frac{-(1+\lambda)\left[F(1+2\lambda)\left(\lambda\eta^2 + (1+\lambda)c^2\right) + c(\eta - c)^2 \lambda(1+\lambda)\right]}{(1+2\lambda)\left[\eta F (1+2\lambda) + c(\eta - c)(1+\lambda)\right]^2} < 0,$$

completing the proof. □

Proof of Proposition 3. Let us assume that $\lambda > 0$ and that the fixed cost satisfies the conditions

$$\frac{(\eta - c)^2 \lambda(1+\lambda)}{\eta(1+2\lambda)^2} < F < \frac{(\eta - c)^2}{4\eta}. \tag{53}$$

The above conditions (53) have two implications: first, they ensure that the Ramsey price is well-defined and, secondly, the conditions imply that the optimal price-insurance policy is an interior solution.

We first prove that $\tilde{W} > W_R$. Define the welfare difference

$$DW(F) \equiv \tilde{W} - W_R$$

$$= \frac{(\eta - c)^2}{2\eta} \frac{(1+\lambda)^2}{1+2\lambda} - F(1+\lambda) - \frac{1}{8\eta}\left(\eta - c + \sqrt{(\eta - c)^2 - 4\eta F}\right)^2.$$

The first derivative of the welfare difference with respect to the fixed cost F is given as

$$DW'(F) = -(1+\lambda) + \frac{\eta - c + \sqrt{(\eta - c)^2 - 4\eta F}}{2\sqrt{(\eta - c)^2 - 4\eta F}},$$

and the second derivative is

$$DW''(F) = \frac{\eta(\eta - c)}{\left(\sqrt{(\eta - c)^2 - 4\eta F}\right)^3} > 0.$$

Therefore, the welfare difference is a strictly convex function of the fixed cost F. The strict convexity of the function $DW(F)$ implies that the unconstrained minimum of the welfare difference must be unique. Solving the first-order condition $DW'(F) = 0$ with respect to F yields the minimum point

(54)
$$F_1 = \frac{(\eta - c)^2}{\eta}\frac{\lambda(1+\lambda)}{(1+2\lambda)^2} \geq 0,$$

which corresponds to the infimum of the interval of the fixed cost (53). This implies that $DW(F) > DW(F_1)$ for all values of the fixed cost that satisfy the condition (53). When evaluated at the minimum point, the value of the welfare difference is zero:

$$\begin{aligned} DW(F_1) &= \frac{(\eta-c)^2}{2\eta}\frac{(1+\lambda)^2}{1+2\lambda} - F_1(1+\lambda) - \frac{1}{8\eta}\left(\eta - c + \sqrt{(\eta-c)^2 - 4\eta F_1}\right)^2 \\ &= \frac{(\eta-c)^2}{2\eta}\left(\frac{1+\lambda}{1+2\lambda}\right)^2(1+2\lambda) - \frac{(\eta-c)^2}{2\eta}\left(\frac{1+\lambda}{1+2\lambda}\right)^2(2\lambda+1) \\ &= 0. \end{aligned}$$

These observations imply that $DW(F) > DW(F_1) = 0$ and $\tilde{W} > W_R$ for all fixed costs satisfying the conditions (53).

Secondly, we have $\bar{W}_f \geq \tilde{W}$ when

(55)
$$\frac{(\eta-c)^2}{2\eta} - F \geq \frac{(\eta-c)^2}{2\eta}\frac{(1+\lambda)^2}{(1+2\lambda)} - F(1+\lambda),$$

which implies that

$$\frac{(\eta-c)^2}{2\eta}\frac{\lambda}{1+2\lambda} \leq F.$$

But now

(56)
$$\frac{(\eta-c)^2}{2\eta}\frac{\lambda}{1+2\lambda} = \frac{(\eta-c)^2}{2\eta}\frac{\lambda(1+2\lambda)}{(1+2\lambda)^2} < \frac{(\eta-c)^2}{\eta}\frac{\lambda(1+\lambda)}{(1+2\lambda)^2},$$

where the last expression corresponds to the infimum of the set of feasible fixed costs (53). Hence, the condition (55) is satisfied as a strict inequality when (53) holds true and $\lambda > 0$, which verifies that $\bar{W}_f > \tilde{W}$. □

Proof of Proposition 4. Let us assume that the conditions $\lambda > 0$ and

(57)
$$\frac{(\eta-c)^2 2(1+\lambda)(1+2\lambda)}{\eta(2+3\lambda)^2} < F$$

hold true in this proof.

As above in the proof of Proposition 1, we will start by analysing the relaxed problem in which the feasibility constraints $p \geq 0$ and $0 \leq r_t \leq 1$ for $t = l, h$ are initially ignored. The Lagrangian function of the relaxed problem is given as follows

$$L = CS(p, r_l, r_h) + (1+\mu)\,\pi(p, r_l, r_h) - (1+\lambda)\,IE(p, r_l, r_h) - \kappa\,(r_h - r_l),$$

where the consumer surplus is

$$CS(p,r_l,r_h)=\int_{\frac{p(1-r_l)}{\eta}}^{\frac{p(1-r_h)}{\eta}}(w\eta-(1-r_l)p)\,dw+\int_{\frac{p(1-r_h)}{\eta}}^{1}(w\eta-(1-r_h)\,p)\,dw,$$

the firm's profit is

$$\pi(p,r_l,r_h)=(p-c)\left(1-\frac{p(1-r_l)}{\eta}\right)-F,$$

and the insurance expenditures are

$$IE(p,r_l,r_h)=r_lp\left(\frac{p(1-r_h)}{\eta}-\frac{p(1-r_l)}{\eta}\right)+r_hp\left(1-\frac{p(1-r_h)}{\eta}\right),$$

and κ is the multiplier of the constraint $r_h \leq r_l$.

The solution of the relaxed problem $(\hat{p}, \hat{r}_l, \hat{r}_h, \hat{\mu}, \hat{\kappa})$ must satisfy the first-order conditions:

$$\frac{\partial L}{\partial p}=\frac{p\,(r_l-r_h)^2}{\eta}-(1-r_h)\left[1-\frac{p\,(1-r_h)}{\eta}\right]$$

$$+(1+\mu)\left[1-\frac{2p(1-r_l)}{\eta}+\frac{(1-r_l)c}{\eta}\right]$$

(58)
$$-(1+\lambda)\left[r_l\left(\frac{2p(r_l-r_h)}{\eta}\right)+r_h\left(1-\frac{2p(1-r_h)}{\eta}\right)\right]=0$$

$$\frac{\partial L}{\partial r_h}=\frac{-p^2(r_l-r_h)}{\eta}+p\left(1-\frac{p(1-r_h)}{\eta}\right)$$

(59)
$$-(1+\lambda)\left[\frac{-p^2(r_l-r_h)}{\eta}+p\left(1-\frac{p(1-r_h)}{\eta}\right)\right]-\kappa=0$$

$$\frac{\partial L}{\partial r_l}=\frac{p^2(r_l-r_h)}{\eta}+(1+\mu)(p-c)\frac{p}{\eta}$$

(60)
$$-(1+\lambda)\left(\frac{p^2(r_l-r_h)}{\eta}+\frac{p^2r_l}{\eta}\right)+\kappa=0$$

Moreover, the solution must satisfy the profit constraint and its complementary slackness conditions $-\pi(p, r_l, r_h) \leq 0$, $\mu \geq 0$ and

(61)
$$\mu\left[F-(p-c)\left(1-\frac{p(1-r_l)}{\eta}\right)\right]=0$$

and the means-testing constraint $r_h \leq r_l$ and its complementary slackness conditions $r_h - r_l \leq 0$, $\kappa \geq 0$ and

(62)
$$\kappa(r_h-r_l)=0.$$

From the perspective of the ensuing analysis it is important to note that effective means-testing, i.e. $r_h < r_l$, occurs in the solution of the regulator's problem only if $\kappa = 0$. If this is not the case and $\kappa > 0$ then by the condition (62) we must have $r_h = r_l$ in the optimal solution. Both low- and high-income patients receive the same insurance reimbursement and means-testing does not take place. Furthermore, it is straightforward to show that the necessary conditions of the problem simplify to those of the optimal price-insurance policy examined in Section 4. Therefore, the following analysis concentrates on the means-testing solution in which $\kappa = 0$.

Lemma 4.1 If $(\hat{p}, \bar{r}_l, \bar{p}_h, \hat{\mu})$ solves the relaxed problem, then $\hat{\mu} = \lambda$.

Proof. Contrary to the claim, suppose that $\mu \neq \lambda$ in the solution of the relaxed problem. Then the first-order conditions (58), (59) and (60) have two solutions $(\bar{p}, \bar{r}_l, \bar{r}_h)$ and $(\check{p}, \check{r}_l, \check{r}_h)$. In the first solution $\bar{p} = 0$ and insurance coverage rates must satisfy the condition (multiple solutions)

$$\bar{r}_h = \frac{1}{\lambda}[\mu + (1+\mu)\frac{c}{\eta}(1-\bar{r}_l)]$$

In the second solution $\check{p} = [\eta(1+\lambda)-c(1+\mu)]/[\lambda-\mu]$ and $\check{r}_h = \check{r}_l = [(\eta-c)(1+\mu)]/[\eta(1+\lambda)-c(1+\mu)]$. When evaluated at these two solutions, the profit of the firm is $-\pi(\bar{p}, \bar{r}_l, \bar{r}_h) = c + F$ and $-\pi(\check{p}, \check{r}_l, \check{r}_h) = F$, respectively. Therefore, the solutions of the first-order conditions (58), (59) and (60) never satisfy the profit constraint. This implies that, if $\mu \neq \lambda$, there are no price-insurance policies that would satisfy the necessary conditions of the relaxed problem. For solutions to exist, we must have $\mu = \lambda$. □

Let us then derive the solution to the regulator's problem. By Lemma 4.1 and the assumption $\lambda > 0$, we must have $\hat{\mu} = \lambda > 0$. Complementary slackness conditions for the profit constraint then imply that $\pi(p, r_l, r_h) = 0$. Solving first-order conditions (58), (59) and (60) together with the zero-profit condition yields the means-tested price-insurance policy and the value of the Lagrange multiplier:

$$\hat{p} = c + \frac{\eta F(2+3\lambda)}{(\eta-c)2(1+\lambda)}$$

$$\hat{r}_l = \frac{\eta F(2+3\lambda)^2 - (\eta-c)^2\, 2\lambda\,(1+\lambda)}{(2+3\lambda)\,[\eta F(2+3\lambda) + c(\eta-c)2(1+\lambda)]}$$

$$\hat{r}_h = \frac{\eta F(2+3\lambda)^2 - (\eta-c)^2\, 2\,(1+\lambda)\,(1+2\lambda)}{(2+3\lambda)\,[\eta F(2+3\lambda) + c(\eta-c)2(1+\lambda)]}$$

$$\hat{\mu} = \lambda.$$

It is worth noting that, because $\lambda < 1 + 2\lambda$, there is effective means-testing and $\hat{r}_h < \hat{r}_l$.

Lemma 4.2 The above solution of the relaxed problem is a local maximum.

Proof. To check that the means-tested price-insurance policy derived above is a local maximum, first note that the relevant bordered Hessian is a 4×4 matrix with the profit constraint binding. When evaluated at the solution of the problem, the determinants of the last two (ie. $n - k = 3 - 1 = 2$) leading principal minors of the bordered Hessian are

$$|\bar{H}_4| = \frac{-\lambda\,[2c(\eta-c)(1+\lambda) + F\eta(2+3\lambda)]^4}{4(\eta-c)^2\eta^4(1+\lambda)^2(2+3\lambda)} < 0, \tag{63}$$

and

$$|\bar{H}_3| = \frac{A(F)}{2(\eta-c)^2\eta^3(1+\lambda)^2(2+3\lambda)^2}, \tag{64}$$

where $A(F) \equiv B + CF + DF^2$ is a quadratic function in the fixed cost. The expressions for B, C and D are defined as follows:

$$B \equiv (1+\lambda)^4(1+2\lambda)\left[8c^2(c^4+6c^2\eta^2+\eta^4)-32c^3\eta(c^2+\eta^2)\right]$$

$$C \equiv 4\eta c(1+\lambda)^2(2+3\lambda)[-c^3(1+\lambda)(2+5\lambda)+\eta^3(2+\lambda)(1+2\lambda)$$

$$-\eta^2 c(6+\lambda(17+9\lambda))+\eta c^2(6+\lambda(19+12\lambda))]$$

$$D \equiv \eta^2(2+3\lambda)^2[c^2(1+\lambda)^2(2+7\lambda)-2c\eta(1+\lambda)(2+\lambda(5+\lambda))$$

$$+\eta^2(2+\lambda(7+4\lambda(2+\lambda)))]$$

To show that the proposed solution is a local maximum point, we need to show that $\bar{H}_3 > 0$. To do this it suffices to show that $A(F) > 0$ for all relevant values. We do this in two steps.

Step 1. First, we first show that $A(F)$ is a strictly convex function of the fixed cost by showing that $D > 0$ for all relevant values. Define $D_p(\eta) \equiv \frac{D}{\eta^2(2+3\lambda)^2}$. Then

$$D_p(\eta) = c^2(1+\lambda)^2(2+7\lambda)-2c\eta(1+\lambda)(2+\lambda(5+\lambda))+\eta^2(2+\lambda(7+4\lambda(2+\lambda)))$$

Since $\eta^2(2+3\lambda)^2 > 0$, to prove that $D > 0$ it suffices to demonstrate that $D_p(\eta) > 0$ for all relevant values of η. The expression $D_p(\eta)$ is a strictly convex function in η with a unique minimum point η_m, which can be found by solving the condition $D'_p(\eta) = 0$ with respect to η. When evaluated at η_m, the value of $D_p(\eta)$ is

$$D_p(\eta_m) = \frac{c^2\lambda(1+\lambda)^2(2+3\lambda)^3}{2+\lambda(7+4\lambda(2+\lambda))} > 0$$

where the strict inequality holds true by the assumptions $c > 0$ and $\lambda > 0$. This implies that $D_p(\eta) \geq D_p(\eta_m) > 0$ for all η and hence also for parameter values $\eta > c$.

Step 2. By the first step, the expression $A(F)$ has a unique minimum point with respect to F, denoted as F_m, which can be found by solving the condition $A'(F) = 0$ with respect to F. When evaluated at the minimum point, the value of the function $A(F)$ is

$$A(F_m) = \frac{4c^2(\eta-c)^2\lambda(1+\lambda)^4(2+3\lambda)(c(1+\lambda)+\eta(1+2\lambda))^2}{E}$$

where

$$E \equiv c^2(1+\lambda)^2(2+7\lambda)-2c\eta(1+\lambda)(2+\lambda(5+\lambda))+\eta^2(2+\lambda(7+4\lambda(2+\lambda))).$$

Note that $E = D_p(\eta)$. Step 1 above thus showed that the denominator of $A(F_m)$ is strictly positive for all relevant parameter values. Similarly, the numerator of the $A(F_m)$ is strictly positive because $c > 0$ and $\lambda > 0$. Therefore $A(F) > 0$ and $|\bar{H}_3| > 0$ when $\lambda > 0$ and $c > 0$, which verifies that the solution of the regulator's relaxed problem is a local maximum. □

To check that the solution of the relaxed problem $(\hat{p}, \hat{r}_l, \hat{r}_h)$ satisfies the feasibility conditions, note that the solution satisfies the constraint $p \geq 0$. The condition (57) ensures that $\hat{r}_h > 0$. That $\hat{r}_l > 0$ follows then from the fact that $\hat{r}_h < \hat{r}_l$. Straightforward calculation shows that $\hat{r}_t < 1$ for both $t = l, h$. Hence, the solution of the relaxed problem also solves the original problem. □

The anticipation effect of a light rail transit line on housing prices in the Helsinki region*

Eetu Kauria
Department of Economics, University of Turku
eekaur@utu.fi

Abstract

I analyze how housing markets in Espoo and Helsinki anticipate the construction of a new light rail transit connection called Jokeri Light Rail. I use geocoded micro-level housing transaction data from 2003–2019. As an econometric identification strategy, I utilize difference-in-differences estimation with a hedonic price model. My main result is that, on average, apartment prices increase by 5 percent more within 800 meters of the Jokeri Light Rail stops than apartments farther away. A rough estimate of the total windfall for homeowners indicates that the anticipated benefits exceed the cost estimate for the investment five to eight years before the Jokeri Light Rail becomes operational.

Keywords: *anticipation effect, difference-in-differences, housing market, light rail transit*
JEL codes: *D61, R41*

I am grateful to Janne Tukiainen, Oskari Harjunen, Elias Oikarinen, Tuukka Saarimaa, two anonymous referees, and Editor Mika Kortelainen for their helpful comments along the way.

1. Introduction

Public services significantly impact citizens' well-being and satisfaction in a particular neighborhood (James 2009). As public investments enhance a neighborhood's service level, the willingness to pay (WTP) for housing near the investment should increase. This leads to higher demand for land and housing near the investment (Alonso 1964; Muth 1969). Consequently, housing prices grow depending on the elasticity of supply for housing. For the policy-makers, whose interest should be to maximize their respective citizens' utility within a limited budget, it is crucial to assess the costs and benefits of various public investment projects and only execute the most effective ones.

In this paper, I assess the impact of a light rail transit (LRT) investment on housing prices. Specifically, I consider the price changes to reflect the desirability of an area. *Ergo*, I deem the changes in monetary values to be an indicator of the public's willingness to pay for the rail investment (see Gibbons and Machin 2005).

This study investigates how housing markets in Espoo and Helsinki react to the decision to build Jokeri Light Rail, a rail line running 25 kilometers from Keilaniemi in the west to Itäkeskus in the east. To my knowledge, this is the first study concerning the subject. To identify the price responses to Jokeri Ligh Rail, I employ rich housing transaction data and combine a difference-in-differences identification strategy with a hedonic price model[1]. My main result shows that the prices of apartments near the investment increase by 5 percent more than apartments farther away.

Moreover, I provide a cost-benefit analysis (CBA) of the rail investment, where I utilize the housing market effect as an indicator of the investment's benefits (see Weisbrod et al. 2016). The CBA of this paper is the first assessment of Jokeri Light Rail that accounts for the overall benefits of the investment so far as I am aware. An earlier CBA, which did not consider the impact on the housing market but rather only on traffic, emissions and accident rate, found the investment's costs to surpass its benefits (Project assessment of Jokeri Light Rail 2019).

I do not explicitly analyze the specific channels through which the rail investment impacts the housing market. For example, the expected urban development might affect the properties more than the initial accessibility improvements. In addition, a rail investment can affect other markets as well, including the labor market. The construction phase of the investment may also cause traffic frictions (see Lewis and Bajari 2011). However, if the housing markets are efficient, all the costs and benefits from a rail investment are reflected in dwelling prices (Gibbons and Machin 2005).

The impact of rail investment is a rather complex problem to analyze. The identification strategy in previous studies varies: some rely merely on the association between the investment and housing prices based on hedonic models, which gives ambiguous results at best, while others build upon more reliable quasi-experimental approaches. Consequently, the studies' results vary significantly (Debrezion et al. 2007; Mohammad et al. 2013). While most studies find a positive relationship between investment and housing prices, some observe adverse or ambiguous effects.

Examples of recent studies using a quasi-experimental research design include Zhou et al. (2019), who find a 3.8 percent price premium when a new metro line opened in Shanghai, and Fesselmeyer and Liu (2018), who study a metro expansion in Singapore and observe a 1.8 percent price increase in apartments within 500 meters of the existing stations.

However, while Ke and Gkritza (2019) find an announcement of a new LRT investment to affect property values positively in Charlotte-Mecklenburg, North Carolina, they find a negative effect during the operations phase. Camins-Esakov and Vandegrift (2018) analyze repeat sales in Bayonne, New Jersey, and detect no significant impact on dwelling prices after the construction of an LRT extension. Analogous to Zhou et al. (2019) and Ke and Gkritza (2019), this paper investigates the anticipation effect of a new LRT line with a quasi-experimental research design.

[1] *Several attributes affect a property's value, but these attributes cannot be purchased separately. Nonetheless, the values can be determined using a hedonic price model. (Rosen 1974; Kain and Quigley 1975.)*

Some modern studies use a state-of-the-art approach, combining a quasi-experimental design with cell phone ping data. Gupta et al. (2020) utilize this technique to assess a metro extension's anticipation effect in New York City. They find a 10 percent price increase in property values, which they credit to reduced commuter times and higher rents.

Research regarding the Finnish housing market is the most relevant to this paper as the housing markets react differently to external shocks, partly due to differences in market structures and the share of wealth used in housing (European Commission 2017). Unfortunately, there are very few studies concerning the impact of public transit investments in Finland employing a quasi-experimental approach, and even those examine only the impact of the metro.

Harjunen (2018) studies how building the West Metro in Espoo and Helsinki impacts dwelling prices. Using a quasi-experimental strategy, he finds a positive effect within 800 meters of the stations, even five years before the operations start. He compares price trends near the West Metro's stations to those near existing railway stations. The dwelling prices within the West Metro's buffer zones experience a 4 percent price increase. In addition, Laakso (1992) finds a positive impact from the Helsinki metro using a hedonic model. The identification strategy used in this paper is similar to the one used by Harjunen (2018), albeit here I examine the impact of Jokeri Light Rail instead of the West Metro.

2. Institutional setting

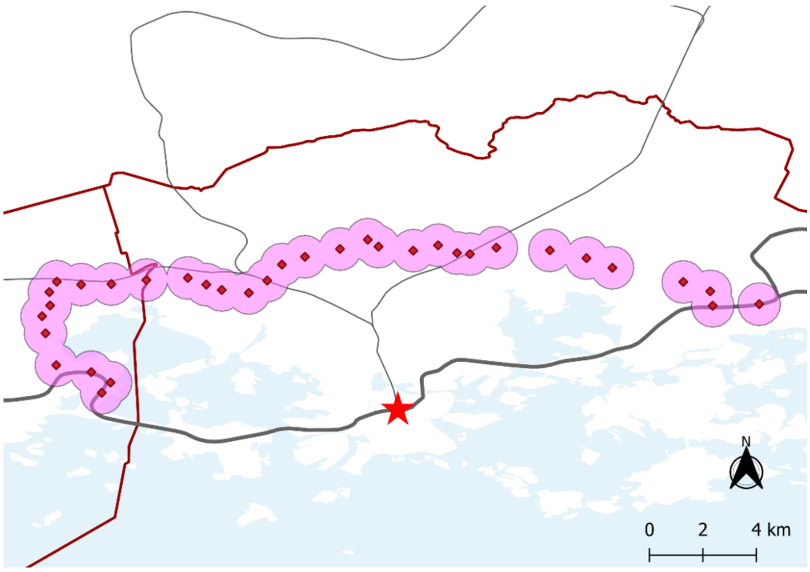
Helsinki is by far the most populous city in Finland: in 2019, it had 648,000 inhabitants, whereas the surrounding Helsinki region had over 1.49 million. These figures are expected to reach 820,000 and 1.92, respectively, by 2050. (Vuori and Kaasila 2019.) Accordingly, housing prices have increased more rapidly in Helsinki than in the rest of the country, and the trend is predicted to continue in the years to come (Laakso and Loikkanen 2016). In the Helsinki region, land values rise exponentially when travel time to the CBD (central business district) is reduced (Laakso and Loikkanen 2013).

Finland's urban rail network consists of tramways in southern Helsinki, a single metro line running from southern Espoo to eastern Helsinki, and commuter trains in the Helsinki region. Moreover, a tramway line is under construction in Tampere. Crosstown public transport in the Helsinki region is limited as only bus lines are in operation. The busiest bus line of the whole region is a crosstown connection named trunk route 550 (Jokeri Light Rail: Information). The trunk route 550 started its operations in August 2003.

The daily ridership of route 550 is already at its upper limit, with 40,000 daily passengers, and this figure is expected to double in the next ten years. Therefore the cities of Espoo and Helsinki have decided to construct a new crosstown LRT line called Jokeri Light Rail to ease the traffic congestion in this growing region. The travel time of Jokeri Light Rail will be identical to the current bus line, but the LRT will be more immune to congestion, and its capacity will be two to three times higher. (Jokeri Light Rail: Information.) Similar rail investments are being planned in other cities, such as in Vantaa (City of Vantaa 2019).

The route of Jokeri Light Rail will run from Keilaniemi in the west to Itäkeskus in the east. The Jokeri Light Rail will have 34 new stops along its 25-kilometer route, of which nine are in Espoo, and sixteen are in Helsinki. A substantial amount of residential housing is planned to be built along its route on top of workplaces for at least 20,000 people. (Jokeri Light Rail: Information.) Figure 1 shows the route of Jokeri Light Rail and 800-meter buffer zones around the stops.

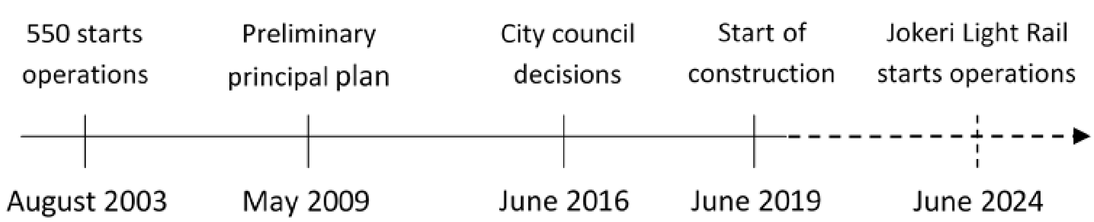
Figure 1: *The route of Jokeri Light Rail and the 800-meter buffer zones*



Note. The red star indicates the CBD (Helsinki Central railway station). The metro line is denoted by the dark grey line, railway lines with light grey, and municipality borders with dark red (General Guide Map of the City of Helsinki). The reference system is WGS 84.

The preliminary principal plan of Jokeri Light Rail was completed in May 2009 (Jokeri Light Rail 2009), while the project plan was completed in January 2016 (Jokeri Light Rail 2016a). The city councils of Espoo and Helsinki approved the project plan in June 2016 (Jokeri Light Rail 2016b). Construction began in June 2019, and the planned beginning of operations is June 2024 (Jokeri Light Rail: Construction). The project's final cost estimate was 386 million euros in 2019 (Jokeri Light Rail 2019). I present the timeline of the Jokeri Light Rail in Figure 2.

