Foreword: Journal of the Finnish Economic Association 1/2020.

We are very happy to provide the very first issue of the Journal of the Finnish Economic Association (JFEA). JFEA is a new peer-reviewed international journal published biannually by the Finnish Economic Association and it replaces the Finnish Economic Papers. The objective of JFEA is to provide a high-quality and fast peer-reviewed publication channel for applied economic papers, and facilitate communication of topical research results to the research community and policymakers.

As the editors of JFEA, we welcome submissions in all fields of economics and closely related sciences with a special emphasis on policy-relevant applications. We are aware that many important and well-executed policy reports and research articles by both faculty and students are never published – perhaps due to lack of statistical significance of the results or perceived lack of novelty. This is often a loss for both academic and policy-making community. We hope to provide a platform that would be open also to such studies. Please keep us in mind for your and your students' work. We particularly encourage studies concerning economies and institutions of Nordic and neighboring countries.

This first issue of JFEA consists of five articles. In the first article, Gilles Saint-Paul argues using a formal model of fiscal policy and empirical evidence, that it is rational for some social groups to support policies that are macroeconomically unsound. This happens if these groups are predictably likely to bear a lower fraction of the costs of the resulting economic crisis, while benefitting from the short-run gains associated with the policies. In the second article, Helena Holmlund summarizes the methodological aspects and substantive findings from the Swedish compulsory school reform. She shows that it provided children from low socio-economic background with better opportunities in life, including higher level of education and earnings as well as reduced likelihood of committing crimes.

John Hassler summarizes his work on the economics of climate change in the third article of the issue. The key conclusion is that a global agreement on a (minimum) price on fossil carbon emission is necessary, sufficient and efficient solution to limiting climate change. In the fourth article, Jacob Lundberg and John Norell survey the quasi-experimental literature on the effects of taxes and benefits on labour force participation. They find strong evidence that individuals respond to incentives on the extensive margin of labour supply. In the fifth article of the first issue, Tuomas Takalo calibrates the switching costs for the Finnish retail deposit market and shows that they manifest large variation across the banks and time, and are high.

Editors of JFEA,
Mika Kortelainen and Janne Tukiainen

From Microeconomic Favoritism to Macroeconomic Populism

Gilles Saint-Paul[1]
Paris School of Economics, ENS-PSL and NYU-AD

Abstract

Why would people support policies that are macroeconomically unsound, in that they are more likely to lead to such events as sovereign crises, balance of payments crises, and the like? This may arise if decisive voters are likely to bear a lower fraction of the costs of the crisis, while benefitting from the short-run gains associated with those policies, such as greater public expenditure or lower taxes.

I first discuss an illustrative model based on Saint-Paul et al. (2017), based on the assumption that in a crisis, not everybody can access his or her entitlement to publicly provided goods, a feature labelled "favoritism". If the decisive voter is relatively favored in this rationing process, then people are more likely to finance public expenditure by debt, the greater the degree of favoritism. Furthermore, favoritism and the likelihood of a crisis raises the level of public spending.

Next, I consider the choice between electing a "populist" who reneges on anonymity when allocating the public good, even in normal times, and a "technocrat" who sticks to anonymity, and does all it takes to balance the budget. I show that the support for the populist is greater, (i) the greater the likelihood of default, (ii) the more depressed the macroeconomic environment, (iii) the greater the inherited level of public debt and (iv) the lower the state's fiscal capacity.

I then argue that the model helps understanding some episodes in French pension reform. Some occupational groups supported unsustainable reductions in the retirement age because they expected that other workers would bear a higher proportion of the burden of future adjustment.

Finally, using a panel of countries, I provide evidence in favor of some of the predictions of the model. As predicted, favoritism raises public debt, budget deficits, and public spending. It also raises the likelihood of a fiscal crisis through its effect on public debt. Furthermore, "populists" are more likely to conquer power, the higher the degree of debt and budget deficits, and the higher the level of government spending – the latter finding being consistent with the model's prediction on the effect of fiscal capacity.