Figure 2: *The timeline of Jokeri Light Rail*



3. Methodology

3.1 Timing of the capitalization

In this paper, I analyze how Jokeri Light Rail capitalizes in property prices. Hence, I assess possible capitalization times by considering two different factors. Firstly, as in many previous studies about the housing market capitalization of an investment, I examine the dates of important political decisions. As discussed in the previous section, the most vital steps for Jokeri Light Rail took place in 2009, 2016, and 2019. However, information about the councils' decisions, for instance, might not immediately reach the public's awareness.

Secondly, I consider the information the public received about the investment to be of significant importance because capitalization is observed in the public's housing transactions. I view Google search figures between 2006 and 2019 as a proxy indicator of the public's information. These figures are presented in Table 1. The second column shows how the relative number of searches per year progresses, where a higher number means more searches. The third column displays

Table 1: *Public's information about Jokeri Light Rail between 2006 and 2019*

Year	Google search index	Number of weeks searched	Increase in number of weeks searched
2006	359	4	–
2007	476	5	25%
2008	743	11	120%
2009	748	13	18%
2010	809	10	-23%
2011	1,354	18	80%
2012	1,462	25	39%
2013	1,885	34	36%
2014	2,089	33	-3%
2015	2,433	48	45%
2016	1,329	69	44%
2017	2,695	75	9%
2018	3,495	77	3%
2019	4,135	90	17%

Note. Search words (in Finnish) "Raide-Jokeri" (term) and "Raide-Jokeri" (subject) inspected in Google Trends (2020). Inspected search weeks per year total 104. Google search index displays how the relative number of searches varies over time, higher meaning that more searches have been conducted. The number of weeks searched shows how many weeks (out of 104) the search words have been used on Google search.

the number of weeks in a year on which the search terms have been used (out of 104), and the fourth column shows the relative increase in the number of weeks searched.

Although there is an upward trend during the whole interval from 2006 to 2019, there are noticeable hikes in 2008, 2011, 2015–2016, and 2019. Subsequently, I regard 2012, 2016, and 2019 as the most likely moments of capitalization. However, to mitigate the risk of using an incorrect year for capitalization, I also estimate yearly effects.

3.2 Research design

I study the effects of the rail investment on the housing market using the investment decision as a quasi-experiment (see Billings 2011; Dubé et al. 2018). The aim is to measure the counterfactual price growth, i.e., how the property prices within the catchment areas of the upcoming Jokeri Light Rail stops would have developed had the investment not been made, and then compare that to the factual price development in the catchment areas.

I estimate the magnitude, timing, and geographical extent of the effect. Furthermore, I use this estimate to assess the value of the total windfall for homeowners. As the start of operations is still yet to come, I estimate the investment's anticipation effect (McDonald and Osuji 1995; Gibbons and Machin 2005).

Hedonic price models are widely used when assessing a rail investment's impact on the housing market. However, as hedonic models do not specify the effect's timing and often suffer from an omitted variable bias, the results cannot be interpreted as causal. (Parmeter and Pope 2013; Mohammad et al. 2017.)

In this paper, my identification strategy is a difference-in-differences (DID). It is one of the most used methods when evaluating the efficacy of policies (Ashenfelter and Card 1984; Billings 2011). In addition, I combine DID with a hedonic price model to achieve more reliable and precise results: the additional control variables should reduce the residual variance (see Papon et al. 2015).

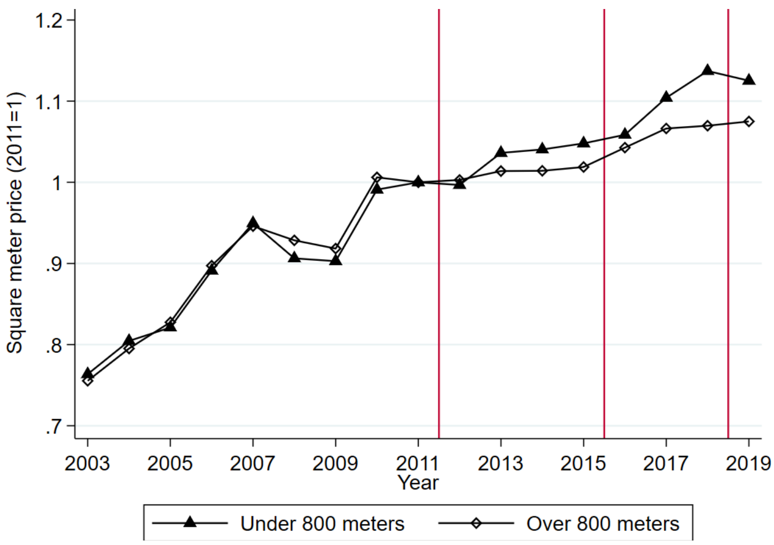
Using a DID model, three assumptions must be satisfied for the results to be interpreted causally: no spillovers between the treatment and the control group, parallel trends in the absence of treatment in both groups, and no coinciding policy reforms or other events that would affect the groups differently.

My primary treatment group comprises observations within 800 meters of the Jokeri Light Rail stops. In contrast, my primary control group includes observations outside these buffer zones but within certain postal code areas around the investment (see Table A1 in Appendix A). Moreover, I assume there are no spillovers between the treatment group and the control group because the dwellings are static, and I consider the impact of the treatment to lessen evenly as the distance to the stops increases.

I consider 800 meters to be the limit of an easily walkable distance (see Olszewski and Wibowo 2005; Harjunen 2018). Later, I proceed to show results for 200-meter bands, which support the usage of the 800-meter limit. Apart from the treatment-control group setting, I also test models using a continuous variable for distances measured in a hundred meters.

To test the assumption of parallel trends, I apply a graphical approach. I inspect the average yearly square meter price trends in Figure 3, both for the observations within the 800-meter buffer zones and the observations outside of these zones. The vertical lines represent possible capitalization times; see the discussion in Section 3.1. The year 2011 is indexed as 1.

Figure 3: *Square meter price trends for the treatment and control group*



Note. The vertical lines represent possible moments of capitalization (see Section 3.1). The year 2011 is indexed as one.

The square meter price trends are similar in both groups until 2012. There is a slight jump in the prices closer to the stops in 2013, which might be due to the increased awareness about Jokeri Light Rail. Nonetheless, according to Figure 3, the assumption of parallel trends holds well until 2012, after which the prices increase somewhat more within the 800-meter buffer zones. Another point of divergence in prices is noticeable after 2016.

Therefore, I consider the assumptions of parallel trends and no spillovers to be satisfied. Moreover, as I estimate the aggregate impact on housing prices along a circular route, there are most likely no coinciding events that would affect only the dwellings near the investment.

This study's econometric identification strategy (DID) estimates how the closeness to Jokeri Light Rail stops affects the dwelling prices. However, that does not necessarily reflect the accessibility changes. On the one hand, the estimates cannot explicitly take into account possible changes in housing supply[2], which can lead to downward skewed estimates; on the other hand, the investment might also shift the demand from the control to the treatment area, lowering the demand in the former, thus causing the estimates to be skewed upward. Finally, also the possible externalities are omitted from the explicit analysis.

3.3 Econometric models

I estimate four different econometric models, of which the initial ones are simple. Then, I test the robustness of these models by adding more control variables. The models differ from one another by interaction terms generated by combining time and distance variables. In addition to using two different types of measures for the distance to the stops, i.e., under and over 800 meters and a continuous variable, I use two variables for time: a before-and-after set-up, where the after-period begins in 2016 (see Section 3.1), and a year-wise set-up.

[2] *This may be negligible because an increase in housing supply presumably affects both the control and the treatment group similarly if the distance to the new housing is constant. Moreover, dwellings near the investments can be considered relatively reliable substitutes for each other.*

Model (1) is a clean DID model containing a before-and-after setting and the two distance groups; model (2) is a modification of the former where a continuous variable replaces the distance groups to estimate the intensity (i.e., gradient) of the treatment. In model (3), I measure time year-wise and distance in two groups. Again, I substitute the continuous variable for the distance groups in model (4). In model (1), represented by Equation (1), the price of an apartment i in year *t* is expressed as follows:

$$ln(SquareMeterPrice_{it}) = \alpha + \beta * JokeriLightRail_i + \gamma * After_t + \tau * (JokeriLightRail_i * After_t) + \delta * X_{it} + \varepsilon_{it} \quad (1)$$

, where the interaction term between the distance to Jokeri Light Rail and the time variable indicates the average treatment effect of the investment. X_{it} is a vector of a set of apartment characteristics used as controls, and ε_{it} are the error terms. Equations for models (3) and (4) are similar to (1), but I estimate yearly effects instead, *ergo* $After_t$ is replaced with $Year_t$.

As primary control variables, I include several housing and neighborhood characteristics (refer to Table 2). As an alternative control, in models with a discrete distance variable, I add postal code (of which there are 62) fixed effects to refine the models by capturing variation in small city district characteristics. When measuring the distance to the stops continuously, I use LRT stop (34) fixed effects, where I assign each observation to its closest Jokeri Light Rail stop. Likewise, standard errors are clustered by the postal code or LRT stop level, respectively. However, because the number of stops is relatively low, I also test the statistical significance based on the cluster generalization of the wild bootstrap in models 2 and 4 (see Cameron et al. 2008).

4. Data

The housing transaction data used in this study is obtained from the database of the Central Federation of Finnish Real Estate Agencies. These data include around 75% of the housing transactions in Finland conducted in the secondary markets (Central Federation of Finnish Real Estate Agencies: Housing markets). The data is very rich, including several micro-level variables about the transactions.

In this study, I include housing transactions made in 2003–2019 in 62 postal code areas along the route of Jokeri Light Rail (see Table A1 in Appendix A). I have geocoded[3] the transactions and calculated Euclidean distances between apartments and the nearest Jokeri Light Rail stop. This study only includes old apartment buildings, low-rise apartment buildings, maisonettes, rowhouses, and duplex houses. I exclude new[4] apartments and single-family houses because those may differ substantially from the other types of buildings.

Moreover, I have omitted observations that contained clear errors, e.g., if the geocoding was not successful on the street level. Observations with square meter prices outside three standard deviations from the mean are omitted because I consider them outliers (e.g., their characteristics might deviate considerably from the rest of the observations) or possible errors. The total number of observations shrinks slightly, from 78,999 to 77,378.

In Table 2, I present descriptive statistics over the housing transactions in Espoo and Helsinki in 2003–2019. Observations within the 800-meter buffer zones are used as a primary treatment group, whereas observations farther than 800 meters away from the Jokeri Light Rail stops are treated as a primary control group. Section 3.2 presents the discussion on the choice of the treatment and the control group.

[3] *Forward geocoding conducted in Stata using command opencagegeo and open data sources.*

[4] *Here, a new apartment refers to a newly constructed apartment, whereas old apartments are sold in secondary markets.*

The second column includes all the observations, while the third and fourth columns include only observations within the catchment areas. The fifth and sixth columns contain observations outside the 800-meter buffer zones. I have deflated all prices to the 2019 level using the CPI (Consumer Price Index 2020).

Table 2: *Descriptive statistics about housing transactions in Espoo and Helsinki in 2003–2019*

	All observations	**Under 800 meters**		**Over 800 meters**	
		2003–2015	**2016–2019**	**2003–2015**	**2016–2019**
N	77,378	16,419	5,192	41,983	13,784
Selling price	218,510 (111,494)	193,548 (92,452)	233,832 (98,528)	216,011 (113,972)	250,084 (120,567)
Price per square meter	3,578 (1,016)	3,420 (806)	4,041 (1,057)	3,466 (974)	3,931 (1,199)
Floor area (m²)	64.9 (29.6)	60.3 (28.6)	64.3 (28.5)	65.6 (30.1)	68.9 (29.2)
Apartment age (years)	38.9 (17.8)	39.5 (16.0)	43.5 (18.9)	37.4 (17.5)	41.2 (19.5)
Maintenance charge (€/m²)	3.76 (1.31)	3.59 (1.14)	4.20 (1.13)	3.65 (1.37)	4.15 (1.30)
Distance to nearest stop (m)	1,666 (1,043)	449 (201)	444 (198)	2,138 (836)	2,137 (827)
Travel time to CBD (min)	33.6 (7.8)	33.1 (5.0)	33.5 (5.1)	33.5 (8.7)	34.2 (8.5)
Freehold site (%)	62	62	58	63	60
Apartment building (%)	81	89	87	79	77
Condition (%)					
- *New or excellent*	1	0	3	0	3
- *Good*	53	49	58	52	59
- *Satisfactory*	37	40	35	36	34
- *Poor*	4	5	3	4	3
- *Unknown*	6	7	1	8	2

Note. Distances are measured to the nearest Jokeri Light Rail stop in a direct straight line. All prices are in euros and deflated to the 2019 level using the consumer price index. Travel time measured in minutes to Helsinki central railway station (CBD) using public transport during the rush hour, taking into account changes and waiting times. Figures are averages, standard errors are reported in parentheses. Travel times are from Tenkanen et al. (2018), other variables taken from the Central Federation of Finnish Real Estate Agencies.

Table 2 shows the number of observations in the control group to be double that of the number of observations in the treatment group. Nonetheless, most variable means in both distance groups are relatively similar, though there is some variation. However, the average square meter prices show more of an increase inside the catchment areas than outside them. The yearly number of observations is comparatively stable in both groups throughout the period, as illustrated in Figure A1 in Appendix A.

In my analysis, I only include those housing characteristics presented in Table 2, even though the data set contains several more. This choice stems from two reasons: firstly, most of the observations have no record of all the variables, and secondly, my models are relatively robust to the rest of the excluded characteristics. I consider travel time (to the CBD using public transport) as a proxy variable for local retail services and regional hubs. It is a static variable measured before the construction of Jokeri Light Rail begun.

5. Results

In this section, I present results from models 1–4. Moreover, I test the robustness of the results. For models 1 and 2, where the after-period begins from 2016, I show the estimates in table form. In models 1 and 2, I omit the observations between 2013 and 2015 to minimize the risk of contaminating the pre-treatment period as the price trends are dissimilar during those years (see Figure 3). When estimating yearly effects in models 3 and 4, I adopt a graphical presentation.

Extensive results for all models can be found in Appendix B. For my analysis, the most intriguing estimates are the interaction terms between the closeness to the Jokeri Light Rail stops, measured group-wise or continuously, and time measured in two groups or yearly. I treat these interactions as measures of how the apartments' square meter prices are affected by the investment.

5.1 Models 1 and 2

The results for models 1a–d are shown in Table 3. The second column (1a) includes only a fully saturated interaction term; the third column (1b) also includes postal code fixed effects, whereas, in the fourth (1c), I control (1a) by adding apartment characteristics as control variables. The fifth column (1d) combines all the former models. The reference group is observations farther than 800 meters from Jokeri Light Rail stops before 2016. Observations between 2013 and 2015 are omitted.

As Figure 3 already suggests, Jokeri Light Rail is estimated to positively impact dwelling prices within the catchment areas. According to model 1d, prices are 4.6 percent higher after 2016 within the 800-meter buffer zones than farther away. The estimate is robust for different housing and neighborhood characteristics used as additional controls. Moreover, the statistical significance of the estimate is at the 5 percent level.

I present the results for models 2a–d in Table 4. Here I measure the distance to the stops with a continuous variable. The second column (2a) includes only the interaction term, the third column (2b) also the LRT (i.e., Jokeri Light Rail) stop fixed effects. In the fourth column (2c), I control the initial model with apartment characteristics. All the former is com-

Table 3: *Model 1. Housing market effect in two distance groups before and after 2016*

Response variable: ln(price per square meter)	**Reference group: over 800 meters and before 2016**			
	(1a)	**(1b)**	**(1c)**	**(1d)**
Under 800 m*After	0.0486* (0.0252)	0.0502* (0.0265)	0.0342* (0.0179)	0.0456** (0.0205)
Under 800 m	-0.00977 (0.0516)	0.00277 (0.0164)	0.00689 (0.0263)	-0.0181 (0.0113)
After	0.139*** (0.0158)	0.144*** (0.0173)	0.117*** (0.0144)	0.139*** (0.0128)
Control variables	No	No	Yes	Yes
Postal code FE	No	Yes	No	Yes
N	63,745	63,745	63,745	63,745
R^2	0.061	0.061	0.476	0.316

*Note. Distances measured to the nearest Jokeri Light Rail stop. The sample is constrained to sales in postal code areas reported in Table A1 in Appendix A. Before-period is 2003–2012, after-period is 2016–2019. Observations from 2013–2015 are not used. Control variables correspond to those reported in Table 2, also taking into account floor area squared. The estimated coefficients statistical significance is marked with * (10%), ** (5%) or *** (1%). Standard errors clustered at the postal code level (62) are reported in parentheses.*

Table 4: *Model 2. Housing market effect with continuous distance before and after 2016*

Response variable: ln(price per square meter)	**Reference group: before 2016**			
	(2a)	**(2b)**	**(2c)**	**(2d)**
Distance*After	-0.00309*** (0.000689)	-0.00269** (0.00107)	-0.00157** (0.000727)	-0.00207** (0.000888)
Distance	0.00172 (0.00299)	0.00200 (0.00173)	-0.00150 (0.00117)	-0.000368 (0.00108)
After	0.204*** (0.0172)	0.205*** (0.0193)	0.152*** (0.0149)	0.179*** (0.0150)
Control variables	No	No	Yes	Yes
LRT stop FE	No	Yes	No	Yes
N	63,745	63,745	63,745	63,745
R^2	0.063	0.063	0.480	0.409

*Note. Distances measured to the nearest Jokeri Light Rail stop. The sample is constrained to sales in postal code areas reported in Table A1 in Appendix A. Before-period is 2003–2012, after- period is 2016–2019. Observations from 2013–2015 are not used. Control variables correspond to those reported in Table 2, also taking into account floor area squared. The estimated coefficients statistical significance is marked with * (10%), ** (5%) or *** (1%). Standard errors clustered at the LRT stop level (34) are reported in parentheses.*

bined in the fifth column (2d). The reference group is observations before 2016. Observations between 2013 and 2015 are omitted.

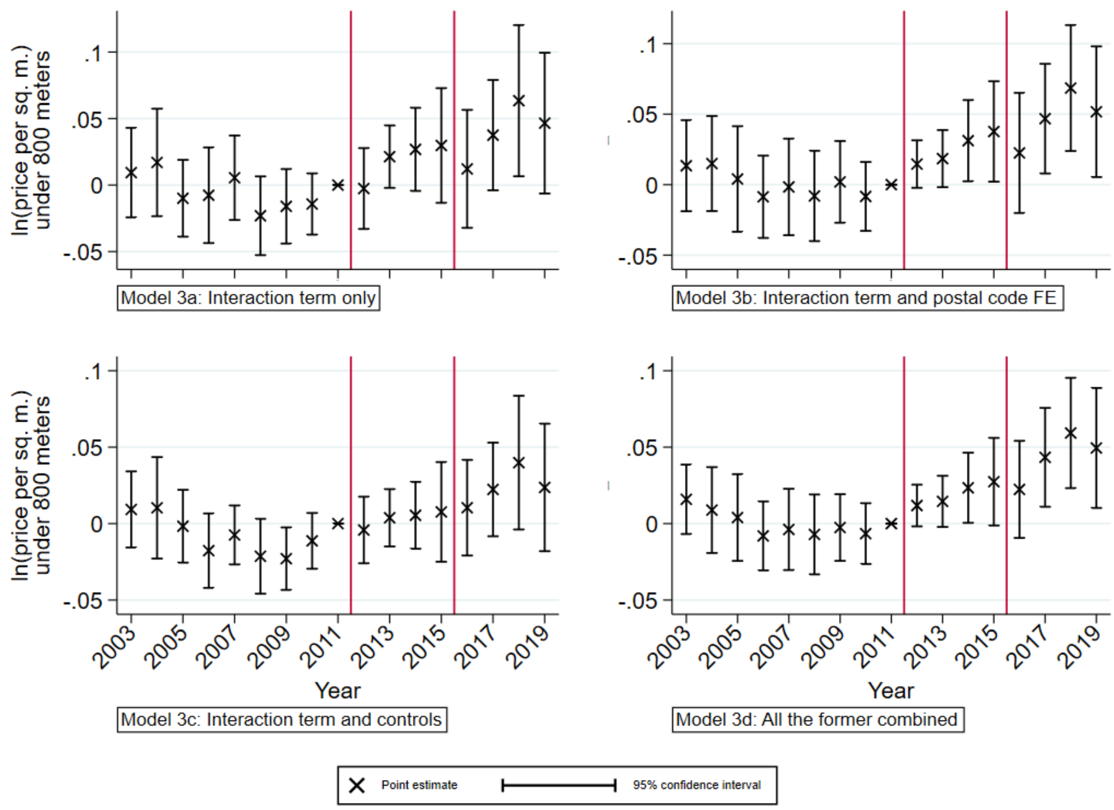
When the distance to the Jokeri Light Rail stops is measured with a continuous variable, the estimate for the investment's impact has a negative sign in every model, meaning that the closer the dwelling is to the stops, the higher the prices are after 2016. Model 2d shows that each hundred-meter increase in distance to the stops decreases prices by 0.2 percent after 2016. This estimate is significant at the 5 percent level. The wild bootstrap test produces comparable results (see Table B2 in Appendix B).

The control variables in models 1 and 2 impact the prices similarly. Travel time, the condition (lower value equals better condition), and floor area of the apartment correlate negatively with the prices; a freehold site and the scarcity of neighbors, on the other hand, correlate positively. The impact of age is non-linear. The maintenance charge does not have a significant effect on the prices.

5.2 Models 3 and 4

Figure 4 provides the interaction terms and their 95 percent confidence intervals for models 3a–d, where I estimate yearly effects with two distance groups. Model 3a includes only the fully saturated interaction term. Model 3b contains the interaction term and postal code fixed effects, whereas model 3c includes housing and neighborhood characteristics but no fixed effects. Finally, model 3d combines all the former. The reference group is observations outside the 800-meter buffer zones and in 2011.

Figure 4: *Models 3a–d. The yearly impact of Jokeri Light Rail in two distance groups*



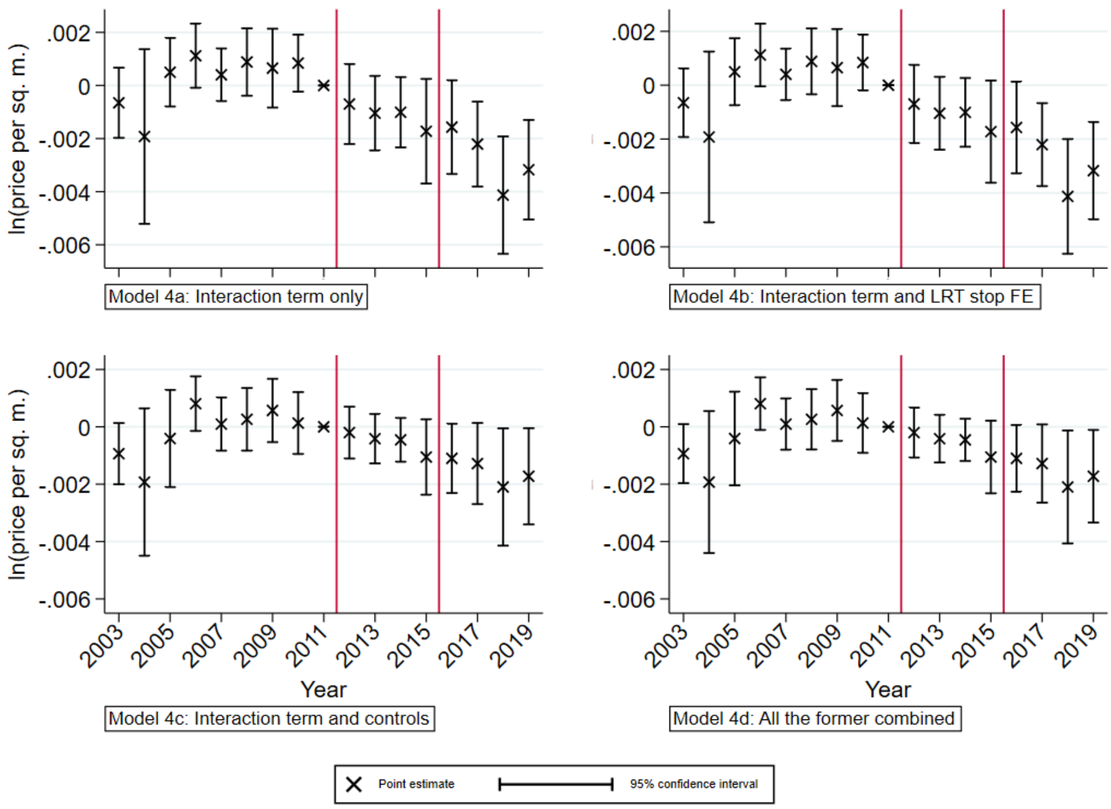
Note. Detailed results are presented in Table B3 in Appendix B. The reference group is transactions over 800 meters away from Jokeri Light Rail stops and the year 2011. The vertical lines represent possible moments of capitalization (see Section 3.1).

A surge in the prices within 800 meters of the stops in 2017–2019 can be distinguished in Figure 4. A more precise analysis of the results shows that prices are 4.5–6.2 percent higher in 2017–2019 within the buffer zones than outside. The point estimates are relatively robust for additional controls as the estimates are statistically significant.

In models 4a–d, I estimate the yearly effects using a continuous distance variable. I present the point estimates for the interaction terms and their 95 percent confidence intervals in Figure 5. Model 4a contains the interaction term solely. In model 4b, I control the former using LRT stop fixed effects. Model 4c combines the interaction term and additional apartment characteristics. Model 4d is a combination of all the former. Here the reference group is observations in 2011.

It can be observed in Figure 5 that moving farther away from the stops lowers the prices from2015 onward. The most potent effect is observable in 2018–2019, when the price reduction is 0.19–0.23 percent *per* hundred meters. These estimates are statistically significant at the 5 percent level. Moreover, the wild bootstrap tests bolster these results.

Figure 5: *Models 4a–d. The yearly impact of Jokeri Light Rail with continuous distance*



Note. Detailed results are presented in Table B4 in Appendix B. The reference group is the year 2011. The vertical lines represent possible moments of capitalization (see Section 3.1).

5.3 Alternative specifications

I conduct balancing tests to account for possible sorting, e.g., more affluent households might buy higher quality homes close to the investment, which affects the sales composition. The tests (refer to Table 2) show that, on average, the dwellings in both the treatment and the control group are statistically similar in terms of observable characteristics. Moreover, I further assess the anticipation effect using alternative distance bands. Of all the models estimated, model 3d[5] has the most statistical power.

I test a modification of model 3d, where I adopt 200-meter bands instead of the original 800-meter ones (see Figure A2 in Appendix A). However, as the yearly number of observations in the 200-meter bands is relatively low, I divide the observations into eight year groups. The reference group in model 5 is observations over 3,000 meters away from the stops and in 2010–2011. Extensive results are shown in Table B5 in Appendix B.

Figure A2 illustrates that the anticipation effect is most abundant in 2018–2019. The positive impact of the investment is observable within 800 meters of the stops. This observation supports the usage of the 800-meter buffer zones as in models 1 and 3. The most substantial impact is felt 0–200 meters from the stops, where the prices are 8.2 percent higher in 2018–2019 than over 3,000 meters away. The impact is between 5 to 7 percent within 200–800 meters of the stops.

[5] *Model 3d estimates yearly effects in two distance groups and includes apartment characteristics and postal code fixed effects, too.*

6. Discussion

When estimating the anticipation effect using the 800-meter catchment areas and a before-and-after setting in model 1, I observe that apartment prices are 4.6 percent higher after 2016 within the 800-meter buffer zones than outside them. In model 3, where I estimate yearly effects, the impact is even more substantial at 4.5–6.2 percent. The latter estimate is notably robust and statistically significant. Higher standard errors closer to the start of operations may be due to dwellings or areas becoming more dissimilar from one another.

However, there may be an alternative explanation for the results. The general house price development cannot be held responsible because I have deflated the prices to the 2019 level using the CPI, and I also incorporate year fixed effects in some of my models. The estimated price increase is probably not due to housing price development near extensive traffic connections either: the control group of my study includes observations where the connections are even more extensive, and thus the urban development is more intense.

The shifts in supply and demand from outside the buffer zones to the inside might have a considerable effect on the estimates. However, Figure A1 in Appendix A depicts no evident supply or demand shifts: the yearly transaction share is relatively stable inside and outside the 800-meter buffer zones. Moreover, although I have omitted new buildings from my sample, it may well be that Jokeri Light Rail will increase the supply of new housing near the investment, affecting the demand for old housing. Nevertheless, I consider this unlikely to have happened so soon.

I present the yearly housing transactions of new buildings in Figure A3 in Appendix A, displaying only a slight increase in sales. Furthermore, if the sales of new apartments increased substantially, my models would underestimate the actual anticipation effect rather than overestimate it.

Dubé et al. (2018) find similar results when studying a rail investment's anticipation effect within 600 meters of stations. Furthermore, Harjunen (2018) observes a 4-percent price premium within 800 meters of metro stations. However, Ransom (2018) finds no significant effects from an LRT investment, though the results vary spatially. All these three studies utilize a difference-in-differences strategy, which makes them good benchmarks for my paper.

When measuring the distance to the stops using a continuous variable, in models 2 and 4, I estimate that housing prices fall by 0.19–0.23 percent for each 100-meter increase in distance to the stops in 2018–2019. These estimates are relatively robust for alternative controls and statistically significant. Even though the number of clusters in these models is only 34, the wild bootstrap tests imply that the significance levels reported are robust.

Dai et al. (2016) find that each hundred meters away from metro stations substantially reduces property prices. However, several studies imply spatial variance in the effect when measuring the distance to the stations continuously. Camins-Esakov and Vandegrift (2018) do not detect any significant effect when assessing the impact of an LRT. In addition, Papon et al. (2015) find that a rail investment may also have a negative effect near the stations.

Moreover, there may be several factors affecting my estimates. Firstly, the shape of the LRT line is an arc around the CBD, which means the accessibility improvements, especially in the regional centers, are not maximized (see Mulley and Tsai 2016). Furthermore, the same route has been operated by a BRT line, and as an LRT, Jokeri Light Rail will not reach as high a service level as a heavy rail line would.

Secondly, the public's information about Jokeri Light Rail is imperfect. The information asymmetry regarding the operations' expected beginning, as the planned start is still many years away, and its impact on other public transportation is unclear. Thus, it might be difficult for the public to anticipate the investment accurately.

Thirdly, the research design may pose problems for validity. It is unlikely that the effects of Jokeri Light Rail end 800 meters from the stops. However, as model 5 shows, 800 meters is the best choice for the buffer zones. The bias is most likely positive as the apartments outside the primary catchment areas used in this study also benefit from the investment.

In this study, I measure the Euclidean distances to the stops, which do not admit any geographical obstacles such as highways or waterways. While I estimate the average treatment effect of Jokeri Light Rail, I surmise that the price changes will be the most drastic in areas where accessibility improves the most (see Diaz and Mclean 1999).

However, observations located over 800 meters from the stops, the primary control group of this study, might not provide the optimal counterfactual. The chosen control group can significantly impact the observed housing market effect (Pilgram and West 2018). Finally, my data cover only six months of transactions after the construction of Jokeri Light Rail started: the actual anticipation effect may be observed at the beginning of the 2020s.

A clear majority of the observations belong to the control group (see Table 2). While I assume the demand to shift from the control group to the treatment group, the fall in demand in the former might not be substantial as the price level has increased in both groups. Ultimately, long-term price changes are likely to exceed the short-term ones in both groups so that the eventual price effect will be greater than the anticipation effect estimated in this study.

6.1 A rough estimate of the total value of capitalization

The cost estimate for Jokeri Light Rail's construction is 386 million euros (Jokeri Light Rail 2019). Here, I provide a feasible estimate for the return on the investment. I assume that the housing transaction data is a representative sample of the total housing stock.

The average impact of Jokeri Light Rail on the square meter prices within 800 meters of the stops in 2019 is 4.95 percent (see Table B3 in Appendix B)[6]. The average square meter price of those apartments is 3,569 euros (Central Federation of Finnish Real Estate Agencies). The existing housing stock within the buffer zones consists of 2.19 million square meters of floor area in 2018 (Registry data: SeutuCD'18).

Therefore, the estimated total windfall generated is between 80 and 723 million euros; when using the point estimate, the estimated value is 396 million. Although it is unclear whether the appreciation of housing stock is distributed optimally, some gain most likely contributes to increased real-estate tax revenue.

My estimate should be taken with some trepidation because I might have overestimated the impact of Jokeri Light Rail due to demand shifts. However, the estimate does not consider the capitalization in commercial properties[7]. For example, there were nearly 850,000 square meters of floor area in office buildings within the catchment areas in 2018 (Registry data: SeutuCD'18). Moreover, I neither explicitly consider possible externalities in this assessment: Jokeri Light Rail may impact, e.g., the number of jobs and traffic congestion.

7. Conclusions

In this paper, I analyze how the political decision of constructing Jokeri Light Rail in Espoo and Helsinki affects the local housing markets between 2003 and 2019. I study whether the housing markets anticipate improved accessibility and forthcoming urban development. To achieve this, I utilize micro-level housing transaction data combined with a difference-in-differences identification strategy. As an added control, I include housing and neighborhood characteristics in my models.

My results imply that the anticipation effect is detectable in the housing markets of Espoo and Helsinki. The positive impact is statistically significant in 2016–2019 when the public had enough information about the investment and its ascertainment, i.e., five to eight years before the operation's planned beginning. Most studies regarding rail investment's

[6] *The 95 percent confidence interval's lower bound is 1.02 percent, while the upper bound is 8.87 percent.*

[7] *A majority of studies have found that the impact of rail investment on commercial properties is positive (Debrezion et al. 2007; Gupta et al. 2020).*

effect on the housing market reach similar conclusions, which is on par with urban economics theories on land value capitalization. However, the results seem to be conditional on the research framework.

I find that the intensity of Jokeri Light Rail's anticipation effect depends on the distance to the stops: the effect is observed most distinctly within 800 meters. The prices of apartments within the 800-meter buffer zones around the stops are, on average, 5 percent higher compared to apartments farther away in 2019. Furthermore, I demonstrate that the average price decrease is around 0.2 percent for each hundred-meter increase in distance to the stops.

My rough estimate for the total windfall for homeowners is 396 million euros, which exceeds the cost estimate for the investment. This hike in housing values reflects households' increased willingness to pay to live near the investment. Moreover, the benefits will most likely increase in the future due to heightened urban development. Besides, the estimate for the total windfall ignores the impact on commercial properties.

My estimates must be regarded with certain reservations. In addition to possibly unrealistic assumptions, there is some level of uncertainty related to the estimates. Standard errors increase the closer the start of operations becomes, perhaps due to observations becoming more dissimilar to each other. While possible externalities are omitted from the explicit analysis, the estimates seem robust when considering different distance bands. Moreover, I find that supply changes, which might distort the results, are, at the most, minor.

This paper can help decision-makers assess new rail investments, especially in Finland, since the housing price increase means demand for dwellings is higher near rail investments. However, I cannot definitely state whether the decision to build Jokeri Light Rail is efficient. Nevertheless, under the assumptions made, my study indicates that property values increase substantially due to the investment and that those benefits outweigh the investment costs.

It would be crucial to increase the number of similar studies conducted, especially in the Finnish context. For instance, the operations phase of West Metro's first stage or the construction phase of West Metro's second stage could be used to measure the efficacy of a rail investment further. Finally, combining a quasi-experimental design with cell phone ping or individual-level data could provide the ultimate tool for assessing an investment's suitability.

References

Alonso, W. (1964), *Location and land use: Toward a general theory of land rent*, Harvard University Press, Cambridge.

Ashenfelter, O. C. and Card, D. (1984), "Using the longitudinal structure of earnings to estimate the effect of training programs", *Review of Economics and Statistics*, 67(4), pp. 648–660.

Billings, S. B. (2011), "Estimating the value of a new transit option", *Regional Science and Urban Economics*, 41(6), pp. 525–536.

Cameron, A. C., Gelbach, J. B. and Miller, D. L. (2008), "Bootstrap-based improvements for inference with clustered errors", *The Review of Economics and Statistics*, 90(3), pp. 414–427.

Camins-Esakov, J. and Vandegrift, D. (2018), "Impact of a light rail extension on residential property values", *Research in Transportation Economics*, 67, pp. 11–18.

Central Federation of Finnish Real Estate Agencies: Housing markets. Retrieved from https://kvkl.fi/ajankohtaista/asuntomarkkinat/ on 21.12.2020.

Central Federation of Finnish Real Estate Agencies, KVKL HSP -hintaseurantapalvelu. Retrieved from https://www.hintaseurantapalvelu.fi/ on 3.1.2020.

City of Vantaa (2019), Vantaan ratikan yleissuunnitelma. Retrieved from https://www.vantaa.fi/instancedata/prime_product_julkaisu/vantaa/embeds/vantaaw wwstruc- ture/147683_Vantaan_ratikan_yleissuunnitelma_raportti.pdf on 3.12.2019.

Consumer Price Index (2020), ISSN=1796-3524. Statistics Finland, Helsinki. Retrieved from http://www.stat.fi/til/khi/index.html on 4.2.2020.

Dai, X., Bai, X. and Xu, M. (2016), "The influence of Beijing rail transfer stations on surrounding housing prices", *Habitat International*, 55, pp. 79–88.

Debrezion, G., Pels, E. and Rietveld, P. (2007), "The impact of railway stations on residential and commercial property value: a meta-analysis", *The Journal of Real Estate Finance and Economics*, 35(2), pp. 161–180.

Diaz, R. B. and Mclean, V. A. (1999), "Impacts of rail transit on property values", American Public Transit Association Rapid Transit Conference Proceedings, pp. 1–8.

Dubé, J., Legros, D. and Devaux, N. (2018), "From bus to tramway: Is there an economic impact of substituting a rapid mass transit system? An empirical investigation accounting for anticipation effect", *Transportation Research Part A: Policy and Practice*, 110, pp. 73–87.

European Commission (2017), Housing market developments. European semester thematic factsheet. Retrieved from https://ec.europa.eu/info/sites/info/files/file_import/europeansemester_thematic- factsheet_housing-market-developments_en.pdf on 17.3.2020.

Fesselmeyer, E. and Liu, H. (2018), "How much do users value a network expansion? Evidence from the public transit system in Singapore", *Regional Science and Urban Economics*, 71, pp. 46–61.

General Guide Map of the City of Helsinki. The maintainer of the dataset is Helsingin kaupunkiym- päristön toimiala / Kaupunkimittauspalvelut. The dataset has been downloaded from Helsinki Region Infoshare service on 10.2.2020 under the license Creative Commons Attribution 4.0. Retrieved from https://hri.fi/data/fi/dataset/helsingin-yleiskartta.

Gibbons, S. and Machin, S. (2005), "Valuing rail access using transport innovations", *Journal of urban Economics*, 57(1), pp. 148–169.

Google Trends (2020), Google. Retrieved from https://trends.google.com/trends/?geo=FI on 29.1.2020.

Gupta, A., Van Nieuwerburgh, S. and Kontokosta, C. (2020), "Take the Q Train: Value Capture of Public Infrastructure Projects", No. w26789, National Bureau of Economic Research.

Harjunen, O. (2018), "Metro investment and the housing market anticipation effect", Working Papers 2018:2, City of Helsinki.

James, O. (2009), "Evaluating the expectations disconfirmation and expectations anchoring approaches to citizen satisfaction with local public services", *Journal of public administration research and theory*, 19(1), pp. 107–123.

Jokeri Light Rail (2009), Alustava yleissuunnitelma. Retrieved from http://raidejokeri.info/wp- content/uploads/2015/05/Raide-Jokeri_raportti.pdf on 22.12.2019.

Jokeri Light Rail (2016a), Raide-Jokerin hankesuunnitelma valmis päätöksentekoon. Retrieved from https://raidejokeri.info/raide-jokerin-hankesuunnitelma-valmis-paatoksentekoon/ on 23.12.2019.

Jokeri Light Rail (2016b), Valtuustot näyttivät vihreää valoa Raide-Jokerille. Retrieved from https://raidejokeri.info/valtuustot-nayttivat-vihreaa-valoa-raide-jokerille/ on 23.12.2019.

Jokeri Light Rail (2019), Raide-Jokerin toteutus etenee kaupunkien päätöksentekoon. Retrieved from https://raidejokeri.info/raide-jokerin-toteutus-etenee-kaupunkien-paatoksentekoon/ on 3.12.2019.

Jokeri Light Rail: Construction. Retrieved from https://raidejokeri.info/rakentaminen/ on 23.12.2019. Jokeri Light Rail: Information. Retrieved from https://raidejokeri.info/mika-raide-jokeri/ on 22.12.2019.

Kain, J. F. and Quigley, J. M. (1975), "A theory of urban housing markets and spatial structure", in *Housing markets and racial discrimination: a microeconomic analysis*, pp. 9–55, National Bureau of Economic Research.

Ke, Y. and Gkritza, K. (2019), "Light rail transit and housing markets in Charlotte-Mecklenburg County, North Carolina: Announcement and operations effects using quasi-experimental methods", *Journal of Transport Geography*, 76, pp. 212–220.

Laakso, S. (1992), "Public transport investment and residential property values in Helsinki", *Scandinavian Housing and Planning Research*, 9(4), pp. 217–229.

Laakso, S. and Loikkanen, H. A. (2013), "Helsingin seudun maankäyttö, kiinteistömarkkinat ja perusrakenteen rahoitus", *Kansantaloudellinen aikakauskirja*, 109(4), pp. 490–511.

Laakso, S. and Loikkanen, H. A. (2016), "Tiivistyvä kaupunkikehitys. tuottavuuden ja hyvinvoinnin kasvun perusta", Tehokkaan Tuotannon Tutkimussäätiö, Helsinki.

Lewis, G. and Bajari, P. (2011), "Procurement contracting with time incentives: Theory and evidence", *The Quarterly Journal of Economics*, 126(3), pp. 1173–1211.

McDonald, J. F. and Osuji, C. I. (1995), "The effect of anticipated transportation improvement on residential land values", *Regional science and urban economics*, 25(3), pp. 261–278.

Mohammad, S. I., Graham, D. J. and Melo, P. C. (2017), "The effect of the Dubai Metro on the value of residential and commercial properties", *Journal of Transport and Land Use*, 10(1), pp. 263–290.

Mohammad, S. I., Graham, D. J., Melo, P. C. and Anderson, R. J. (2013), "A meta-analysis of the impact of rail projects on land and property values", *Transportation Research Part A: Policy and Practice*, 50, pp. 158–170.

Mulley, C. and Tsai, C. H. (2016), "When and how much does new transport infrastructure add to property values? Evidence from the bus rapid transit in Sydney, Australia", *Transport policy*, 51, pp. 15–23.

Muth, R. F. (1969), *Cities and housing*, University of Chicago Press, Chicago.

Olszewski, P. and Wibowo, S. S. (2005), "Using equivalent walking distance to assess pedestrian accessibility to transit stations in Singapore", *Transportation research record*, 1927(1), pp. 38–45.

Papon, F., Nguyen-Luong, D. and Boucq, E. (2015), "Should any new light rail line provide real estate gains, or not? The case of the T3 line in Paris", *Research in Transportation Economics*, 49, pp. 43–54.

Parmeter, C. F. and Pope, J. C. (2013), "Quasi-Experiments and Hedonic Property Value Methods", in *Handbook on experimental economics and the environment*, eds. List, J. A. – Price, M. K., pp. 3–66, Edward Elgar Publishing, Cheltenham.

Pilgram, C. A. and West, S. E. (2018), "Fading premiums: The effect of light rail on residential property values in Minneapolis, Minnesota", *Regional Science and Urban Economics*, 69, pp. 1–10.

Project assessment of Jokeri Light Rail (2019), FLOU Oy. Retrieved from https://raidejokeri.info/wp- content/uploads/2019/01/RJ_000_KPT-PRJ_Raportti-hankearviointi-ID-24372.pdf on 3.4.2021.

Ransom, M. R. (2018), "The effect of light rail transit service on nearby property values", *Journal of Transport and Land Use*, 11(1), pp. 387–404.

Registry data: SeutuCD'18, HSY and Statistics Finland, 2018.

Rosen, S. (1974), "Hedonic prices and implicit markets: product differentiation in pure competition", *Journal of political economy*, 82(1), pp. 34–35.

Tenkanen, H., Espinosa, J.L., Willberg, E., Heikinheimo, V., Tarnanen, A., Jaakkola, T., Järvi, J., Salonen, M. and Toivonen, T. (2018), Helsinki Region Travel Time Matrix 2018. DOI: 10.13140/RG.2.2.20858.39362

Vuori, P. and Kaasila, M. (2019), "Helsingin ja Helsingin seudun väestöen- nuste 2019–2050", Helsingin kaupunki, Tilastoja, 14. Retrieved from https://www.hel.fi/hel2/tietokeskus/julkaisut/pdf/19_10_25_Tilastoja_14_Vuori_Kaasila.pdf on 6.1.2020.

Weisbrod, G., Mulley, C. and Hensher, D. (2016), "Recognising the complementary contributions of cost benefit analysis and economic impact analysis to an understanding of the worth of public transport investment: A case study of bus rapid transit in Sydney, Australia", *Research in Transportation Economics*, 59, pp. 450–461.

Zhou, Z., Chen, H., Han, L. and Zhang, A. (2019), "The effect of a subway on house prices: Evidence from Shanghai", *Real Estate Economics*, 49(S1), pp. 199–234.

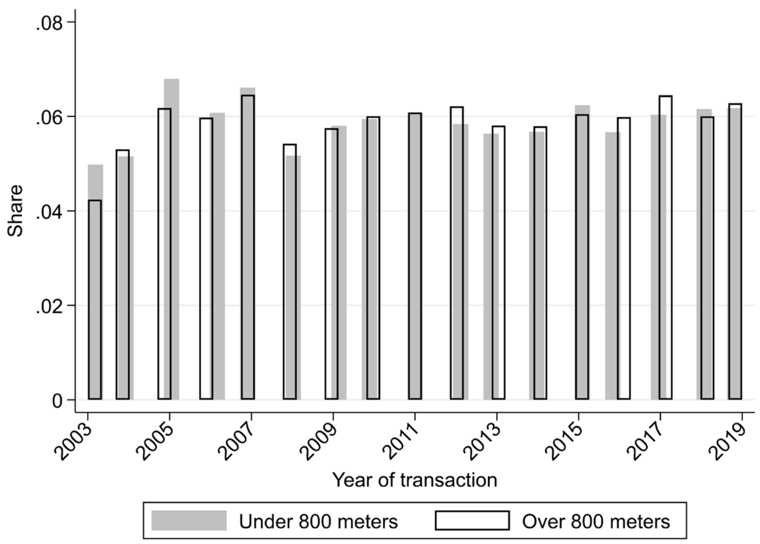
A Appendix: Data appendix

Table A1: *Postal code areas included in this study*

	Helsinki		Espoo
00200	00420	00780	02100
00230	00430	00790*	02110
00240	00440	00800*	02120
00270	00560	00810	02130*
00280	00600	00820	02140*
00300	00610	00830	02150*
00310	00620*	00880*	02160
00320*	00630*	00900*	02180
00340	00640*	00910	02200
00350*	00650*	00920*	02600*
00360*	00660	00930*	02610
00370*	00670	00940	02620
00380*	00680	00950	02630
00390	00700	00980	02650
00400*	00710*		02660
00410	00720		02680
	46 areas		16 areas

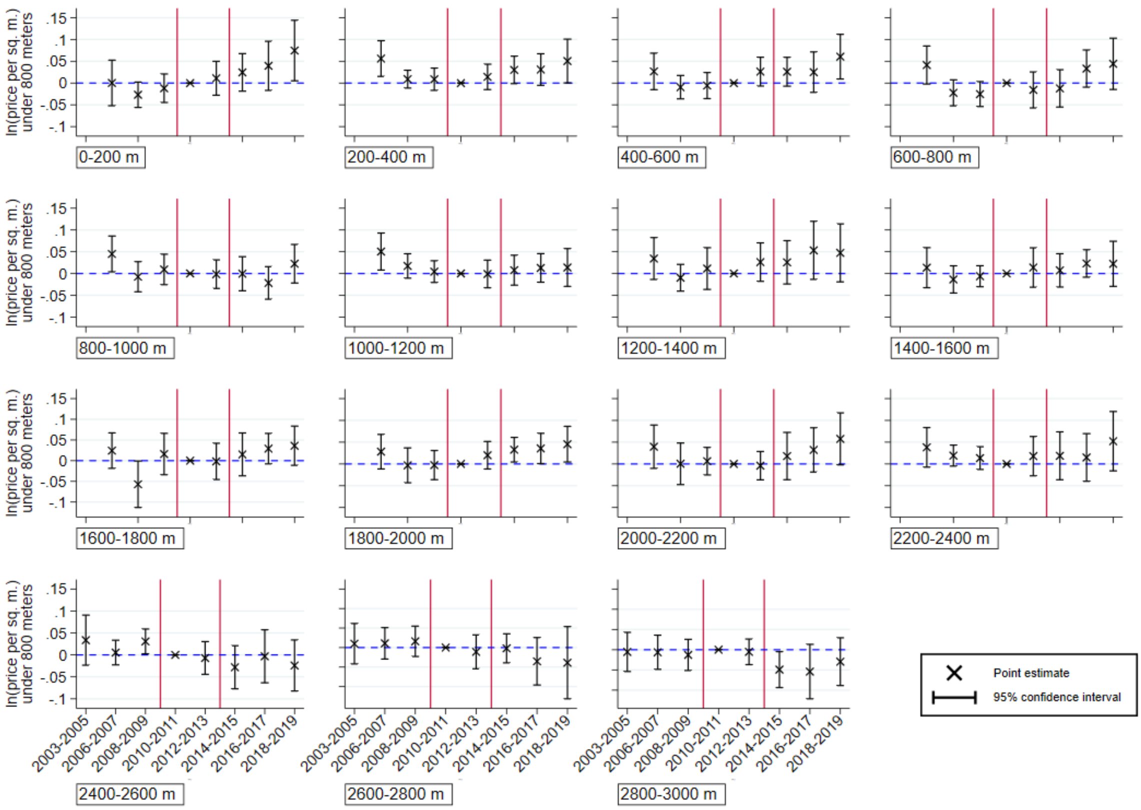
*Note. Postal code areas that contain Jokeri Light Rail stops marked with *.*

Figure A1: *The yearly share of housing transactions in the treatment and control group between 2003 and 2019*



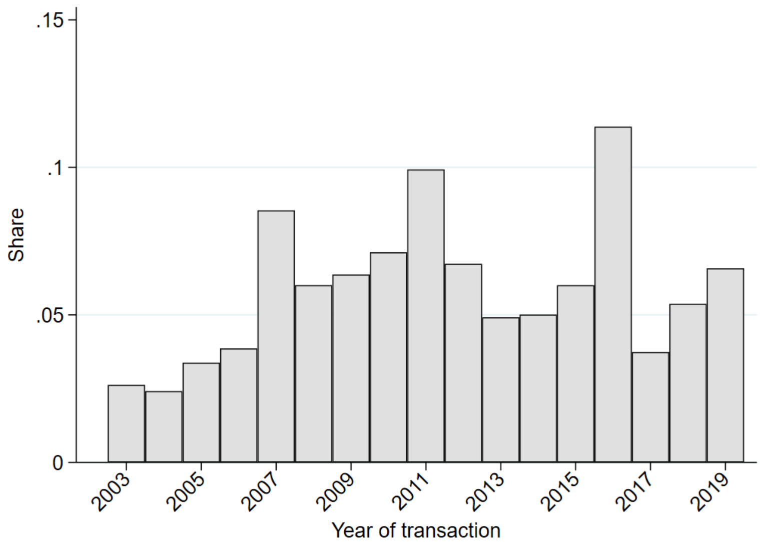
Note. Here, the share of transactions during the whole period sum up to one separately for both distance groups.

Figure A2: *Model 5. Housing market effect in 200-meter bands and year groups*



Note. Model 5 is a modification of model 3d; here, I have divided observations into 200-meter bands and adopt two-year groups instead. Vertical lines represent possible moments of capitalization (see Section 3.1). Zero marked with a blue dashed line.

Figure A3: *The yearly share of housing transactions of new apartments between 2003 and 2019 in postal code areas where Jokeri Light Rail stops are located*



Note. The transactions of December 2019 are excluded due to data availability.