Keywords: political economy, fiscal crises, favoritism, entitlements, public debt, inequality, state capacity
JEL Classification: E620, F340, H120, H600, O110, P160

[1] *I am grateful to Abhijit Banerjee, François Bourguignon, Andrea Ichino, Davide Ticchi, Mathilde Viennot, and Andrea Vindigni, as well as seminar participants at the Bank of Finland, for helpful comments and suggestions. This work has been funded by a French government subsidy managed by the Agence Nationale de la Recherche under the framework of the Investissements d'avenir programme reference ANR-17-EURE-001.*

Editor: Kari Heimonen

1. Introduction

Why would people support policies that are macroeconomically unsound, in that they are more likely to lead to such events as sovereign crises, balance of payments crises, and the like? Dornbusch and Edwards (1991) have noted that such policies, which they label “macroeconomic populism”, are recurrent, in particular in Latin America, and typically end in severe crises and painful internal and external adjustment. Perhaps the popular support for such platforms is a result of irrationality or short-sightedness. But I want to argue that it is rational for some social groups to support those policies. This is because they are predictably likely to bear a lower fraction of the costs of the crisis, while benefitting from the short-run gains associated with the policies, such as greater public expenditure or lower taxes. We can think of a fiscal crisis as implying expenditure cuts and restricted access of citizens to their entitlements of publicly provided goods.[2] This paper’s central insight is two-fold: First, favoritism is more likely to occur in a fiscal crisis than absent a crisis. That is, to the extent that a crisis involves the government reneging on its commitments to citizens, it is natural to assume that favoritism and suspension of equal treatment are likely to arise, if only as the outcome of competition between people to access their entitlement to publicly provided goods. Second, and consequently, expectations of the burden of the crisis being allocated in an uneven way, generates an ex-ante political support in favor of fiscal indiscipline.

This mechanism sheds light on the insights of Dornbusch and Edwards. I show that greater favoritism favors higher public expenditure, debt financing, and raises the likelihood of fiscal crises. In other words, by pursuing unsound fiscal policies, the favored groups somehow “engineers” future crises. Thanks to crises, these favored groups manage to have their entitlements financed on average by the unfavored groups.

To illustrate these effects, I first discuss a short illustrative model based on Saint-Paul et al. (2017). All agents are entitled to consuming a certain level of a publicly provided good – that is, in normal times, public good provision is based on principles of equity and anonymity. In a crisis, however, budget cuts imply that not everybody can access his or her entitlement, and I assume that some groups are better than others at getting it. For this feature, called favoritism, to reduce the sustainability of fiscal policy, the decisive voter must be favored relative to the mean, a central assumption underlying my results. If this holds, I show that people are more likely to finance public expenditure by debt, the greater the degree of favoritism, the lower income inequality, and the greater the probability of a crisis. One can also show that favoritism and the likelihood of a crisis raises the level of public spending, which also goes up (as in the standard literature that follows from Meltzer and Richard (1981)) with income inequality.

I also tackle the issue of favoritism arising endogenously as the outcome of collective choice. I consider the choice between electing a “populist” who reneges on anonymity when allocating the public good, even in normal times, and a “technocrat” who sticks to anonymity, and does all it takes to balance the budget so as to ensure that as many citizens as possible can access their entitlement – thus the technocrat will restrict access to public goods only if there is a fiscal crisis, and will do so by implementing random rationing. I show that the support for the populist is greater, the greater the likelihood of a crisis. This means in particular that populists are more likely to conquer power (i) the greater the likelihood of default, (ii) the more depressed the macroeconomic environment, (iii) the greater the inherited level of public debt and (iv) the lower the state’s fiscal capacity. Somehow, there is a connection between populism in a sense often used by the popular mainstream press – a populist party favors some groups at the expense of others – and macroeconomic populism in the Dornbusch and Edwards sense.[3]

I then provide empirical evidence supporting the theory. I first discuss a case study, that of French pension reform, and show how it can be interpreted in light of the model: That is, civil servants and wage earners covered by specific pension regimes supported unsustainable reductions in the retirement age because they had good reasons to believe that other workers would bear a higher proportion of the burden of future pension reforms. I then provide evidence based on a

[2] *However, as briefly discussed below, favoritism also potentially arises when one consider tax hikes, although the definition of a fiscal crises used here assumes such tax hikes to be impossible.*