B Appendix: Regression tables for models 1–5

Table B1: *Model 1. Housing market effect in two distance groups before and after 2016*

Response variable: ln(price per square meter)	Reference group: over 800 meters and before 2016			
	(1a)	**(1b)**	**(1c)**	**(1d)**
Under 800 m*After	0.0486* (0.0252)	0.0502* (0.0265)	0.0342* (0.0179)	0.0456** (0.0205)
Under 800 m	-0.00977 (0.0516)	0.00277 (0.0164)	0.00689 (0.0263)	-0.0181 (0.0113)
After	0.139*** (0.0158)	0.144*** (0.0173)	0.117*** (0.0144)	0.139*** (0.0128)
Travel time	–	–	-0.0121*** (0.00211)	-0.00133 (0.00122)
Age	–	–	-0.0183*** (0.00205)	-0.0159*** (0.00121)
Age^2	–	–	0.000246*** (2.90e-05)	0.000180*** (1.78e-05)
Floor area	–	–	-0.00321*** (0.000383)	-0.00340*** (0.000302)
Freehold site	–	–	0.144*** (0.0201)	0.0658*** (0.00845)
Maint. charge (€/m^2)	–	–	0.00321 (0.00419)	0.000955 (0.00298)
Condition FE	No	No	Yes	Yes
Building type FE	No	No	Yes	
Intercept	8.089*** (0.0441)	8.101*** (0.0191)	8.911*** (0.0654)	8.645*** (0.0570)
Postal code FE	No	Yes	No	Yes
N	63,745	63,745	63,745	63,745
R^2	0.061	0.061	0.476	0.316

*Note. Distances measured to the nearest Jokeri Light Rail stop. The sample is constrained to sales in postal code areas reported in Table A1 in Appendix A. Before-period is 2003–2012, after-period is 2016–2019. Observations from 2013–2015 are not used. Control variables correspond to those reported in Table 2, also taking into account floor area squared. The estimated coefficients statistical significance is marked with * (10%), ** (5%) or *** (1%). Standard errors clustered at the postal code level (62) are reported in parentheses.*

Table B2: *Model 2. Housing market effect with continuous distance before and after 2016*

Response variable: ln(price per square meter)	**Reference group: before 2016**					
	(2a) reg	**(2b) xtreg**	**(2b.II) reg with wbs**	**(2c) reg**	**(2d) xtreg**	**(2d.II) reg with wbs**
Distance*After	-0.00309*** (0.000689) [0.000]	-0.00269** (0.00107) –	-0.00269** (0.00107) [0.025]	-0.00157** (0.000727) [0.038]	-0.00207** (0.000888) –	-0.00207** (0.000889) [0.029]
Distance	0.00172 (0.00299)	0.00200 (0.00173)	0.00201 (0.00173)	-0.00150 (0.00117)	-0.000368 (0.00108)	-0.000364 (0.00108)
After	0.204*** (0.0172)	0.205*** (0.0193)	0.205*** (0.0194)	0.152*** (0.0149)	0.179*** (0.0150)	0.179*** (0.0150)
Travel time	–	–	–	-0.0125*** (0.00204)	-0.0078*** (0.00136)	-0.00575*** (0.00137)
Age	–	–	–	-0.0186*** (0.00194)	-0.0176*** (0.00106)	-0.0176*** (0.00106)
Age²	–	–	–	0.000249*** (2.75e-05)	0.000209*** (1.59e-05)	0.000209*** (1.59e-05)
Floor area	–	–	–	-0.00321*** (0.000470)	-0.00331*** (0.000378)	-0.00332*** (0.000377)
Freehold site	–	–	–	0.148*** (0.0225)	0.0742*** (0.0123)	0.740*** (0.123)
Maint. charge (€/m²)	–	–	–	0.00420 (0.00437)	0.00277 (0.00283)	0.00274 (0.00282)
Condition FE	No	No	No	Yes	Yes	Yes
Building type FE	No	No	No	Yes	Yes	Yes
Intercept	8.057*** (0.0349)	8.056*** (0.0301)	8.319*** (0.0486)	8.954*** (0.0579)	8.772*** (0.0588)	9.001*** (0.0575)
RT stop FE	No	Yes	Yes	No	Yes	Yes
N	63,745	63,745	63,745	63,745	63,745	63,745
R^2	0.063	0.063	0.417	0.480	0.409	0.642

*Note. reg and xtreg refer to the command used; wbs corresponds to wild bootstrap. Distances measured to the nearest Jokeri Light Rail stop. The sample is constrained to sales in postal code areas reported in Table A1 in Appendix A. Before-period is 2003–2012, after-period is 2016–2019. Observations from 2013–2015 are not used. Control variables correspond to those reported in Table 2, also taking into account floor area squared. The estimated coefficients statistical significance is marked with * (10%), ** (5%) or *** (1%). Standard errors clustered at the LRT stop level (34) are reported in parentheses. The p-value for the wild bootstrap is reported in square brackets.*

Table B3: *Model 3. Yearly housing market effect in two distance groups*

Response variable: ln(price per square meter)	Reference group: over 800 meters and year 2011			
	(3a)	(3b)	(3c)	(3d)
Under 800 m*2003	0.00938 (0.0168)	0.0134 (0.0165)	0.00926 (0.0124)	0.0159 (0.0116)
Under 800 m*2004	0.0170 (0.0202)	0.0150 (0.0172)	0.0103 (0.0166)	0.00885 (0.0143)
Under 800 m*2005	-0.00995 (0.0144)	0.00402 (0.0191)	-0.00172 (0.0119)	0.00397 (0.0145)
Under 800 m*2006	-0.00765 (0.0180)	-0.00861 (0.0149)	-0.0177 (0.0122)	-0.00805 (0.0115)
Under 800 m*2007	0.00551 (0.0159)	-0.00162 (0.0175)	-0.00743 (0.00964)	-0.00381 (0.0136)
Under 800 m*2008	-0.0231 (0.0148)	-0.00801 (0.0163)	-0.0214* (0.0123)	-0.00708 (0.0133)
Under 800 m*2009	-0.0160 (0.0140)	0.00198 (0.0148)	-0.0229** (0.0102)	-0.00258 (0.0111)
Under 800 m*2010	-0.0143 (0.0115)	-0.00834 (0.0125)	-0.0113 (0.00913)	-0.00657 (0.0101)
Under 800 m*2012	-0.00259 (0.0152)	0.0146* (0.00859)	-0.00420 (0.0109)	0.0118* (0.00696)
Under 800 m*2013	0.0214* (0.0117)	0.0185* (0.0103)	0.00381 (0.00937)	0.0145* (0.00853)
Under 800 m*2014	0.0269* (0.0156)	0.0312* (0.0147)	0.00542 (0.0109)	0.0234* (0.0117)
Under 800 m*2015	0.0298 (0.0216)	0.0377** (0.0182)	0.00765 (0.0163)	0.0274* (0.0146)
Under 800 m*2016	0.0122 (0.0222)	0.0225 (0.0217)	0.0104 (0.0157)	0.0224 (0.0162)
Under 800 m*2017	0.0376* (0.0207)	0.0468** (0.0198)	0.0224 (0.0153)	0.0433* (0.0165)
Under 800 m*2018	0.0634** (0.0284)	0.0685*** (0.0228)	0.0399* (0.0219)	0.0593*** (0.0184)
Under 800 m*2019	0.0466* (0.0265)	0.0517** (0.0236)	0.0237 (0.0209)	0.0495** (0.0200)
Under 800 m	-0.00183 (0.0541)	0.00225 (0.0159)	0.0153 (0.0290)	-0.0219* (0.0118)
Control variables	No	No	Yes	Yes
Yearly effects	Yes	Yes	Yes	Yes
Postal code FE	No	Yes	No	Yes
N	77,378	77,378	77,378	77,378
R^2	0.112	0.112	0.525	0.342

*Note. Distances measured to the nearest Jokeri Light Rail stop. The sample is constrained to sales in postal code areas reported in Table A1 in Appendix A. The time period is 2003– 2019. Control variables correspond to those reported in Table 2, also taking into account floor area squared. The estimated coefficients statistical significance is marked with * (10%), ** (5%) or *** (1%). Standard errors clustered at the postal code level (62) are reported in parentheses.*

Table B4: *Model 4. Yearly housing market effect with continuous distance*

Response variable: ln(price per square meter)	Reference group: year 2011					
	(4a) reg	(4b) xtreg	(4b.II) reg with wbs	(4c) reg	(4d) xtreg	(4d.II) reg with wbs
Distance*2003	-0.000652 (0.000650)	-0.000652 (0.000650)	-0.000318 (0.000762)	-0.000936* (0.000525)	-0.000936* (0.000525)	-0.000318 (0.000762)
Distance*2004	-0.00192 (0.00162	-0.00192 (0.00162)	-0.000822 (0.000890)	-0.00193 (0.00126)	-0.00193 (0.00126)	-0.000822 (0.000890)
Distance*2005	0.000499 (0.000634)	0.000499 (0.000634)	-0.000461 (0.00137)	-0.000408 (0.000833)	-0.000408 (0.000833)	-0.000461 (0.00137)
Distance*2006	0.00112* (0.000594)	0.00112* (0.000594)	0.000925 (0.000705)	0.000807* (0.000468)	0.000807* (0.000468)	0.000925 (0.000705)
Distance*2007	0.000402 (0.000487)	0.000402 (0.000487)	0.000349 (0.000697)	9.51e-05 (0.000456)	9.51e-05 (0.000456)	0.000349 (0.000697)
Distance*2008	0.000884 (0.000624)	0.000884 (0.000624)	0.000590 (0.000709)	0.000262 (0.000537)	0.000262 (0.000537)	0.000590 (0.000709)
Distance*2009	0.000653 (0.000730)	0.000653 (0.000730)	0.000423 (0.000840)	0.000571 (0.000543)	0.000571 (0.000543)	0.000423 (0.000840)
Distance*2010	0.000842 (0.000528)	0.000842 (0.000528)	0.000410 (0.000706)	0.000132 (0.000531)	0.000132 (0.000531)	0.000410 (0.000706)
Distance*2012	-0.000699 (0.000740)	-0.000699 (0.000740)	-0.00083 (0.000515)	-0.000202 (0.000444)	-0.000202 (0.000444)	-0.00083 (0.000515)
Distance*2013	-0.00104 (0.000690)	-0.00104 (0.000690)	-0.00106* (0.000486)	-0.000413 (0.000424)	-0.000413 (0.000424)	-0.00106** (0.000486)
Distance*2014	-0.00101 (0.000651)	-0.00101 (0.000651)	-0.00135* (0.000585)	-0.000454 (0.000375)	-0.000454 (0.000375)	-0.00135** (0.000585)
Distance*2015	-0.00172* (0.000968)	-0.00172* (0.000968)	-0.00146 (0.000840)	-0.00105 (0.000646)	-0.00105 (0.000646)	-0.00146* (0.000840)
Distance*2016	-0.00157* (0.000867)	-0.00157* (0.000867)	-0.00186* (0.000831)	-0.00110* (0.000594)	-0.00110* (0.000594)	-0.00186** (0.000831)
Distance*2017	-0.00221*** (0.000786)	-0.00221*** (0.000786)	-0.00205** (0.000826)	-0.00128* (0.000696)	-0.00128* (0.000696)	-0.00205** (0.000826)
Distance*2018	-0.00413*** (0.00109)	-0.00413*** (0.00109)	-0.00311*** (0.000942)	-0.00210** (0.00101)	-0.00210** (0.00101)	-0.00311*** (0.000942)
Distance*2019	-0.00317*** (0.000921) [0.001]	-0.00317*** (0.000921) –	-0.00281*** (0.00100) [0.004]	-0.00172** (0.000824) [0.047]	-0.00172** (0.000824) –	-0.00281*** (0.00100) [0.007]
Distance	0.00142 (0.00304)	0.00142 (0.00304)	0.00180 (0.00189)	-0.00148 (0.00129)	-0.00148 (0.00129)	0.00180 (0.00189)
Control variables	No	No	No	Yes	Yes	Yes
Yearly effects	Yes	Yes	Yes	Yes	Yes	Yes
LRT stop FE	No	Yes	Yes	No	Yes	Yes
N	77,378	77,378	77,378	77,378	77,378	77,378
R^2	0.114	0.114	0.207	0.530	0.530	0.702

*Note. reg and xtreg refer to the command used; wbs corresponds to wild bootstrap. Distances measured to the nearest Jokeri Light Rail stop. The sample is constrained to sales in postal code areas reported in Table A1 in Appendix A. The time period is 2003–2019. Control variables correspond to those reported in Table 2, also taking into account floor area squared. The estimated coefficients statistical significance is marked with * (10%), ** (5%) or *** (1%). Standard errors clustered at the LRT stop level (34) are reported in parentheses. The p-value for the wild bootstrap is reported in square brackets.*

Table B5: *Model 4. Housing market effect in 200-meter bands*

Response variable: ln(price per square meter)	**Reference group: over 3 000 meters and 2010–2011**						
	2003–2005	**2006–2007**	**2008–2009**	**2012–2013**	**2014–2015**	**2016–2017**	**2018–2019**
0–200 m	0.00293 (0.0267)	-0.0268* (0.0149)	-0.0118 (0.0167)	0.0110 (0.0199)	0.0245 (0.0220)	0.0395 (0.0287)	0.0747** (0.0355)
200–400 m	0.0562** (0.0209)	0.00899 (0.0104)	0.00883 (0.0131)	0.0142 (0.0149)	0.0302* (0.0161)	0.0310* (0.0185)	0.0506** (0.0256)
400–600 m	0.0267 (0.0214)	-0.00942 (0.0137)	-0.00573 (0.0153)	0.0263 (0.0169)	0.0260 (0.0169)	0.0251 (0.0237)	0.0606** (0.0262)
600–800 m	0.0415* (0.0223)	-0.0224 (0.0152)	-0.0253* (0.0146)	-0.0158 (0.0211)	-0.0122 (0.0219)	0.0333 (0.0218)	0.0442 (0.0300)
800–1 000 m	0.0450** (0.0210)	-0.00740 (0.0177)	0.00937 (0.0178)	-0.00151 (0.0167)	-0.000576 (0.0200)	-0.0216 (0.0191)	0.0224 (0.0226)
1 000–1 200 m	0.0504** (0.0217)	0.0173 (0.0143)	-0.000432 (0.0127)	-0.00104 (0.0162)	0.00759 (0.0176)	0.0128 (0.0167)	0.0138 (0.0222)
1 200–1 400 m	0.0344 (0.0245)	-0.00974 (0.0156)	0.0114 (0.0244)	0.0262 (0.0225)	0.0256 (0.0254)	0.0532 (0.0340)	0.0473 (0.0340)
1 400–1 600 m	0.0133 (0.0235)	-0.0136 (0.0159)	-0.00624 (0.0123)	0.0137 (0.0231)	0.00717 (0.0195)	0.0231 (0.0161)	0.0222 (0.0264)
1 600–1 800 m	0.0243 (0.0219)	-0.0571** (0.0286)	0.0160 (0.0255)	-0.00180 (0.0224)	0.0152 (0.0264)	0.0291 (0.0188)	0.0361 (0.0242)
1 800–2 000 m	0.0282 (0.0202)	-0.00320 (0.0204)	-0.00264 (0.0171)	0.0200 (0.0161)	0.0327** (0.0144)	0.0356** (0.0178)	0.0454** (0.0209)
2 000–2 200 m	0.0396 (0.0254)	0.00347 (0.0243)	0.00634 (0.0161)	-0.00402 (0.0166)	0.0181 (0.0277)	0.0322 (0.0259)	0.0575* (0.0304)
2 200–2 400 m	0.0381 (0.0232)	0.0191 (0.0123)	0.0134 (0.0133)	0.0179 (0.0230)	0.0187 (0.0280)	0.0149 (0.0279)	0.0522 (0.0348)
2 400–2 600 m	0.0336 (0.0291)	0.00535 (0.0143)	0.0307** (0.0144)	-0.00708 (0.0190)	-0.0281 (0.0250)	-0.00317 (0.0308)	-0.0241 (0.0297)
2 600–2 800 m	0.0983 (0.0264)	0.0108 (0.0206)	0.0160 (0.0198)	-0.0111 (0.0222)	-0.00199 (0.0190)	-0.0352 (0.0310)	-0.0388 (0.0473)
2 800–3 000 m	-0.00550 (0.0247)	-0.00629 (0.0215)	-0.0130 (0.0196)	-0.00527 (0.0161)	-0.0492** (0.0227)	-0.0540 (0.0344)	-0.0296 (0.0301)
Control variables				Yes			
Distance group coeffs				Yes			
Year group FE				Yes			
Postal code FE				Yes			
N				77,378			
R^2				0.536			

*Note. The interaction terms of model 5 are shown in five columns, one column for each year group. Distances measured to the nearest Jokeri Light Rail stop. The sample is constrained to sales in postal code areas reported in Table A1 in Appendix A. The time period is 2003–2019. Control variables correspond to those reported in Table 2, also taking into account floor area squared. The estimated coefficients statistical significance is marked with * (10%), ** (5%) or *** (1%). Standard errors clustered at the postal code level (62) are reported in parentheses.*

Nowcasting Finnish GDP growth using financial variables: a MIDAS approach*

Olli-Matti Laine
Bank of Finland and Tampere University
olli-matti.laine@bof.fi

Annika Lindblad
Bank of Finland
annika.lindblad@bof.fi

Abstract

We analyse the performance of financial market variables in nowcasting Finnish quarterly GDP growth. In particular, we assess if prediction accuracy is affected by the sampling frequency of the financial variables. Therefore, we apply MIDAS models that allow us to nowcast quarterly GDP growth using monthly or daily data without temporal aggregation in a parsimonious way. We find that financial market data nowcasts Finnish GDP growth relatively well: nowcasting performance is similar to industrial production, but financial market data is available much earlier. Our results suggest that the sampling frequency of financial market variables is not crucial: nowcasting accuracy of daily, monthly and quarterly data is similar.

Keywords: *MIDAS, Nowcasting, Financial markets, GDP*
JEL codes: *E44, G00, E37*

**We thank Eleonora Granziera and two anonymous referees for useful and insightful comments on this research. We also thank Rasmus Rannikko for helping us with the data. This paper represents the views of the authors which are not necessarily those of the Bank of Finland. Any remaining errors are solely our own responsibility.*

1. Introduction

Financial markets provide high-frequency information about investors' expectations. This information may be useful in predicting or nowcasting GDP growth because asset prices are based on expected future cash flows, which in turn are linked to macroeconomic conditions.[1] In situations where the economic environment is changing quickly, daily or weekly updates on economic conditions can be crucial for forming an accurate and timely view of the economy.

The previous literature has shown that including financial market variables in short-term forecasting or nowcasting models for economic activity is useful (e.g., Friedman and Kuttner, 1993; Estrella and Mishkin, 1998; Henry, Olekalns and Thong, 2004; Chionis, Gogas and Pragidis, 2010; Nyberg, 2010). Junttila and Korhonen (2011) show that dividend yields and short-term interest rates are relevant for forecasting output growth in multiple countries. According to their results, financial variables are especially useful in turbulent times. Kuosmanen and Vataja (2014) find similar results for Finland. Although many studies provide evidence of the usefulness of financial market information in forecasting, Alessi, Ghysels, Onorante, Peach and Potter (2014) argue that central banks have not utilised this information in the best possible way.

Despite many studies analysing the predictive power of financial market variables, the optimal sampling frequency for financial market data in a nowcasting model for GDP has not been widely studied. Should we care about, for example, daily fluctuations, or is it better to consider temporally aggregated data? Typically, if the variables in a model are measured at different frequencies, the high-frequency variables are aggregated to the level of the lowest frequency variable. However, temporal aggregation inevitably ignores some of the information in the daily financial market data. Some of this information could, nevertheless, be important for nowcasting GDP growth and ignoring it could confound the relationship between the variables (see, e.g., Lütkepohl (2010) for a discussion on forecasting with aggregated data). Time series data can have a decaying memory structure, and thus giving the same weight to each daily observation within a quarter is not necessarily optimal. Higher frequency data also enables nowcasts to be updated frequently and during the on-going quarter, which is naturally a crucial feature of nowcasting models. This could be especially useful during the beginning of a recession, when traditional macroeconomic variables are slow to react and additionally suffer from publication lags. On the other hand, daily data tends to be noisier than temporally aggregated data. As a lower signal-to-noise ratio could obscure the relationship between the variables and therefore lead to a deterioration in forecast accuracy, high-frequency data does not automatically improve predictions. It is therefore clear that the sampling frequency could impact the accuracy of a nowcasting model. Analysing the effect of the sampling frequency is also important for understanding the relationship between variables. Due to the availability of daily data and mixed-frequency methods, it is possible to include high-frequency data in a quarterly model for GDP growth and provide policy makers with (close to) real-time GDP nowcast updates.

With the development of modelling frameworks being able to handle data sampled at different frequencies, the question of whether using higher frequency data directly to nowcast and forecast lower frequency data improves accuracy has naturally arisen. Ghysels, Santa-Clara and Valkanov (2004) propose a Mixed Data Sampling (MIDAS) regression framework to overcome the issues related to different sampling frequencies without relying on temporal aggregation. This framework has turned out to be useful in forecasting returns and volatility in financial markets. In these applications, the sampling frequency of the explanatory variable seems to have an impact on the results (Ghysels, Sinko and Valkanov, 2007). For example, Ghysels, Santa-Clara and Valkanov (2006) conclude that intra-daily data does not outperform daily data (aggregated from the high-frequency data) in forecasting realized volatility. They also show that the most recent days are the most important when predicting monthly or weekly financial market volatility.

MIDAS models have also been widely used to nowcast and forecast quarterly GDP growth using mainly monthly predictors. For example, Clements and Galvão (2008, 2009) (using US data), Marcellino and Schumacher (2010) (using German data) and Kim and Swanson (2017) (using Korean data) show that monthly macroeconomic data, and especially monthly data on the current quarter, improves quarterly output growth forecasts in different MIDAS specifications. In addition,

[1] *Nowcasting means the prediction of the current state of the variable of interest. Many economic time series are published with a substantial lag. Therefore, economists do not only predict the future but also the present and the past.*

Armesto, Engemann and Owyang (2009) concludes that using a MIDAS weighting scheme for monthly employment growth improves forecast accuracy over an equal-weighted quarterly average when forecasting US GDP growth, especially over a short forecasting horizon. There is, therefore, some evidence that increasing the sampling frequency from quarterly to monthly improves nowcasting performance.

The literature closest to our paper is that which considers the relationship between financial market data and output growth in a mixed frequency setting. For example, Andreou, Ghysels and Kourtellos (2013) (using US data), Ferrara and Marsilli (2013) (using euro area data and data for the four largest euro area countries), Ferrara, Marsilli and Ortega (2014) (using US, UK and French data) and Marsilli (2014) (using US data) have shown using MIDAS models that daily or monthly financial market variables include useful information for GDP nowcasts and forecasts not included in macroeconomic data. However, these papers do not focus on assessing the relevance of the sampling frequency or discuss in detail the relationship between individual financial variables and GDP growth. In other words, most of the earlier literature does not consider whether, for example, daily or monthly variation in asset prices includes useful information for nowcasting GDP growth not included in the same data aggregated to a lower frequency. For example, a decline in asset prices mid-quarter could provide a useful indication of near-term economic activity, which could be obscured if a quarterly average was used. The choice of sampling frequency is, however, discussed in Tsui, Xu and Zhang (2018), who conclude that the best frequency of stock market returns is weekly when forecasting output growth for Singapore with a MIDAS model. In addition, Gómez-Zamudio and Ibarra (2017) (using Mexican data) and Doğan and Midiliç (2019) (using Turkish data) conclude that increasing the frequency of financial market data from quarterly to daily improves the accuracy of output growth forecasts using MIDAS and factor MIDAS models. On the other hand, Tay (2007) finds that using daily stock market returns aggregated over a year generally leads to better output growth nowcasts for the US than using a non-parametric MIDAS specification. The evidence regarding the optimal sampling frequency of different financial variables is thus scarce. In addition, most of the previous literature does not explore whether any specific days within a quarter are more important than others. Thus, based on the previous literature it is not possible to conclude whether an equal-weighted average is a suitable aggregation method for financial data when nowcasting GDP growth.

In this paper we study the usefulness of including daily financial market variables in a nowcasting model for quarterly Finnish GDP growth in MIDAS framework. We also consider whether the choice of sampling frequency matters for nowcasting Finnish GDP growth. Determining the impact of the sampling frequency amounts to studying whether any important information available in the higher frequency data is lost in the aggregation. However, as seen in the previous literature, there could be a point beyond which increasing the frequency of the explanatory data is not beneficial. Determining this point is an empirical question and therefore specific to the country and variable under consideration. Here we consider the question for Finnish GDP growth using a relatively broad range of both domestic and foreign financial variables. MIDAS models allow us to study how temporal aggregation of variables affects prediction accuracy and it allows us to determine whether a simple average is a satisfactory aggregation method, or whether a more flexible weighting scheme is necessary to correctly characterise the relationship.

We also assess different ways to utilise financial market variables in the MIDAS framework. The earlier literature has considered, for example, individual variables, forecast combinations and principal components (Andreou et al., 2013; Ferrara et al., 2014). The previous papers discussing the gains of high-frequency data for nowcasting GDP growth focus on the US or other large economies. In order to broaden the literature and the applicability of previous results, our aim here is to provide evidence of these potential gains from the perspective of a small open economy. Therefore, we also include foreign variables, namely from Germany and the US, as the international financial markets might include important information for nowcasting growth in a small open economy.

Our results suggest that it does not matter significantly and systematically for nowcast accuracy whether one uses financial market data at the daily, monthly or quarterly frequency when nowcasting Finnish GDP growth. However, there may be some practical reasons to prefer higher frequency data, such as the ability to update the nowcast on a monthly, weekly or even daily frequency. On the other hand, increasing the frequency also brings some challenges, in particular by increasing the noisiness of the data. Ultimately, our results suggest that the choice of frequency can be made based on data availability and the needs of the forecaster, without significantly compromising nowcast accuracy.

To gauge the importance of financial market variables for nowcasting GDP growth we compare their nowcasting ability to that of industrial production growth, which is a traditional predictor of GDP growth. Our results imply that financial market variables predict GDP growth as accurately as industrial production. Because industrial production is observed with more than a one-month lag, the results suggest that we can, without loss of accuracy, nowcast GDP earlier using financial variables. Different kinds of financial ratios – like the dividend yield – nowcast Finnish GDP growth well, which is in line with the results in Junttila and Korhonen (2011). Our results also provide some evidence that nowcast accuracy can be improved by combining a financial market based nowcasts to a nowcast based on macroeconomic data. However, these improvements are not statistically significant.

The rest of the article is organised as follows. Section 2 introduces the MIDAS regression framework. Section 3 summarises the data. Section 4 studies the nowcasting performance of individual financial market variables and considers the effects of increasing the frequency of the financial market data. Section 5 discusses the best way of utilising financial market data to nowcast Finnish GDP growth. Section 6 concludes.

2. The MIDAS framework

MIDAS models were introduced by Ghysels et al. (2004), Ghysels et al. (2005), Ghysels et al. (2006) and Ghysels et al. (2007). The central idea of the MIDAS approach is to explain a low-frequency variable by variables sampled at higher frequencies. The MIDAS framework is used in this paper because it is a simple nowcasting framework which allows the parsimonious inclusion of several lags of the explanatory data and enables data sampled at different frequencies to be included into the same model.[2] Although the modelling framework is linear, the weight functions allow complex relationships between the dependent and independent variables. An important and useful feature of the MIDAS framework is that the number of parameters to be estimated does not depend on the sampling frequency of the data or the length of the sample period. The chosen framework therefore enables us to compare in a straightforward way the nowcasting performance of several variables and compare their individual performance when sampled at different frequencies.

The standard MIDAS model with one explanatory variable can be written as follows:

$$y_t = \beta_0 + \beta_1 \sum_{h=0}^{d} \theta_h \, x_{tm-h} + u_t, \tag{1}$$

where y_t is a low-frequency variable (GDP growth in our models), x_{tm-h} is a high-frequency variable (a financial market variable or industrial production growth in our models) and d is the number of lags of the explanatory data included in the model.[3] There are m observations of the high-frequency variable to one observation of the low-frequency variable. For example, if we explained a quarterly variable by a monthly variable, m would be 3. If the number of lags, d, was 2, then we would explain the quarterly variable by all the monthly observations of the high-frequency variable from the given quarter. If $\beta_1 \neq 0$, there is a connection between the low-frequency and the high-frequency variables. The function θ_h is a polynomial that weights the contemporaneous observation of the high-frequency variable and its lags in a parsimonious way. In this paper we use the (normalised) exponential Almon lag polynomial:

[2] *The MIDAS model is not the only mixed-frequency method available (see, for example, Foroni and Marcellino (2013) for a survey on various mixed-frequency methods). For example, Kuzin, Marcellino and Schumacher (2011) compares the MIDAS model and a mixed-frequency VAR (MF-VAR) approach for nowcasting and forecasting monthly GDP in the euro area. They find in their empirical application that the MIDAS model tends to perform better for shorter horizons, while the MF-VAR model performs better for longer horizons. On the other hand, Franta, Havrlant and Rusnák (2016) finds that the dynamic factor model with mixed frequency data performs better than MIDAS models for nowcasting GDP in the Czech Republic. The results are similar in Galli, Hepenstrick and Scheufele (2019) for Switzerland, except that the MIDAS models slightly outperform the dynamic factor model after the financial crisis.*

[3] *The AR-MIDAS model by Andreou et al. (2013), which includes lagged values of the dependent variable, could be used for determining whether financial market data include useful information for nowcasting GDP in addition to lagged GDP. However, in this paper we concentrate on evaluating the usefulness of different financial market variables and the impact of varying their sampling frequency in a MIDAS model for GDP. Thus, exploring the AR-MIDAS model in this context is left for future work.*

$$\theta_h = \frac{e^{\lambda_1(h+1)+\lambda_2(h+1)^2}}{\sum_{s=0}^{d} e^{\lambda_1(s+1)+\lambda_2(s+1)^2}}. \quad (2)$$

The normalisation ensures that the weights sum up to one, which allows the separate identification of β_1. The exponential Almon lag polynomial allows flexible, such as decaying or hump-shaped, weighting schemes. Parameters λ_1 and λ_2 are estimated simultaneously with the other model parameters, and together with the number of lags they govern the shape of the weighting scheme. If $\lambda_1 = \lambda_2 = 0$, the lags have equal weights ($1/d$) and we essentially include a moving average of the past d lags in the MIDAS model. Note that in this case using temporally aggregated data, by taking an average over the lowest frequency, yields the same result. The benefit of the MIDAS framework over temporal aggregation is thus that the MIDAS model allows the data to decide the weights and therefore enables taking into account the decaying memory structure often present in time series data. The MIDAS model is estimated using non-linear least squares (NLS).

Due to the lag polynomial, MIDAS models are especially useful when the number of lags is large, as they allow including, for example, many daily lags without increasing the parameter space. However, when only a few lags are included an unrestricted MIDAS (U-MIDAS) model, which does not include a weighting scheme but estimates a separate regression parameter for each lag, can be used (Foroni, Marcellino and Schumacher, 2015). The U-MIDAS regression model can be written as:

$$y_t = \beta_0 + \sum_{h=0}^{d} \beta_h \, x_{tm-h} + u_t. \quad (3)$$

In this case the model can be estimated using OLS.

3. Data

We use daily, monthly and quarterly data from Q2/2002 to Q3/2019. The beginning of the sample period is, firstly, restricted by the availability of daily data. For example, many financial ratios are available to us from 2002 onwards. Secondly, using data only from the early 21st century reduces the risk of structural breaks. This sample period does, however, include two recessions for Finland and thus covers, in our opinion, enough business cycle variation for drawing conclusions. The Finnish GDP data is from Statistics Finland. We use the latest vintage of the quarterly real GDP growth rate.[4]

We use a comprehensive set of 28 financial market variables, covering both domestic and foreign data. All the financial data is obtained via Bloomberg. Andreou et al. (2013) uses a much wider data set of 991 daily series, but, for example, Ferrara et al. (2013), Ferrara et al. (2014) and for Finland Kuosmanen and Vataja (2014) use a significantly narrower set of financial market data. We assess the predictive power of stock indices and interest rates from Finland (OMX Helsinki), Germany (DAX) and the USA (S&P 500). In addition, we use the average price-to-earnings ratios, price-to-book ratios and dividend yields from the same countries. The averaging utilises the same weights that have been used in the construction of the stock market indices. The ratios are calculated using past information about earnings and dividends from the past 12 months.[5] The last available information is used for the book value. The predictive power of the oil price, expected stock market volatility implied by Eurostoxx 50 index options and the EUR/USD exchange rate are also considered.

The stock indices and the oil price are in log-differences. Interest rates and the exchange rate are in differences. We also calculate the spread between the German 10-year yield and the 12-month yield. Financial ratios, stock market volatility and the interest rate spread are in levels. For comparison, we nowcast GDP growth also using the monthly growth rate

---

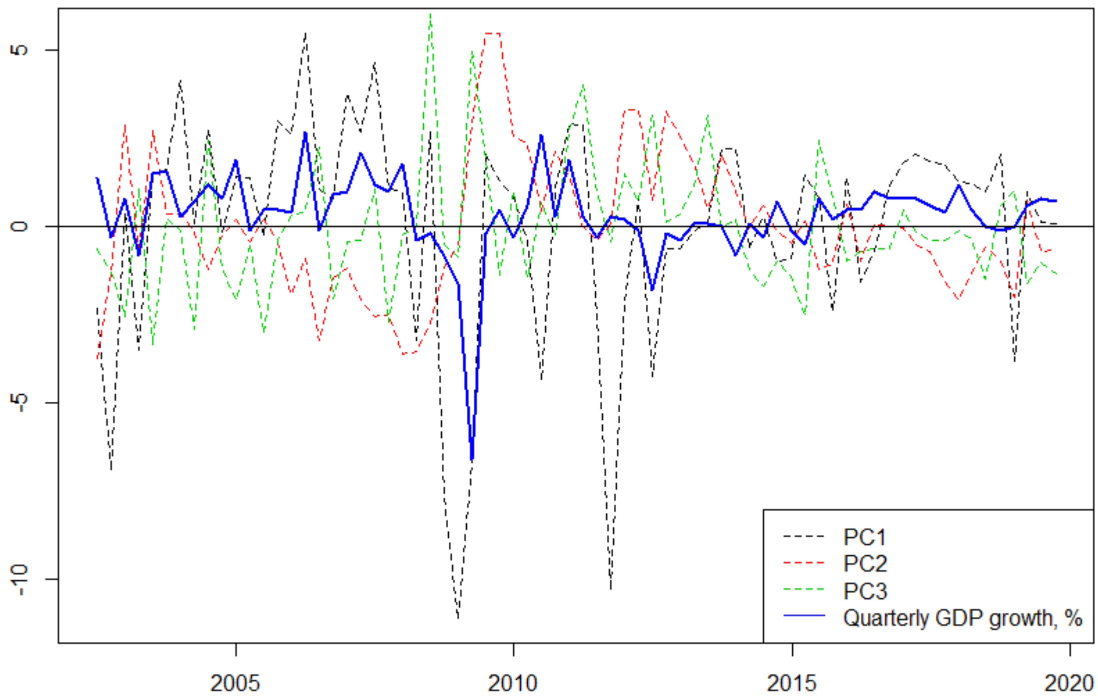
[4] *GDP statistics are typically revised substantially. We use the latest available vintage at the time of writing (data collected on December 31st, 2019) as we consider it the best available estimate for the actual, final growth rate. However, we recognize that there are also arguments for using an earlier vintage, such as first release data.*

[5] *This calculation method causes some minor revisions to financial ratios. The series are recalculated every day using the latest information about the past equity level fundamentals. This may potentially overstate the nowcasting performance of financial ratios.*

of industrial production, for which the latest vintage is used.[6] The transformations are done to achieve stationarity and are based on previous studies (for example, Bernanke, Boivin and Eliasz, 2005; Becker, Lee and Gup, 2012; Lewellen, 2004; Koustas and Serletis, 2005; Marcellino and Schumacher, 2010). The time series included in the data set are listed (together with their sources and transformations) and plotted in Appendix A.

To summarise the financial market information, we use principal component analysis to extract common factors from the financial market data. Figure 1 plots the quarterly GDP growth rate together with the first three principal components (PCs) based on standardised data (excluding industrial production). The principal components are here calculated from quarterly data for visual reasons. The principal components capture the common variation on the financial markets over time. Figure 1 shows that the relationship between financial markets and GDP growth seems to be time-varying and at its strongest during turbulent times.[7] Especially the first PC seems to capture well some of the (negative) spikes in GDP growth. Corresponding figures for the individual financial market variables can be found in Appendix A.

Figure 1. *This figure shows the development of the first three principal components together with quarterly GDP growth. The principal components are calculated from the data summarised in Appendix A (excluding industrial production). The sample is from Q2/2002 to Q3/2019 and the frequency is quarterly.*



[6] *Using revised data may potentially overstate the nowcasting performance of industrial production.*

[7] *The first PC has a correlation of 0.48 with GDP growth, the second PC a correlation of -0.24, and the third PC a correlation of -0.35.*

4. Nowcasting GDP growth using financial market data

In this section we assess how well different variables sampled at different frequencies nowcast GDP growth. We conduct a rolling window analysis, where the first estimation sample is from Q2/2002 to Q4/2011 and the first out-of-sample observation is Q1/2012. We have chosen 2011 to be the end of the first estimation sample in order to see how well the models predict the sharp decline in Finnish quarterly GDP growth in Q2 2012.[8] Altogether we produce 31 pseudo out-of-sample nowcasts, which, although a relatively small number of out-of-sample observations, is of similar length as the out-of-sample period in, for example, Andreou et al. (2013). To plausibly estimate the parameters for the first out-of-sample observation, the length of the out-of-sample period cannot be significantly increased, as increasing the number of out-of-sample observations naturally decreases the number of observations used in the estimation. As a robustness check Appendix E reports the results when 40 out-of-sample observations are used.

In models using daily data we include 62 daily lags. In our sample, every quarter includes at least 63 working days. By choosing 62 daily lags we avoid technical issues relating to "missing data" in the estimation.[9] In models with monthly data the number of lags is two and in models with quarterly data the number of lags is zero, which means that only the contemporaneous quarter is included. Regardless of the frequency used, all models therefore rely on the same information set and utilise data from one quarter only. We include additional lags in Section 5.

The models using quarterly data are naturally unrestricted as there are no lags. For the monthly models we consider both the restricted and the unrestricted versions as the number of lags is small. The daily models are restricted, as it is infeasible to estimate 62 individual regression parameters. The parameters are re-estimated every period in the rolling window analysis. Appendix B reports some estimation results for the whole sample (Q2/2002 to Q3/2019).

[8] *As Q2/2012 can be considered an outlier, Appendix E reports the results excluding the out-of-sample forecast for Q2/2012. These results are discussed briefly in Section 5.*

[9] *We use the R package midasr by Virmantas Kvedaras and Vaidotas Zemlys-Balevicius to estimate the models (see also Ghysels, Kvedaras and Zemlys, 2016).*

Table 1: *Out-of-sample RMSEs of MIDAS regression models. The models are ordered based on the RMSEs of the quarterly models. The abbreviations for the variables are explained in Appendix A. The RMSEs are calculated from a rolling window analysis, in which the first estimation sample is from Q2/2002 to Q4/2011 and the first out-of-sample nowcast is Q1/2012. Thus, the results are based on 31 out-of-sample observations.*

Explanatory variable	Quarterly	Monthly	Monthly (unrestricted)	Daily
SP500 p/e	0.52	0.55	0.56	0.55
DAX Dividend yield	0.58	0.58	0.60	0.58
OMX Hels Dividend yield	0.60	0.66	0.66	0.61
OMX Hels p/b	0.60	0.63	0.72	0.64
PC1	0.63	0.61	0.61	0.64
SP500 p/b	0.64	0.66	0.76	0.67
OMX Hels p/e	0.65	0.69	0.64	0.68
FI_10y	0.66	0.85	0.95	0.67
OIL	0.66	0.66	0.70	0.67
DE 10y-1y	0.66	0.66	0.65	0.66
DE_10y	0.67	0.77	0.91	0.66
DE_5y	0.67	0.72	0.72	0.68
DE_7y	0.67	0.77	0.83	0.68
EURUSD	0.67	0.77	0.78	0.67
DAX p/e	0.68	0.67	0.85	0.68
Industrial production	0.68	0.62	0.66	–
OMX Hels Telec	0.69	0.73	0.77	0.83
DE_1y	0.70	0.64	0.71	0.65
FI_5y	0.70	0.76	0.79	0.69
OMX Hels Industrials	0.71	0.73	0.75	0.70
OMX Hels Utilities	0.71	0.66	0.71	0.77
SP500 Dividend yield	0.71	0.70	0.73	0.70
DAX p/b	0.72	0.67	0.80	0.78
OMX Hels Hlth Care	0.72	0.81	0.99	0.67
OMX Hels Technology	0.72	0.70	0.85	0.75
SP500	0.74	0.77	0.72	0.64
OMX Hels	0.75	0.71	0.79	0.67
Eurostoxx 50 volatility	0.77	0.74	0.78	0.77
DAX	0.84	0.78	0.84	0.65
OMX Hels Basic Metal	0.84	0.82	0.87	0.72

Table 1 reports average root-mean-square errors (RMSEs) for the whole out-of-sample period. The predictors are ordered based on the RMSEs of the quarterly models. The average price-to-earnings ratio in the United States has, perhaps surprisingly, been the best predictor for Finnish GDP growth, regardless of the frequency used. The price-to-earnings and price-to-book ratios especially for the Finnish and the US stock markets have also performed well. The dividend yields for the Finnish and German stock markets produce accurate nowcasts, while implied volatility, the EUR/USD exchange rate, and the stock indices tend to perform worse. However, some of the results are sensitive to the sampling frequency of the explanatory variables. For example, industrial production performs relatively better on the monthly frequency, while the S&P 500, OMX Helsinki and DAX indices seem to improve as predictors as the sampling frequency increases. The relative performance of the various interest rates also varies depending on the sampling frequency.

In addition to the individual financial market variables, we also include the first principal component based on the financial market data. It also produces accurate nowcasts, indicating that combining information from a large set of financial market indicators provides useful information for nowcasting Finnish GDP growth. At least the average price-to-earnings ratio in the United States, the German dividend yield and the first PC produce more accurate nowcasts on the monthly frequency than the benchmark model, where industrial production growth is used as a predictor. Thus, regardless of the frequency, some financial variables seem useful for nowcasting Finnish GDP growth, in particular considering that they are available much earlier than industrial production.

In a MIDAS model the impact of an explanatory variable is determined by β_1 in equation (1) as well as by the weighting scheme (θ_h) in equation (2). Their combined effect (β_1* θ_h) is shown in Appendix B for the daily and monthly models reported in Table 1. For several daily financial variables most of the weight is on just the last few days of the quarter. This implies that the daily intra-quarterly variation is not a significant factor for explaining GDP growth and that all the relevant economic information is included in the recent values of the variables. For some variables, such as the 5-year Finnish interest rate, all weight is on the first days of the quarter, implying that potentially a longer lag length would be needed. For series with counterintuitive weighting schemes there could also be an issue with too much noise in the data for the Almon polynomial to be able to discern a stable relationship between the financial variable and GDP growth. Aggregating to a lower frequency could aid in separating the signal from the noise. For some variables, such as the price-to-book ratios and the dividend yields, the weighting schemes are hump-shaped, implying that these variables affect GDP with a lag. The largest weight is attained in all these cases for lags between 10 and 20 days. As expected, the relationship between GDP growth and the price-to-book ratios is positive while the relationship between GDP growth and the dividend yield is negative. For the price-to-earnings ratio we would expect the relationship to be positive, and this is true for the German and US price-to-earnings ratios. Financial market volatility has a negative relationship to GDP growth and the largest weight is on approximately the 50th daily lag. The relationship between the first PC and GDP growth is positive and most of the weight is on lags between approximately 15 and 30 days.

For the monthly models two monthly lags is generally not enough to make the weights decay to zero and thus the daily models seem better specified. On the other hand, for example, for the price-to-book ratios most weight is on the first monthly lag, which is roughly in line with the daily data. For the dividend yields most weight is on the current and first lags, also in line with the daily results. The weighting schemes for the restricted and unrestricted monthly models are mostly similar.

Overall, looking at the weighting schemes it seems that equal weights (such as in the quarterly models) do not represent the relationship between GDP growth and financial market data accurately. Whether this impairs nowcasting performance depends on the variable. As the differences in nowcast accuracy are, however, mostly small, the impact of the sampling frequency on nowcasting performance seems minor.

The results regarding the importance of individual financial market variables are largely in line with the previous literature. For example, the dividend yields perform well, as concluded by Junttila and Korhonen (2011). As a small open economy Finland is significantly affected by global fluctuations, which might explain why foreign variables, such as the average price-to-earnings ratio of the S&P 500, nowcast well. Forecasting US GDP growth Andreou et al. (2013) finds no differences between the importance of various asset classes. Ferrara et al. (2013) use stock returns, oil prices and the term spread on a monthly frequency to nowcast and forecast GDP growth in the euro area, France, Germany, Italy and Spain with mixed results. For nowcasting, stocks and oil prices receive some support in some countries when compared to the benchmark MIDAS model with confidence data.

Some caution is needed when interpreting the results for the models using financial ratios. It is reasonable to assume that, for example, dividends and stock prices have equal growth rates in the long run. Nevertheless, the ratio may deviate from its long run mean very persistently, and there has been some discussion about the stationarity of these variables (e.g., Koustas and Serletis, 2005; Lewellen, 2004; Becker et al., 2012). Therefore, one could argue that financial ratios should not be used in levels but in differences. Table 2 shows that the nowcasting accuracy of these variables tends to be worse when they are considered in log-differences.

One should also note that in this section we only use information from the contemporaneous quarter to nowcast GDP growth. However, some variables (for example, interest rate spreads) could be correlated to GDP growth with a lag. We discuss including further lags in a more aggregated setting in the next section, but for individual variables we do not consider nowcasts using additional lags, as in their case we concentrate on the choice of sampling frequency.

Table 2: *Out-of-sample RMSEs of regression models which have financial ratios or implied volatility as an explanatory variable. The frequency is quarterly. The abbreviations for the variables are explained in Appendix A. The RMSEs are calculated from a rolling window analysis, in which the first estimation sample is from Q3/2002 to Q4/2011 and the first out-of-sample nowcast is Q1/2012. Thus, the results are based on 31 out-of-sample observations.*

	level	**log-difference**
OMX Hels p/e	0.66	0.71
DAX p/e	0.69	0.67
SP500 p/e	0.53	0.71
OMX Hels Dividend yield	0.59	0.73
DAX Dividend yield	0.59	0.80
SP500 Dividend yield	0.72	0.69
OMX Hels p/b	0.60	0.88
DAX p/b	0.71	0.82
SP500 p/b	0.64	0.72
Eurostoxx 50 volatility	0.77	0.71

All in all, using a higher sampling frequency does not seem to clearly and consistently improve nowcasts. The differences in the RMSEs also tend to be small, and whether performance improves or deteriorates as the sampling frequency increases depends on the variable. This contrasts to some extent the results in, for example, Gómez-Zamudio and Ibarra (2017) and Doğan and Midiliç (2019), where on an aggregate level daily financial data improved forecasts compared to quarterly financial data. The average RMSE of our quarterly models is 0.68, of the restricted monthly models 0.70, of the unrestricted monthly models 0.76 and of the daily models 0.68. Importantly, despite the small number of monthly lags the MIDAS model utilising a weighting scheme performs generally better than the unrestricted model.

5. Does financial market data improve GDP nowcasts?

The results in Section 4 showed that financial market variables can be useful for nowcasting Finnish GDP growth. In this section our aim is to find the best nowcasting specification when using a combination of financial market variables in a MIDAS framework. In addition to the first PC, we combine the information in financial market variables by simple forecast averaging, i.e., we take an average of the MIDAS model nowcasts from Section 4. However, as some of the models in Section 4 produced consistently inferior nowcasts, we also combine only the nowcasts produced by MIDAS models driven by financial ratios (price-to-earnings, price-to-book, dividend yield), which were the most accurate class of predictors in Section 4. We also consider how increasing the lag length in the MIDAS models impacts nowcasting performance. As benchmark models we consider the MIDAS model based on monthly growth in industrial production as well as a simple AR(1) model for GDP growth.

We use monthly data as a compromise between daily and quarterly data because our previous results did not suggest a preference for any specific frequency. Monthly data provides relatively timely information, allowing the nowcast to be updated within the quarter, while being less noisy than daily data. In addition, technical issues limited our use of daily data to within-quarter data. As some variables, such as interest rate spreads or stock indices, could affect GDP with a significant lag, it could be important to include lags beyond the current quarter in a nowcasting model. With monthly data we now consider models with 2, 5 and 11 monthly lags. Since we allow very flexible weights in the weighting scheme, and the weights are estimated based on the data, the estimated weight functions will take into account the leading properties of the explanatory data up to eleven months.

Table 3 reports the RMSEs of these models.[10] The average nowcast based on only financial ratios produces the most accurate nowcasts during the out-of-sample period from Q1/2012 to Q3/2019 regardless of the number of lags included. Summarising the information by principal component analysis leads to slightly lower RMSEs than combining all the MIDAS model nowcasts, and the PC based nowcast thus provides a viable alternative if one does not wish to preselect the models to be combined.[11] This finding is in line with Andreou et al. (2013), whose results also show a slight preference to using factors instead of forecast combinations. However, none of the three purely financial market based nowcasts in Table 3 outperforms the nowcast based on industrial production growth in a statistically significant way.[12] In this sample, all the models nowcast GDP roughly as accurately as an AR(1) model (RMSE 0.58).[13]

[10] *All the weight functions are plotted in Appendices B, C, D and G.*

[11] *As pointed out by an anonymous referee, machine learning methods could be used to improve nowcasting performance. For example, Babii et al. (2021) introduces a sparse-group LASSO estimator which enables a large number of predictors to be simultaneously included in a MIDAS model. ML methods could also be used to estimate the weights of a U-MIDAS model. The application of this to the Finnish case is left for future research and an even larger number of explanatory variables, including financial variables which we have in this paper shown to be useful for nowcasting Finnish GDP growth. Here we concentrate on analysing the performance of individual financial market variables in Finland and the importance of the sampling frequency.*

[12] *However, we are aware that tests for statistical significance can be unreliable for small sample sizes, and thus we focus on the ranking of the models. This follows the approach taken in other studies with small sample sizes, such as Andreou et al. (2013) (for their small sample), Ferrara et al. (2013) and Ferrara et al. (2014).*

[13] *It should be noted that, first of all, the latest vintage of the GDP data is used, i.e., the data includes revisions. Thus, the performance of the AR(1) model might be better than it would have been in real time. Secondly, GDP is released with a roughly two-month lag in Finland.*

Table 3: *RMSEs of MIDAS regression models. In the model using industrial production the only explanatory variable is the MoM growth rate of the volume of industrial production. In the PC model, the only explanatory variable is the first principal component of the financial market variables (see Appendix A). 'Average nowcast' is the simple average of all the financial variable based nowcasts (nowcasts produced using the financial market variables listed in Appendix A one at a time). 'Average nowcast based on financial ratios only' is the average of the models in which the explanatory variable is the price-to-earnings ratio, the price-to-book ratio or the dividend yield. The RMSEs are calculated from a rolling window analysis, in which the first estimation sample is from Q1/2003 to Q4/2011 and the first out-of-sample nowcast is Q1/2012. Thus, the results are based on 31 out-of-sample observations. To test the statistical significance of the RMSE differences between the industrial production-based nowcast and the other nowcasts we use the Diebold-Mariano test assuming no heteroscedasticity or autocorrelation because the forecast horizon is zero. Asterisks *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.*

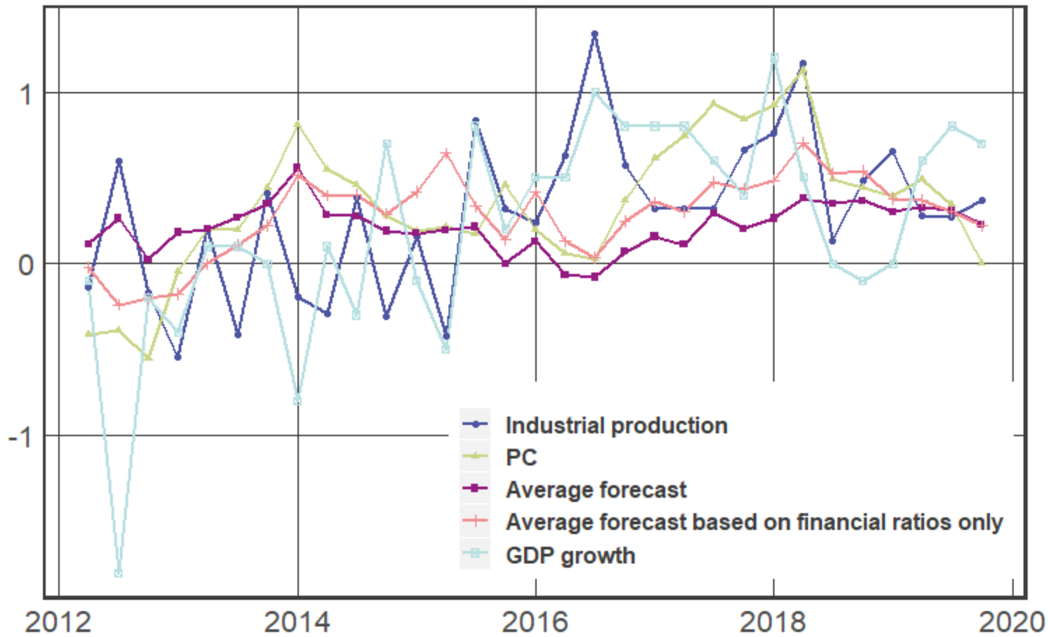
	2 lags	5 lags	11 lags
Industrial production	0.71	0.57	0.60
PC	0.60	0.62	0.59
Average nowcast	0.63	0.66	0.65
Average nowcast: financial ratios only	0.55	0.56	0.58
Average nowcast: financial ratios only and industrial production	0.60*	0.51	0.54

Regarding the choice of lag length, including only two monthly lags of industrial production growth leads to inferior nowcasting performance. From the weight functions (see Appendix G) we can see that when eleven lags are included most of the weight is on lags three and four and the lags after lag seven are virtually zero. This might also explain why increasing the lag length from two to five improves nowcast accuracy, while increasing the lag length from five to eleven does not: lags three and four receive large weights while lags past the fifth lag do not impact the nowcast much anymore compared to the first five lags. For most of the other models in Table 3 nowcasting performance is similar regardless of the lag length chosen. A longer lag length can be generally preferred based on the idea that the exponential Almon polynomial should be able put the weight of any excess lags to zero, and therefore it can be argued that including too many lags is less detrimental than including too few lags. By comparing the weighting schemes plotted in Appendices B, C and D we can clearly see that two monthly lags tends to be too few for the weights to decay to zero, whereas for many variables, such as the first PC, eleven lags seems to be the most suitable lag length. Using elven lags also leads to the lowest RMSE in Table 3 for the model including the first PC. The weight functions are in many cases hump-shaped, implying that the variables are leading indicators compared to GDP. On the other hand, it could be argued that the importance of financial data lies in it being able to quickly reflect the current economic situation and changes to it, which would indicate that putting at least most of the weight on near-term data is desirable. This would advocate including only two lags, as nowcasting performance is similar regardless of the number of included lags.

Financial market data and macroeconomic variables might include different types of information, useful for nowcasting GDP during different time periods, depending, for example, on the origin of the downturn or upturn. Therefore, it could be useful to combine the nowcasts produced by the MIDAS model based on industrial production growth and the MIDAS models based on financial data. This nowcast results in the lowest RMSEs when using at least 5 lags, while the improvement compared to the nowcast based on only industrial production is weakly statistically significant when using two lags. This is in line with the results in Andreou et al. (2013), who find that combining forecasts from models including macroeconomic and financial data tend to improve short term forecasts over models using only macroeconomic data. Ferrara et al. (2014) find that including financial market volatility and macroeconomic data in a MIDAS model leads to lower RMSEs for nowcasts of GDP growth in France, the UK and the US.

Finally, in Figure 2 we compare the MIDAS model nowcasts to realised GDP growth, in order to assess whether there are differences in model performance over time. It is clear that none of the models capture the sharp decline in GDP in 2012 or the recent slowdown in growth in 2018 very well. Especially in the second quarter of 2012 all the models produce very large nowcast errors. The model driven by industrial production growth performed particularly badly while financial variables were relatively more useful for nowcasting GDP. However, especially due to the limited number of observations available for evaluation, we also consider the RMSE excluding the forecast error for Q2/2012 in Appendix F. The results confirm that in particular the RMSE of the industrial production driven model improves by the exclusion of Q2/2012, but the average nowcast combining financial ratios and industrial production growth still outperforms this nowcast in a weakly statistically significant way when two lags are used. Industrial production growth nowcasts GDP growth particularly well around 2015-2016. It also seems more volatile than the nowcasts based on financial variables. As expected, especially the two combination nowcasts are relatively stable, and thus fail to account for the strong variation in quarterly GDP growth.

Figure 2. *Out-of-sample nowcasts of the different models (number of lags 11) and realised GDP growth. A rolling window, in which the first estimation sample is from Q2/2002 to Q4/2011 (GDP data from Q1/2003), is used to produce the nowcasts.*



Overall, financial market variables and more traditionally used predictors, here represented by industrial production, seem to nowcast Finnish GDP approximately equally well. Therefore, the usefulness of financial variables for nowcasting Finnish GDP growth arises from the fact that they are available earlier than macroeconomic variables. In fact, many of them are available almost immediately and can even be updated on a daily frequency. This issue is highlighted by the fact that the analysis above is conducted using the final, revised vintage of industrial production data, which could give an advantage to industrial production as a predictor of GDP. According to Kuosmanen and Vataja (2014) financial market variables are especially useful during turbulent times. Combined with the fact that when uncertainty is high or the economy is plunging into a recession, timely nowcasts are extremely valuable for economic agents, this highlights the benefits of our MIDAS models based on financial market data.

6. Conclusion

We conclude that financial market variables are useful for nowcasting Finnish quarterly GDP growth. The main advantage of financial market data is its immediate availability. Unlike many other variables that have been traditionally used for forecasting or nowcasting GDP growth, nowcasts based on financial market data may be updated daily. This is particularly useful at the time of crisis, when financial markets might react immediately, but many macroeconomic variables are published with a delay. Using financial market variables in combination with macroeconomic data when nowcasting Finnish GDP growth thus seems beneficial.

On the other hand, our results show that one cannot improve nowcast accuracy for Finnish GDP growth by increasing the sampling frequency of financial market data from quarterly to monthly or daily frequency. This indicates that high-frequency fluctuations of financial markets do not contain additional useful information for nowcasting GDP, which is not already included in corresponding lower frequency data. On the other hand, using higher frequency data does not overall lead to a deterioration in forecast accuracy either, allowing the choice of frequency to be determined by the needs of the researcher.

This paper focuses on the results for Finland. One may argue that this could limit the generalizability of our results. Ultimately, finding the best nowcasting model or the optimal sampling frequency is an empirical question, which depends, among other things, on the sample period used, the explanatory variables chosen as well as the structure of the financial markets and the economy in general. As the previous literature shows, the benefits of individual financial market variables for nowcasting GDP growth varies across countries. For example, Ferrara and Marsilli (2013) shows that the forecasting accuracy of financial market variables differs across four large euro area countries, implying that results are not easily generalisable across countries. However, as discussed earlier in the paper, some of our findings lend support to earlier results, conducted using data mostly on large economies and smaller or more aggregated data sets. Thus, our results broaden the applicability of these results. We also confirm that foreign financial market variables can be important for nowcasting GDP growth in a small open economy using a MIDAS framework. If the results vary between countries as much as the still relatively limited previous literature suggest, it would be interesting to consider the theoretical underpinnings of these results. There could, for example, be some structural differences in the economies driving the variations in the results. This interesting question is left for future research to explore.

Regarding the optimal sampling frequency, which we consider the most interesting research question of our paper, one could argue that the results should not vary significantly across countries where the structure of the economy and the financial markets are similar to Finland. However, factors which could impact the choice of sampling frequency include, for example, persistent differences in the noisiness and the information content of financial market data. One limitation of our analysis, complicating also the comparison to previous papers on other countries, is that these factors could also vary over time. For example, using a higher sampling frequency could improve (by providing more information) or weaken (by increasing noise) forecasting performance during turbulent periods. There is some evidence that increasing the sampling frequency improves forecast accuracy for output growth, but to our knowledge there does not yet exist broad-based evidence of this on an international scale. The relatively short time series of daily data available in many countries complicate the analysis. Overall, we believe determining whether the optimal sampling frequency varies between countries and over time systematically is a highly interesting avenue for future research. If variation of this kind exists, determining the reasons behind this variation is also important.

One issue that may impact the generalizability of our results is the chosen sample period. The results by, for example, Junttila and Korhonen (2011) suggest that financial market variables are especially relevant for forecasting macroeconomic developments during crisis periods. Our in-sample period includes the financial crisis, but the out-of-sample period is less turbulent. Therefore, it is possible that our results understate the nowcasting accuracy of financial market variables, compared to studies where the financial crisis or other turbulent time periods are included.

Although we have used a relatively broad set of variables which we believe to represent the Finnish financial markets well, it is possible that an even more comprehensive set of variables would reveal different results. Another interesting extension left for future research is the use of machine learning methods, which would allow including a larger number of explanatory variables in the same model or enable more flexible weight estimation in the U-MIDAS model.

References

Alessi, L., Ghysels, E., Onorante, L., Peach, R., & Potter, S. (2014). "Central bank macroeconomic forecasting during the global financial crisis: The European Central Bank and Federal Reserve Bank of New York experiences." *Journal of Business & Economic Statistics*, 32(4), 483–500.

Andreou, E., Ghysels, E., & Kourtellos, A. (2013). "Should macroeconomic forecasters use daily financial data and how?" *Journal of Business & Economic Statistic*s, 31(2), 240–251.

Armesto, M. T., Engemann, K. M., & Owyang, M. T. (2010). "Forecasting with mixed frequencies." *Federal Reserve Bank of St. Louis Review*, 92(6), 521–36.

Babii, A., Ghysels E., and Striaukas, J., (2021). "Machine Learning Time Series Regressions with an Application to Nowcasting." Unpublished manuscript. Available at SSRN 3503191

Becker, R., Lee, J., & Gup, B. E. (2012). "An empirical analysis of mean reversion of the S&P 500's P/E ratios." *Journal of Economics and Finance*, 36(3), 675–690.

Bernanke, B. S., Boivin, J., & Eliasz, P. (2005). "Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach." *The Quarterly Journal of Economics*, 120(1), 387–422.

Chionis, D., Gogas, P., & Pragidis, I. (2010). "Predicting European Union recessions in the euro era: The yield curve as a forecasting tool of economic activity." *International Advances in Economic Research*, 16(1), 1–10.

Clements, M. P., & Galvão, A. B. (2008). "Macroeconomic forecasting with mixed-frequency data: Forecasting output growth in the United States." *Journal of Business & Economic Statistics*, 26(4), 546–554.

Clements, M. P., & Galvão, A. B. (2009). "Forecasting US output growth using leading indicators: An appraisal using MIDAS models." *Journal of Applied Econometrics*, 24(7), 1187–1206.

Doğan, B. Ş., & Midiliç, M. (2019). "Forecasting Turkish real GDP growth in a data-rich environment." *Empirical Economics*, 56(1), 367–395.

Estrella, A., & Mishkin, F. S. (1998). "Predicting US recessions: Financial variables as leading indicators." *Review of Economics and Statistics*, 80(1), 45–61.

Ferrara, L., & Marsilli, C. (2013). "Financial variables as leading indicators of GDP growth: Evidence from a MIDAS approach during the Great Recession." *Applied Economics Letters*, 20(3), 233–237.

Ferrara, L., Marsilli, C., & Ortega, J. P. (2014). "Forecasting growth during the Great Recession: is financial volatility the missing ingredient?" *Economic Modelling*, 36, 44–50.

Foroni, C., & Marcellino, M. G. (2013). "A survey of econometric methods for mixed-frequency data." *Norges Bank Working Paper* 06/2013. Available at SSRN 2268912.

Foroni, C., Marcellino, M., & Schumacher, C. (2015). "Unrestricted mixed data sampling (MIDAS): MIDAS regressions with unrestricted lag polynomials." *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(1), 57–82.

Franta, M., Havrlant, D., & Rusnák, M. (2016). "Forecasting Czech GDP using mixed-frequency data models." *Journal of Business Cycle Research*, 12(2), 165–185.

Friedman, B. M., & Kuttner, K. (1993). "Why does the paper-bill spread predict real economic activity?" In Business cycles, indicators and forecasting (pp. 213–254). University of Chicago Press.

Galli, A., Hepenstrick, C., & Scheufele, R. (2019). "Mixed-frequency models for tracking short-term economic developments in Switzerland." *International Journal of Central Banking*, 58.

Ghysels, E., Kvedaras, V., & Zemlys, V. (2016). "Mixed Frequency Data Sampling Regression Models: The R Package midasr." *Journal of Statistical Software*, 72(4), 1–35.

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2004). "The MIDAS touch: Mixed data sampling regression models." Discussion paper, UCLA.

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2005). "There is a risk-return trade-off after all." *Journal of Financial Economics*, (76), 509-548.

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2006). "Predicting volatility: getting the most out of return data sampled at different frequencies." *Journal of Econometrics*, 131(1–2), 59–95.

Ghysels, E., Sinko, A., & Valkanov, R. (2007). "MIDAS regressions: Further results and new directions." *Econometric Reviews*, 26(1), 53–90.

Gomez-Zamudio, L. M., & Ibarra, R. (2017). "Are daily financial data useful for forecasting GDP? Evidence from Mexico." *Journal of the Latin American and Caribbean Economic Association*, 17(2), 173–203.

Henry, O. T., Olekalns, N., & Thong, J. (2004). "Do stock market returns predict changes to output? Evidence from a nonlinear panel data model." *Empirical Economics*, 29(3), 527–540.

Junttila, J., & Korhonen, M. (2011). "Utilizing financial market information in forecasting real growth, inflation and real exchange rate." *International Review of Economics & Finance*, 20(2), 281–301.

Kim, H. H., & Swanson, N. R. (2018). "Methods for backcasting, nowcasting and forecasting using factor-MIDAS: With an application to Korean GDP." *Journal of Forecasting*, 37(3), 281–302.

Koustas, Z., & Serletis, A. (2005). "Rational bubbles or persistent deviations from market fundamentals?" *Journal of Banking & Finance*, 29(10), 2523–2539.

Kuosmanen, P., & Vataja, J. (2014). "Forecasting GDP growth with financial market data in Finland: Revisiting stylized facts in a small open economy during the financial crisis." *Review of Financial Economics*, 23(2) 90–97.

Lewellen, J. (2004). "Predicting returns with financial ratios." *Journal of Financial Economics*, 74(2), 209–235.

Lütkepohl, H. (2010). "Forecasting aggregated time series variables: A survey." *OECD Journal: Journal of Business Cycle Measurement and Analysis*, 2010(2), 1–26.

Marcellino, M., & Schumacher, C. (2010). "Factor MIDAS for nowcasting and forecasting with ragged-edge data: A model comparison for German GDP." *Oxford Bulletin of Economics and Statistics* 72(4), 518–550.

Marsilli, C. (2014). "Variable selection in predictive MIDAS models." *Banque de France Working Paper*, 520, Available at SSRN: https://ssrn.com/abstract=2531339.

Nyberg, H. (2010). "Dynamic probit models and financial variables in recession forecasting." *Journal of Forecasting*, 29(1-2), 215–230.

Tay, A. S. (2007). "Financial variables as predictors of real output growth." Singapore Management University Economics and Statistics Working Paper No. 14-2007.

Tsui, A. K., Xu, C. Y., & Zhang, Z. (2018). "Macroeconomic forecasting with mixed data sampling frequencies: Evidence from a small open economy." *Journal of Forecasting*, 37(6), 666–675.

Appendix

A

The following table lists all the used variables, their abbreviations, sources and transformations. All the data are obtained via Bloomberg.

Table A 1

Variable	Category	Abbreviation	Source	Transformation
Average ratio of the price of a stock and the company's earnings per share in OMX Helsinki	Financial ratio	OMX Hels p/e	NASDAQ OMX Helsinki	No transformation
Average ratio of the stock price to the book value per share in OMX Helsinki.	Financial ratio	OMX Hels p/b	NASDAQ OMX Helsinki	No transformation
Average dividend yield in OMX Helsinki	Financial ratio	OMX Hels Dividend yield	NASDAQ OMX Helsinki	No transformation
Average ratio of the stock price to the book value per share in DAX	Financial ratio	DAX p/e	Deutsche Börse	No transformation
Average ratio of the stock price to the book value per share in DAX.	Financial ratio	DAX p/b	Deutsche Börse	No transformation
Average dividend yield in DAX	Financial ratio	DAX Dividend yield	Deutsche Börse	No transformation
Average ratio of the stock price to the book value per share in S&P 500	Financial ratio	SP500 p/e	Standard and Poor's	No transformation
Average ratio of the stock price to the book value per share in S&P 500	Financial ratio	SPX p/b	Standard and Poor's	No transformation
Average dividend yield in S&P 500	Financial ratio	SP500 Dividend yield	Standard and Poor's	No transformation
The yield of Finland government bond with maturity of 10 years	Interest rate	FI_10y	Bloomberg	Difference
The yield of Finland government bond with maturity of 5 years	Interest rate	FI_5y	Bloomberg	Difference
The yield of Germany government bond with maturity of 12 months	Interest rate	DE_1y	Bloomberg	Difference
The yield of Germany government bond with maturity of 5 years	Interest rate	DE_5y	Bloomberg	Difference
The yield of Germany government bond with maturity of 7 years	Interest rate	DE_7y	Bloomberg	Difference
The yield of Germany government bond with maturity of 10 years	Interest rate	DE_10y	Bloomberg	Difference
The spread between German 10 year yield and 12 month yield	Interest rate	DE 10y-1y	Bloomberg	No transformation

Variable	**Category**	**Abbreviation**	**Source**	**Transformation**
Implied volatility on Eurostoxx 50 index options with a rolling 30 day expiry	Other	Eurostoxx 50 volatility	Deutsche Börse	No transformation
The price of oil (brent)	Other	OIL	Deutsche Börse	Log-difference
The price of euro in dollars	Other	EURUSD	Deutsche Börse	Difference
Finland Industrial Production Volume, MoM growth rate, SA	Real economy	Industrial production	Statistics Finland	No transformation
OMX Helsinki, price index	Stock index	OMX Hels	NASDAQ OMX Helsinki	Log-difference
OMX Helsinki Technology, price index	Stock index	OMX Hels Technology	NASDAQ OMX Helsinki	Log-difference
OMX Helsinki Utilities, price index	Stock index	OMX Hels Utilities	NASDAQ OMX Helsinki	Log-difference
OMX Hels Industrials, price index	Stock index	OMX Hels Industrials	NASDAQ OMX Helsinki	Log-difference
OMX Hels Telecommunication, price index	Stock index	OMX Hels Telec	NASDAQ OMX Helsinki	Log-difference
OMX Helsinki Basic Materials, preice index	Stock index	OMX Hels Basic Matl	NASDAQ OMX Helsinki	Log-difference
OMX Helsinki Health Care, price index	Stock index	OMX Hels Hlth Care	NASDAQ OMX Helsinki	Log-difference
S&P 500, price index	Stock index	SP500	Standard and Poor’s	Log-difference
DAX, price index	Stock index	DAX	Deutsche Börse	Log-difference

The following figures plot all the (transformed) variables (solid lines) aggregated to a quarterly frequency together with quarterly GDP growth (dashed lines). All variables have been demeaned and divided by their standard deviations (in the figures) to make interpretation easier.

Figure A 1

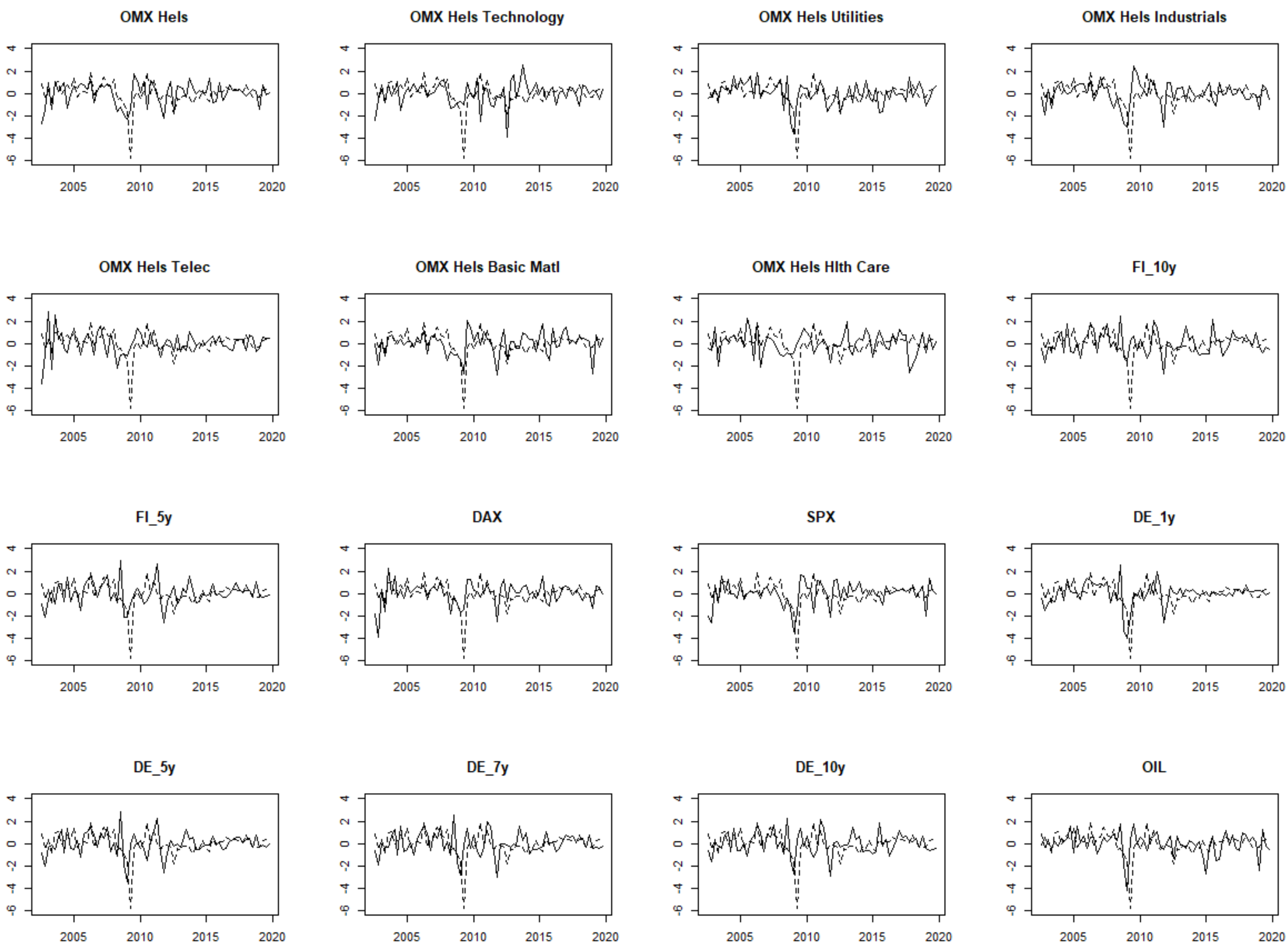
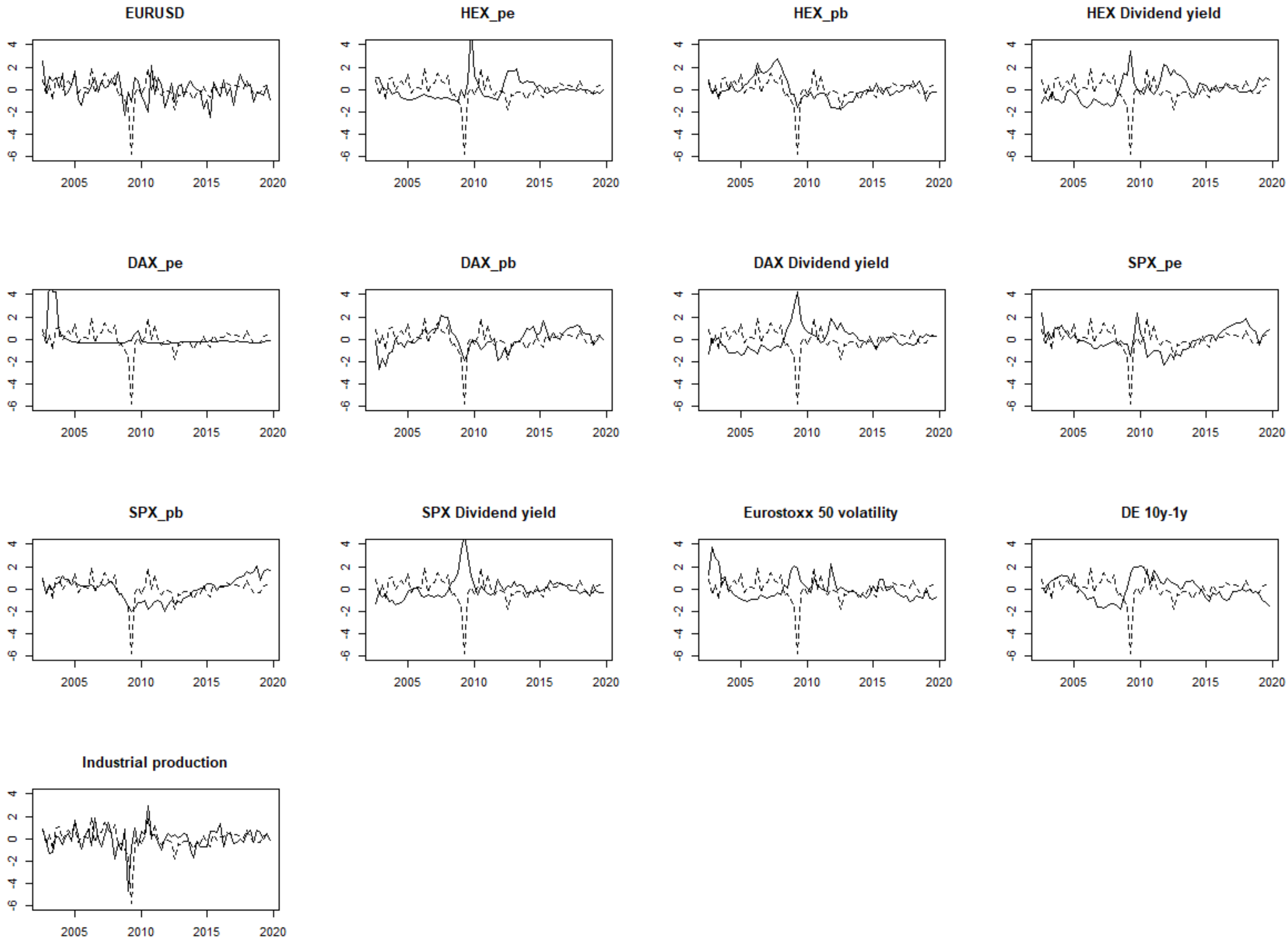


Figure A 2


EURUSD
HEX_pe
HEX_pb
HEX Dividend yield
DAX_pe
DAX_pb
DAX Dividend yield
SPX_pe
SPX_pb
SPX Dividend yield
Eurostoxx 50 volatility
DE 10y-1y
Industrial production
4 2 0 -2 -4 -6
2005 2010 2015 2020

B

The following figures show the weighting schemes that are estimated using daily data from Q2/2002 to Q3/2019.

Figure B 1

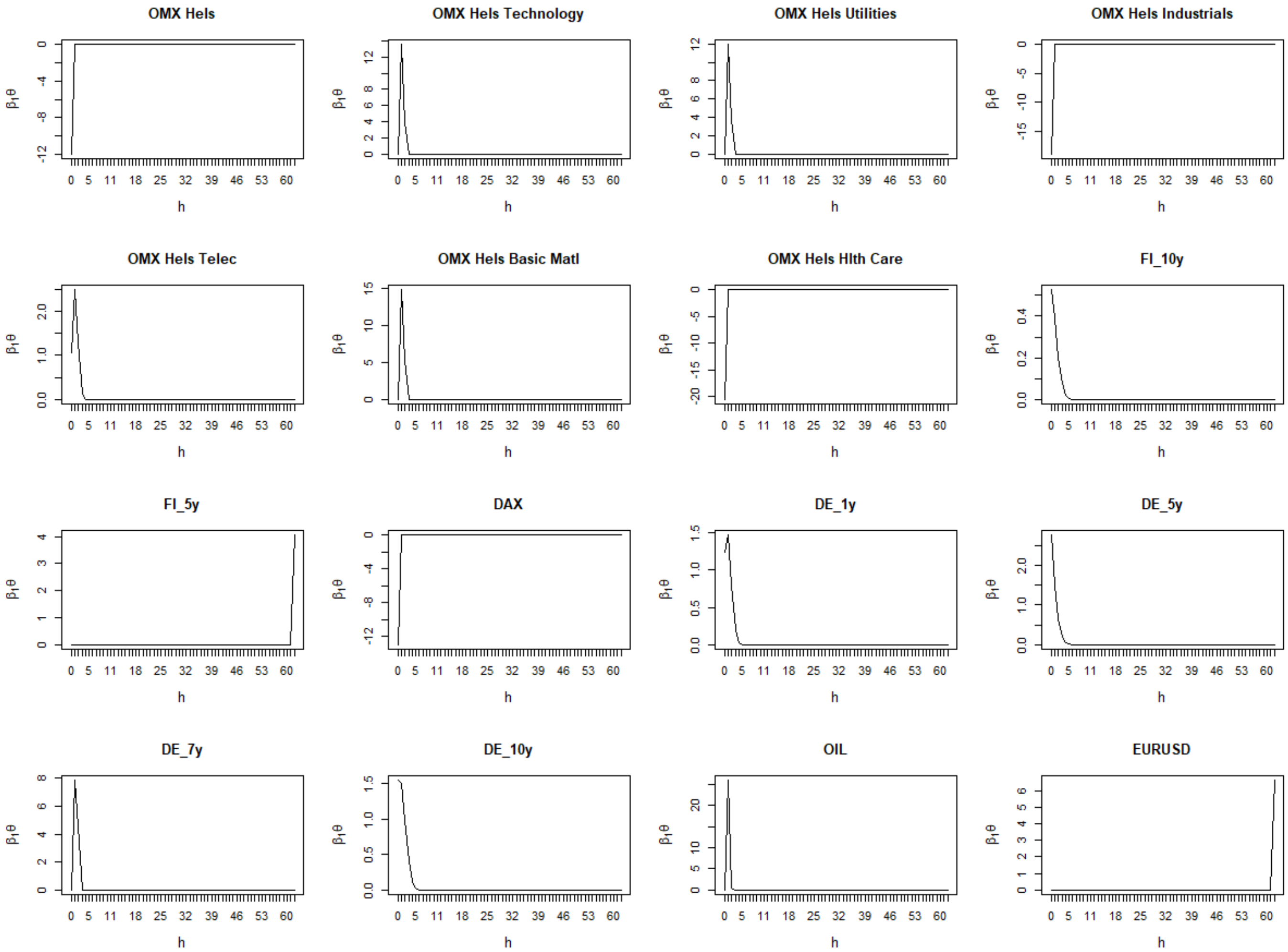
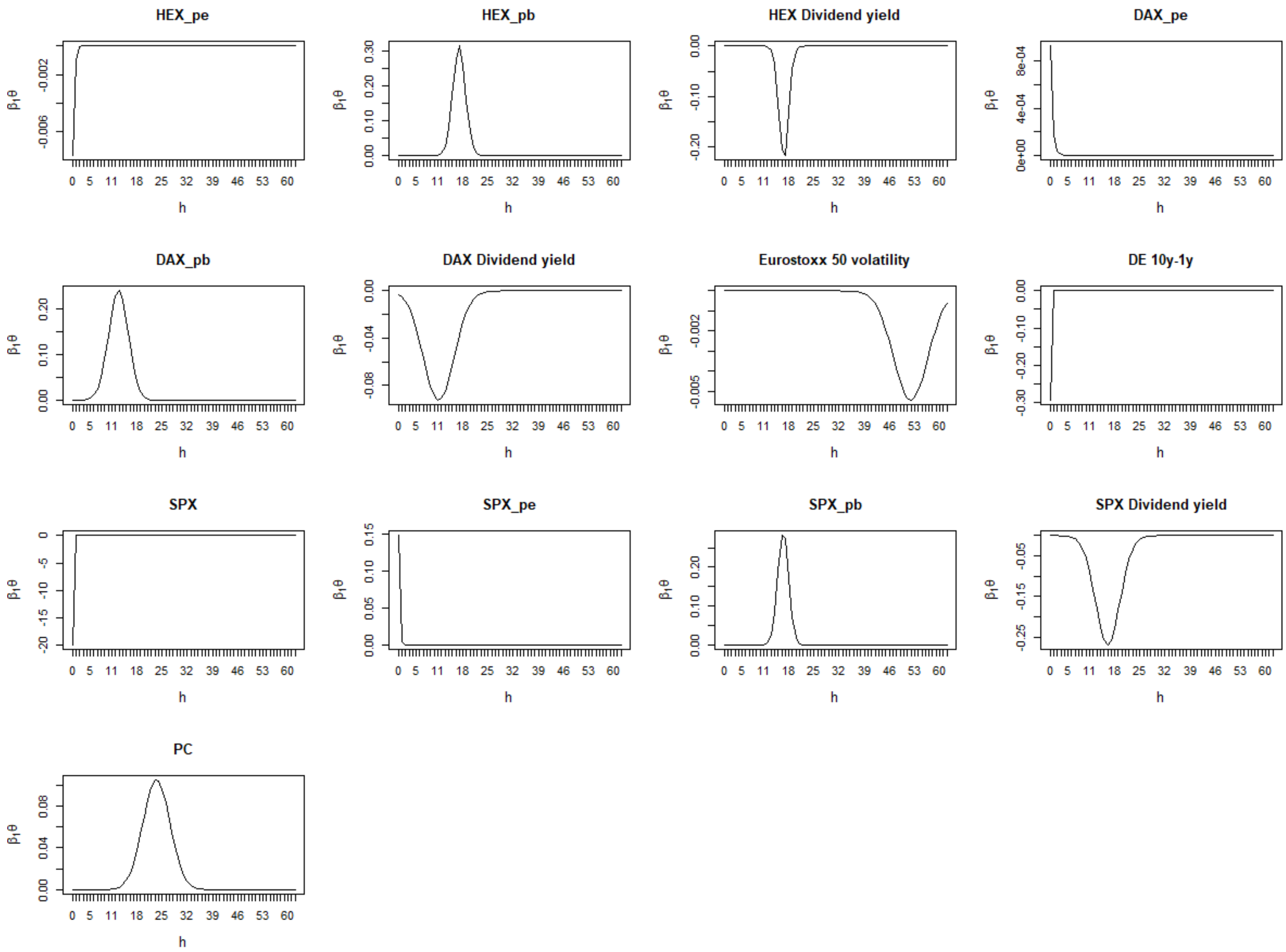


Figure B 2


HEX_pe
HEX_pb
HEX Dividend yield
DAX_pe
DAX_pb
DAX Dividend yield
Eurostoxx 50 volatility
DE 10y-1y
SPX
SPX_pe
SPX_pb
SPX Dividend yield
PC
β₁θ
h
0 5 11 18 25 32 39 46 53 60

The following figures show the weighting schemes that are estimated using monthly data from Q2/2002 to Q3/2019.

Figure B 3

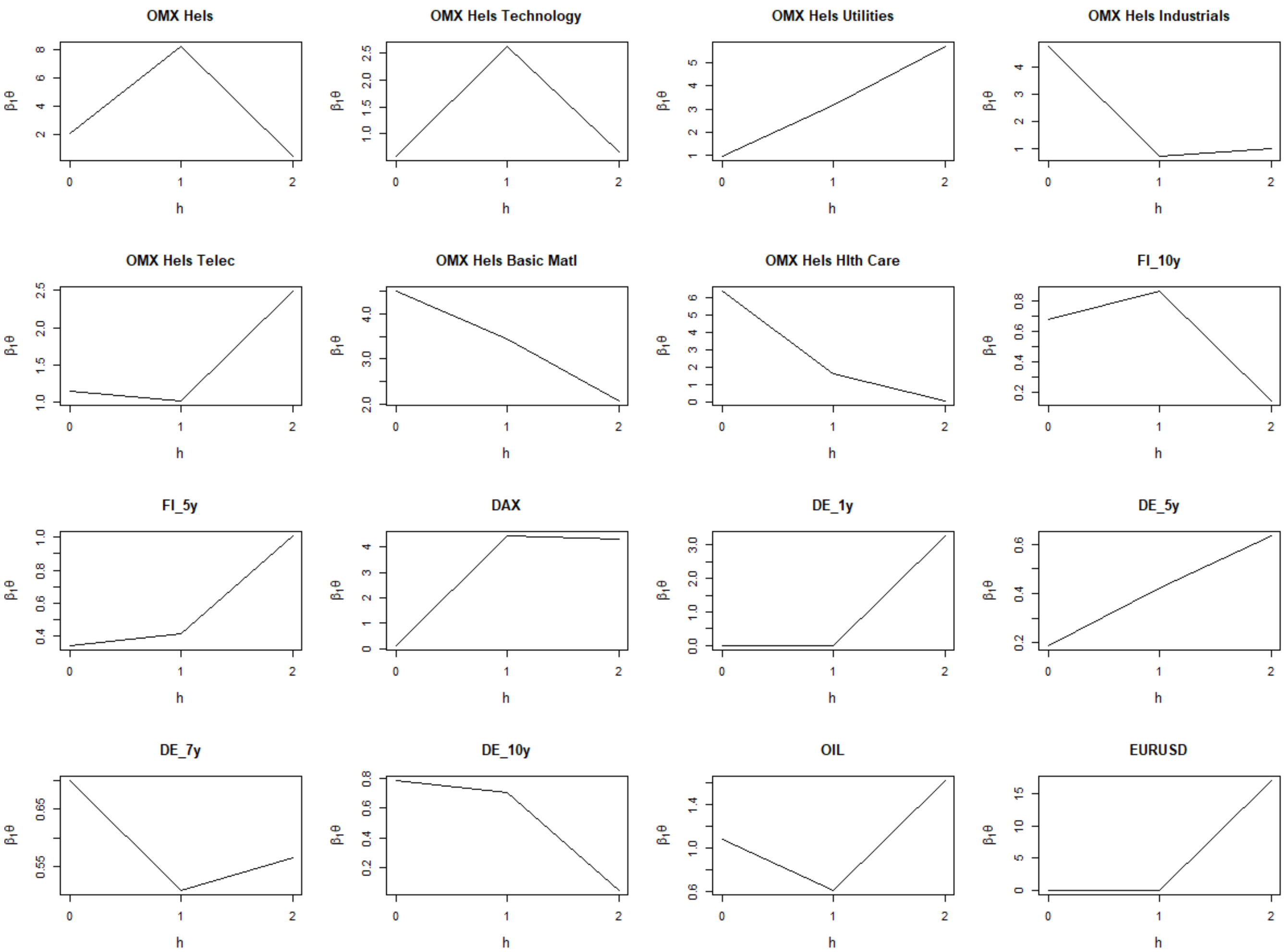
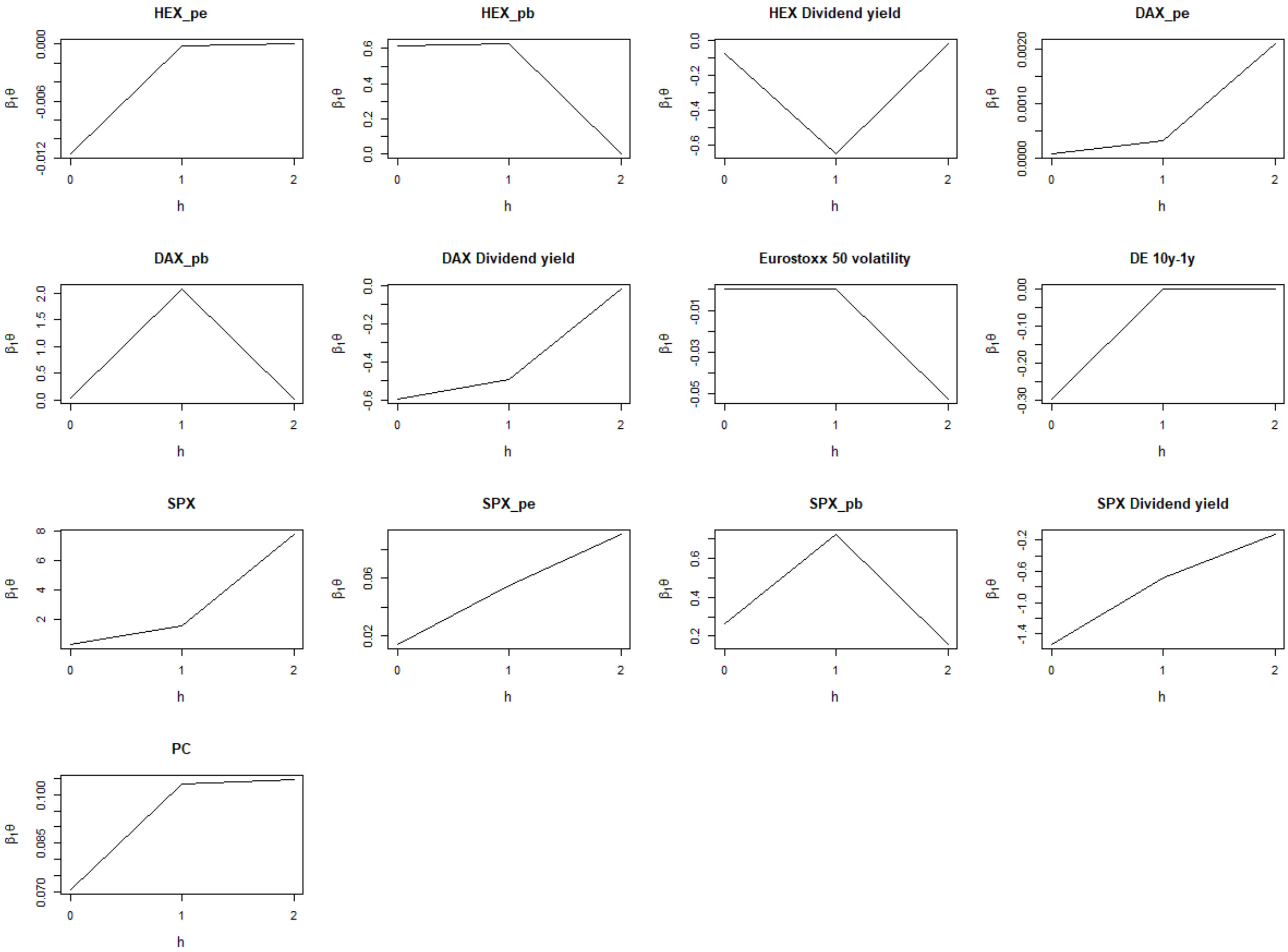


Figure B 4


HEX_pe
HEX_pb
HEX Dividend yield
DAX_pe
DAX_pb
DAX Dividend yield
Eurostoxx 50 volatility
DE 10y-1y
SPX
SPX_pe
SPX_pb
SPX Dividend yield
PC
h

The following figures show the coefficient estimates of the unrestricted model that is estimated using monthly data from Q2/2002 to Q3/2019.

Figure B 5

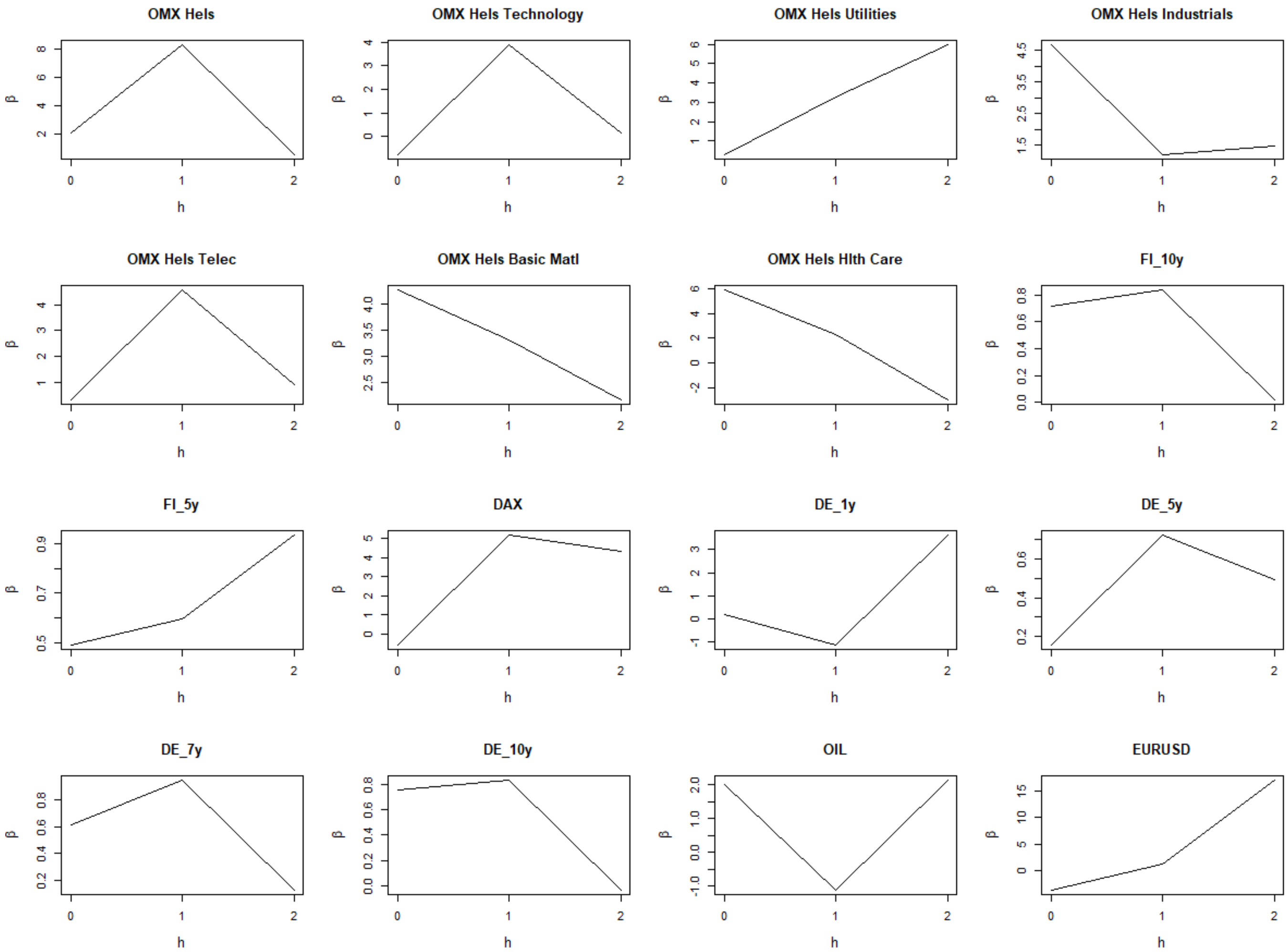


Figure B 6

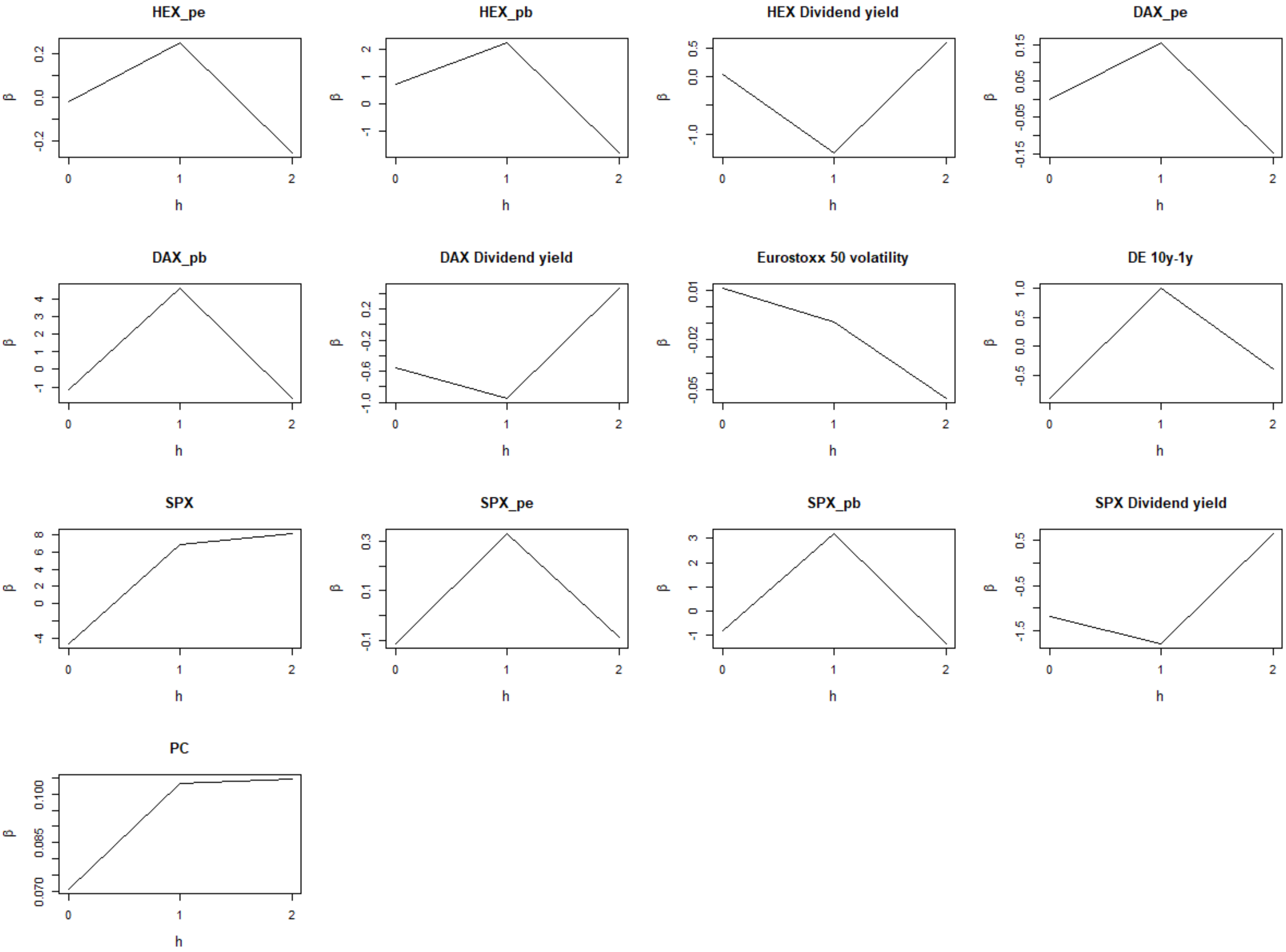
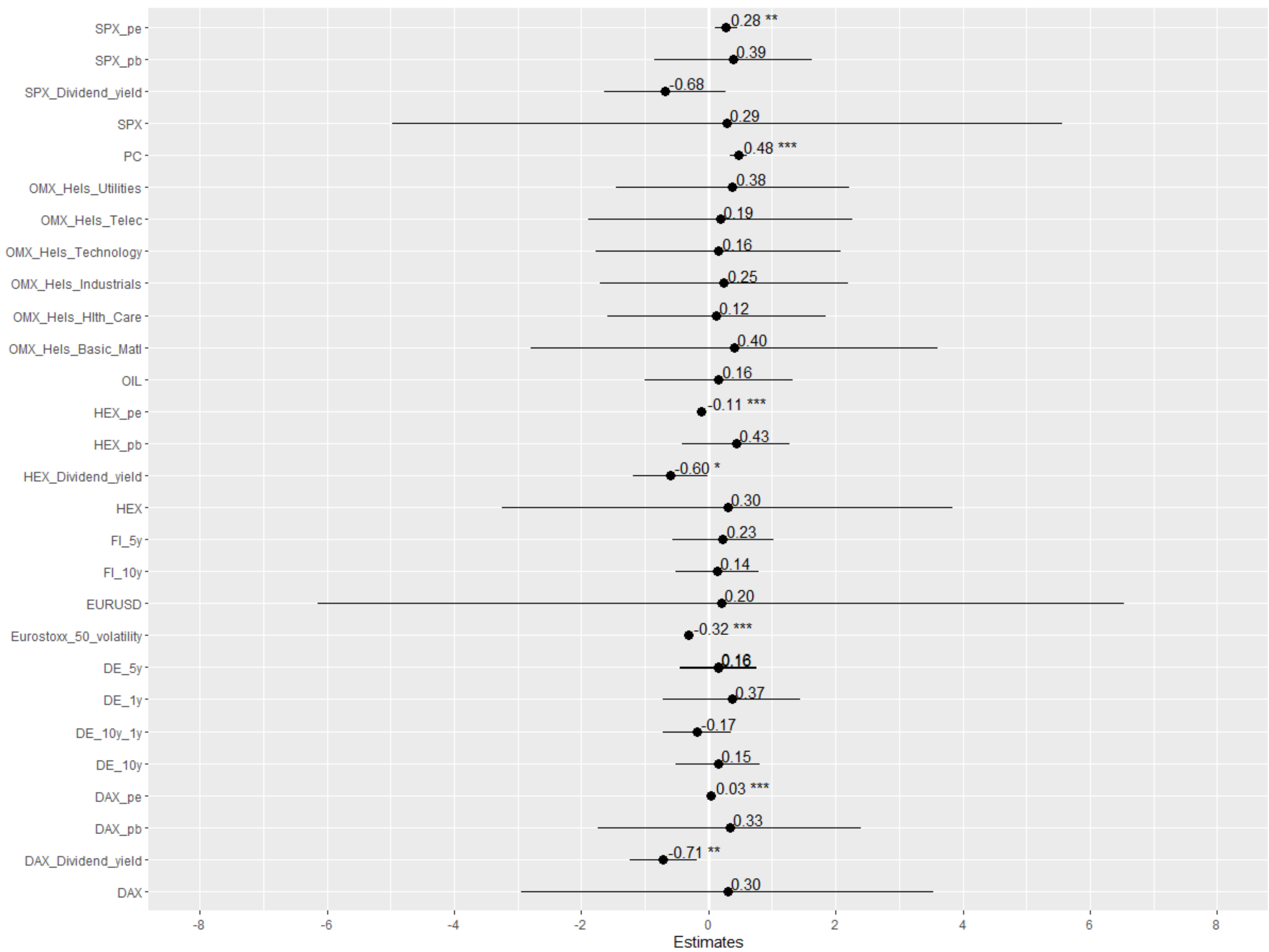

HEX_pe
HEX_pb
HEX Dividend yield
DAX_pe
DAX_pb
DAX Dividend yield
Eurostoxx 50 volatility
DE 10y-1y
SPX
SPX_pe
SPX_pb
SPX Dividend yield
PC

Figure B 7 reports standardised coefficient estimates for the quarterly models. The lines represent 95 per cent confidence intervals based on heteroscedasticity and autocorrelation robust standard errors.

Figure B 7



C

The following figures show the weighting schemes that are estimated using monthly data from Q3/2002 to Q3/2019 assuming 5 lags.

Figure C 1

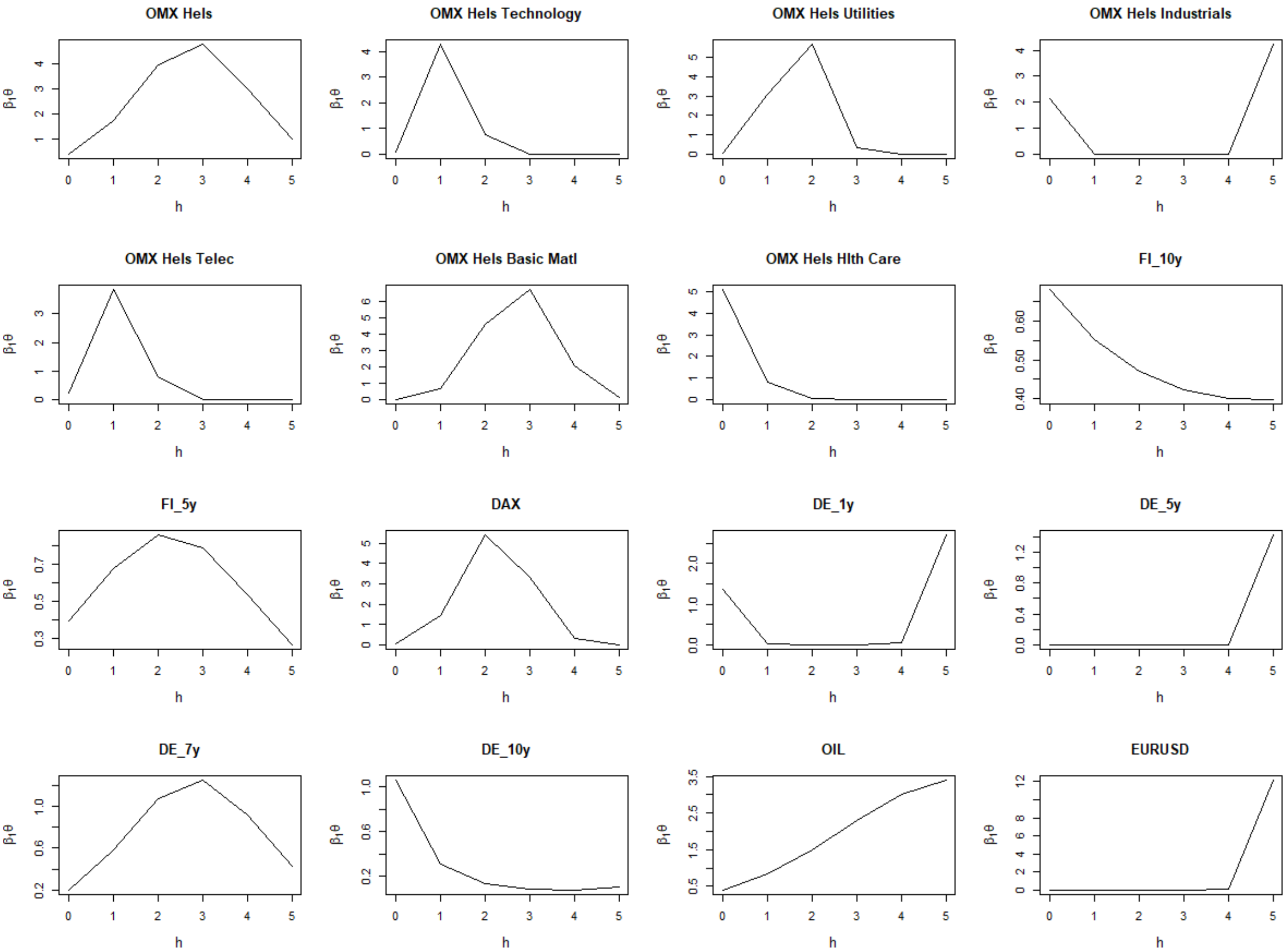
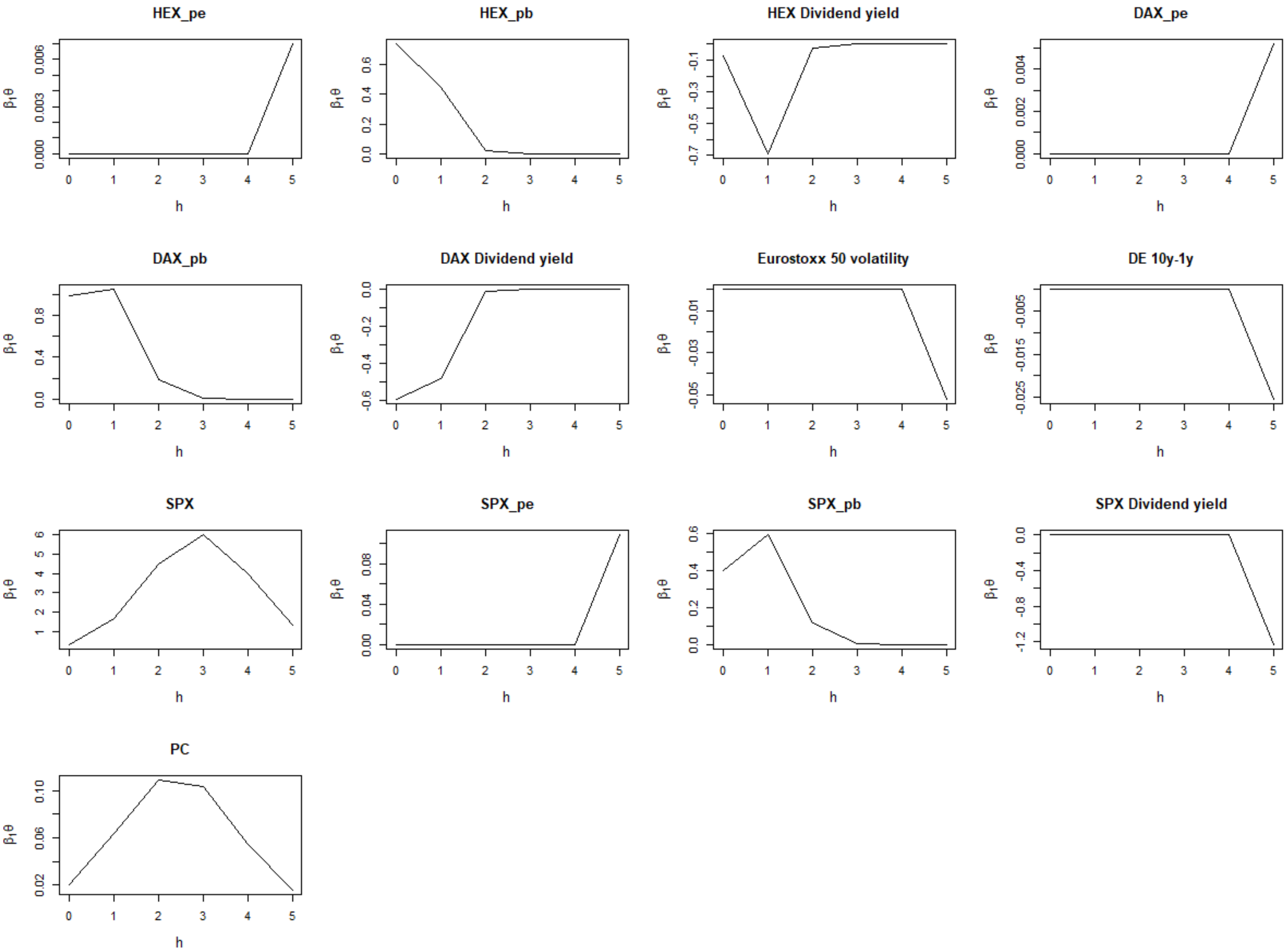


Figure C 2


HEX_pe
HEX_pb
HEX Dividend yield
DAX_pe
DAX_pb
DAX Dividend yield
Eurostoxx 50 volatility
DE 10y-1y
SPX
SPX_pe
SPX_pb
SPX Dividend yield
PC
h
$\beta_1\theta$

D

The following figures show the weighting schemes that are estimated using monthly data from Q1/2003 to Q3/2019 assuming 11 lags.

Figure D 1

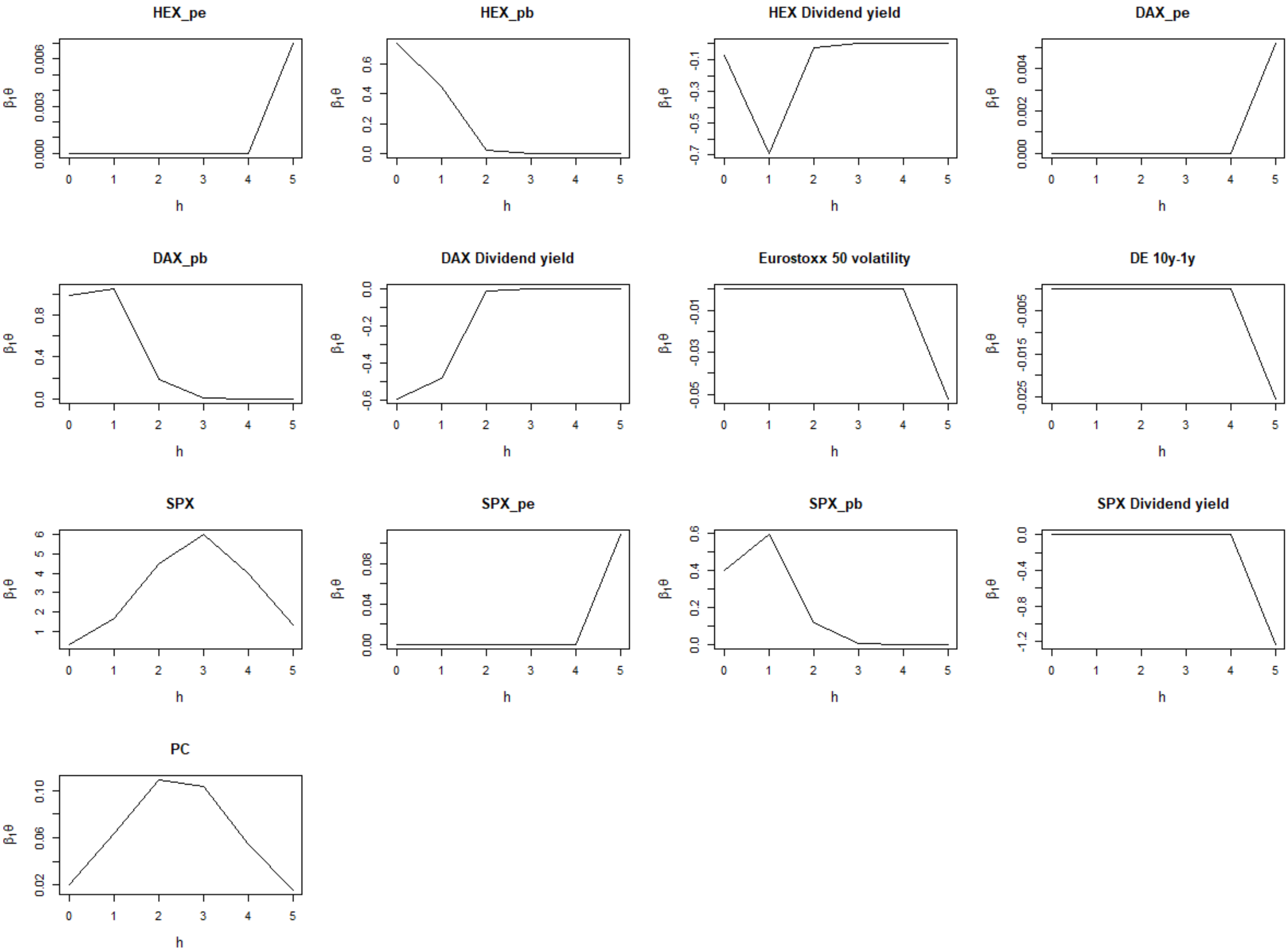
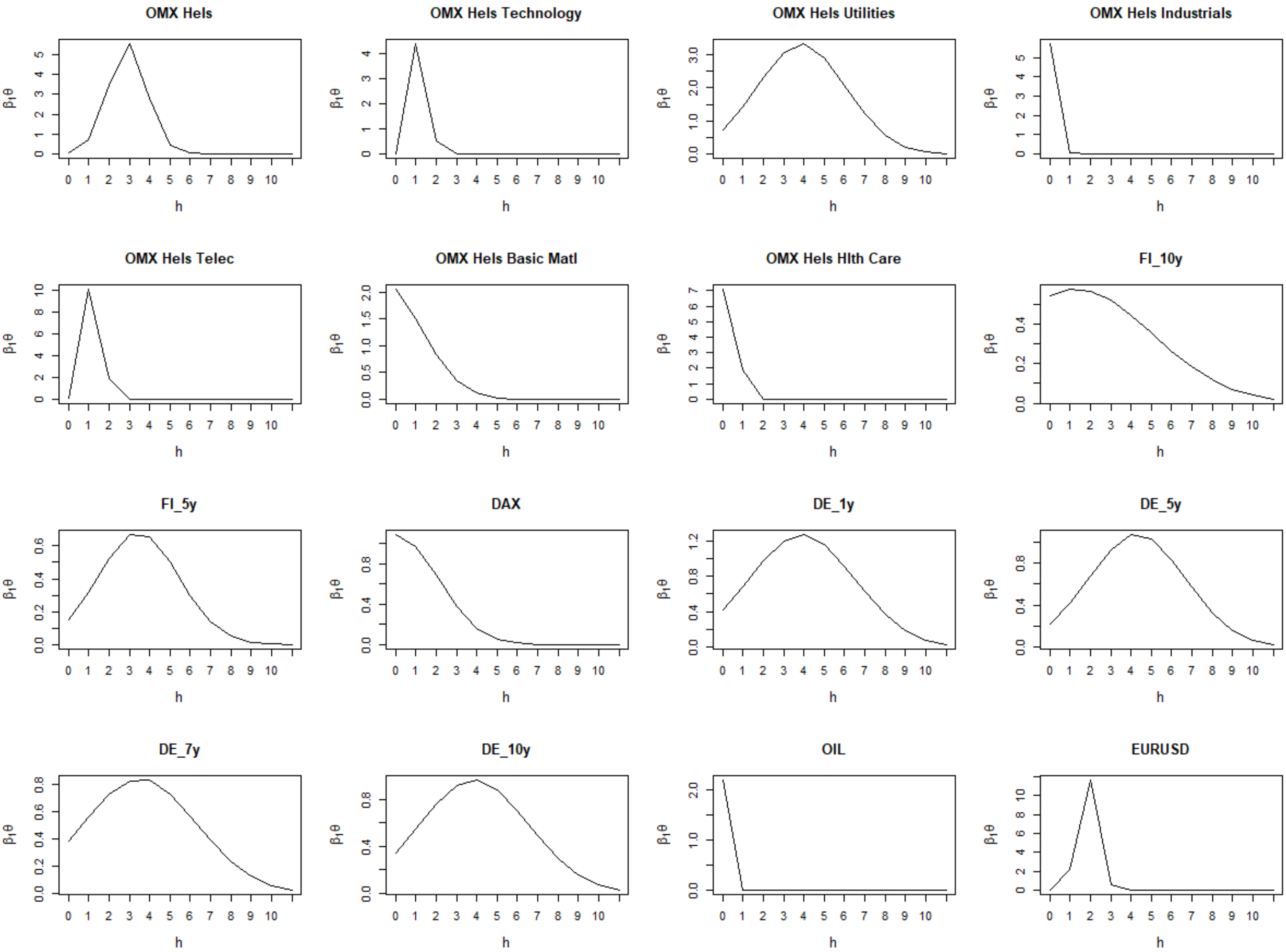


Figure D 2


OMX Hels
OMX Hels Technology
OMX Hels Utilities
OMX Hels Industrials
OMX Hels Telec
OMX Hels Basic Matl
OMX Hels Hlth Care
FI_10y
FI_5y
DAX
DE_1y
DE_5y
DE_7y
DE_10y
OIL
EURUSD
h
$\beta_1\theta$

E

Table E 1 replicates Table 1 excluding the out-of-sample forecast for the period Q2/2012. The forecast error in Q2/2012 is clearly larger than at other times and is excluded here in order to control for the effect of this potential outlier.

Table E 1: *Out-of-sample RMSEs of MIDAS regression models. The models are ordered based on the RMSEs of the quarterly models. The abbreviations for the variables are explained in Appendix A. The RMSEs are calculated from a rolling window analysis, in which the first estimation sample is from Q2/2002 to Q4/2011 and the first out-of-sample forecast is Q1/2012. The out-of-sample forecast for the period Q2/2012 is excluded. Thus, the results are based on 30 out-of-sample observations.*

Explanatory variable	Quarterly	Monthly	Monthly (unrestricted)	Daily
SP500 p/e	0.42	0.47	0.49	0.45
OMX Hels p/e	0.53	0.56	0.64	0.55
DE 10y-1y	0.54	0.54	0.57	0.54
Industrial production	0.54	0.46	0.50	–
DE_10y	0.55	0.59	0.77	0.54
FI_10y	0.55	0.69	0.83	0.54
FI_5y	0.55	0.57	0.59	0.53
DAX Dividend yield	0.56	0.57	0.60	0.57
DAX p/e	0.56	0.56	0.78	0.56
DE_5y	0.56	0.60	0.60	0.55
DE_7y	0.56	0.62	0.71	0.55
OIL	0.56	0.56	0.56	0.62
HEX p/b	0.57	0.61	0.72	0.63
DE_1y	0.58	0.54	0.63	0.52
EURUSD	0.58	0.69	0.69	0.54
PC	0.59	0.58	0.58	0.65
SP500 p/b	0.59	0.63	0.74	0.63
OMX Hels Dividend yield	0.60	0.66	0.65	0.61
OMX Hels Telec	0.61	0.68	0.72	0.59
OMX Hels Industrials	0.62	0.60	0.61	0.58
OMX Hels Hlth Care	0.62	0.72	0.91	0.62
SPX Dividend yield	0.62	0.61	0.66	0.61
SP500	0.66	0.68	0.69	0.57
DAX p/b	0.67	0.62	0.79	0.73
Eurostoxx 50 volatility	0.67	0.65	0.67	0.67
OMX Hels Utilities	0.69	0.64	0.71	0.62
OMX Hels	0.72	0.71	0.80	0.58
OMX Hels Technology	0.72	0.71	0.86	0.61
DAX	0.79	0.75	0.82	0.59
OMX Hels Basic Matl	0.82	0.79	0.84	0.64

Table E 2 is otherwise identical to Table 1, but the number of out-of-sample observations is increased from 31 to 40.

Table E 2: *Out-of-sample RMSEs of MIDAS regression models. The models are ordered based on the RMSEs of the quarterly models. The abbreviations for the variables are explained in Appendix A. The RMSEs are calculated from a rolling window analysis, in which the first estimation sample is from Q2/2002 to Q4/2011 and the first out-of-sample forecast is Q3/2009. Thus, the results are based on 40 out-of-sample observations.*

Variable	Quarterly	Monthly	Monthly (unrestricted)	Daily
Industrial production	0.67	0.75	0.72	0.75
OMX Hels p/e	0.73	0.98	0.92	1.26
DAX p/e	0.75	0.76	0.97	0.76
OIL	0.75	0.75	0.81	0.75
DE_10y	0.76	0.95	1.06	0.74
FI_10y	0.76	0.99	1.14	0.85
HX6000PI	0.77	0.85	0.93	0.73
HX7000PI	0.77	0.81	0.85	0.74
DE 10y-1y	0.77	0.77	0.81	0.83
DE_5y	0.78	0.86	0.89	0.82
DE_7y	0.78	0.92	1.00	0.87
HX2000PI	0.78	0.78	0.83	0.76
HX4000PI	0.80	0.84	1.01	0.75
SP500 Dividend yield	0.82	0.80	0.90	0.81
SP500 p/e	0.82	0.74	1.01	0.88
FI_5y	0.83	0.86	0.91	0.90
DE_1y	0.84	0.83	0.90	0.76
HEX p/b	0.84	0.85	0.99	0.84
DAX Dividend yield	0.85	0.85	0.92	0.85
DAX p/b	0.86	0.80	1.03	0.88
EURUSD	0.89	0.97	1.04	0.74
OMX Hels	0.90	1.02	1.03	0.72
PC1	0.93	1.02	1.00	0.83
SP500	0.94	0.93	1.03	0.73
HX1000PI	0.95	0.94	1.07	0.77
OMX Hels Dividend yield	0.97	0.92	0.91	0.91
DAX	0.99	0.90	1.16	0.74
HX9000PI	0.99	0.96	1.22	0.83
Eurostoxx 50 volatility	1.00	0.83	1.02	1.06
SP500 p/b	1.07	1.09	1.19	1.07

F

Table F 1 replicates Table 2 excluding the out-of-sample forecast for the period Q2/2012. The forecast error in Q2/2012 is clearly larger than in other periods and is excluded here in order to control for the effect of this potential outlier. Comparing Table 2 to Table F 1 reveals that especially the RMSE of the industrial production driven model improves by the exclusion of Q2/2012. However, the average forecast using the financial ratios and industrial production still outperforms the industrial production driven forecast in a weakly statistically significant way when including two monthly lags and performs equally well when using more lags. Using only financial ratios for forecasting still produces competitive forecasts as well.

Table F 1: *RMSEs of MIDAS regression models. In the model using industrial production the only explanatory variable is the MoM growth rate of the volume of industrial production. In the PC model, the only explanatory variable is the first principal component of the financial market variables (see Appendix A). 'Average forecast' is the simple average of all the financial variable based forecasts (forecasts produced using the financial market variables listed in Appendix A one at a time). 'Average forecast based on financial ratios only' is the average of the models in which the explanatory variable is the price-to-earnings ratio, the price-to-book ratio or the dividend yield. The RMSEs are calculated from a rolling window analysis, in which the first estimation sample is from Q1/2003 to Q4/2011 and the first out-of-sample forecast is Q1/2012. The out-of-sample forecast for the period Q2/2012 is excluded. Thus, the results are based on 30 out-of-sample observations. To test the statistical significance of the RMSE differences between the forecast based on industrial production and the other forecasts, we use the Diebold-Mariano test assuming no heteroscedasticity or autocorrelation because the forecast horizon is zero. Asterisks *,** and*** indicate statistical significance at the 10%, 5% and 1% levels, respectively.*

	2 lags	**5 lags**	**11 lags**
Industrial production	0.56	0.39	0.42
PC	0.57	0.56	0.54
Average forecast	0.54	0.55*	0.55
Average forecast: financial ratios only	0.50	0.50	0.52
Average forecast: financial ratios only and industrial production	0.50*	0.39	0.42

G

Figure G 1 shows the weighting schemes of industrial production that are estimated using monthly data from Q2/2002 to Q3/2019 with 3, 5 and 11 lags.

Figure G 1