[3] *Both definitions differ from that of Acemoglu et al. (2013), who consider populism as a left-wing phenomenon.*

panel of countries. I match four datasets: the IMF's World Economic Outlook for macro indicators, the Institutional Profiles Database (IPD) for indicators of favoritism at the micro/institutional level, the Database of Political Institutions (DPI) for indicators of party ideology, and the CRAG-Bank of Canada database of sovereign defaults to get proxies for fiscal crises. Based on an IPD indicator of equality of treatment, I show that favoritism raises public debt, budget deficits, and public spending, as predicted by the theory. Favoritism raises the likelihood of a fiscal crisis through its effect on public debt. Finally, to test for the effects on collective choices, I define a "populist" party as either nationalist, regional, rural or religious. This definition can be criticized but, based on the indicators available in DPI, comes closest to a measure of whether a party favors some groups (defined by non-economic characteristics) at the expense of others. I then show that populists are more likely to conquer power, the higher the degree of debt and budget deficits, and the higher the level of government spending. The two first findings are consistent with the model's predictions. The last one may capture other mechanisms, but is also consistent with my predictions if interpreted as the effect of fiscal capacity.

The present paper contributes to the literature on the effect of institutions on the performance of macroeconomic policy. In contrast to the work of e.g. Persson et al. (2000) or Cukierman (2003), the focus here is not on the role of formal institutions such as political constitutions or the status of the Central Bank, but on the more informal one of favoritism.

Before proceeding to the analysis, it is useful to discuss the different forms of favoritism.

2. Forms of favoritism

We can enumerate a number of mechanisms by which favoritism would arise, especially so in a crisis. We may distinguish between *explicit* and *implicit* favoritism.

2.1 Explicit favoritism

A most salient form of explicit favoritism is *ethnic discrimination*. Hitler's Nuremberg laws, for instance, included provisions for expropriating Jews. They were deprived of their nationality, right to vote, and access to many professions. They were also banned from using public goods; for example, Jewish children were excluded from public schools in Berlin in 1937; they were banned from owning a car (thus using public infrastructures) or having a driver's license in 1938.[4] Historically, many governments have regularly struck the Jewish community with discretionary tax levies, especially in times of fiscal crises.[5] In modern democracies, anonymity and equity prevail in principle, but ethnic favoritism may arise in an opaque way as an outcome of identity politics (recall the controversy in the US about the Community Reinvestment Act).[6]

While nationality and ethnicity are obvious criteria, the frontier between favored groups and disadvantaged ones may obey other criteria, such as for example occupational ones. The example of French pensions discussed below shows that civil servants have enjoyed preferential statutory treatment in the pension reforms, and still enjoy better terms that are indirectly financed by other occupational groups. It is useful to interpret such discrepancies as an outcome of differences in *bargaining power* between civil servants and private sector employees that lead to the latter bearing most of the burden of adjustment in a fiscal consolidation, much in the fashion of Alesina and Drazen (1991).

[4] *A concise account of those laws can be found at http://alphahistory.com/holocaust/anti-jewish-laws/*

[5] *See for example Gerber, 1980.*

[6] *See Banerjee and Pande (2007) for an analysis of how ethnic preferences in political parties may lead to a reduction in the quality of elected officials, in particular in the dimension of corruption. See also Franck and Rainier (2012) on ethnicity and Grim and Finke (2006) on religion.*

A natural source of explicit favoritism is the one associated with income and embedded in the *progressivity of the tax system*. The more progressive this system, the more people expect adjustment to fall upon a small set of people – the richest – to the extent that adjustment involves tax hikes.[7] As shown by Ferrière (2015) both empirically and theoretically, more progressive tax systems raise the political support for and level of public debt which in turn makes default more likely. The theory outlined here focuses on differences in access to publicly provided goods that are not due to income.

2.2 Implicit favoritism

A natural source of favoritism is *corruption*.[8] 0ne may consider corruption as a perfectly competitive market: any publicly provided good (say a driving license) can be obtained by any citizen at the going market bribe. In such a situation favoritism arises only insofar that there is an income effect: the rich's willingness to pay for the good being higher than the poor's. If one considers, on the other hand, that some individuals are more corrupt than others, then more corrupt citizens lose less from restricted access to publicly provided goods in a crisis, since they are disproportionately likely to be granted their entitlement by paying bribes. Also, on the supply side, more corrupt officials may also benefit from a fiscal crisis to the extent that tighter rationing may help them raise the bribes they charge.

Similarly, differential access to *tax evasion* in the case of a tax hike may provide a mechanism for favoritism to operate.[9] Clearly, groups who have better access to tax evasion, when expecting that crises would lead to tax increases, are less likely to oppose macroeconomic populism. This may have some relevance in explaining the policies that led to the Greek sovereign crisis, for example.

Another source of inequities in access to publicly provided goods is the role of *social networks*, in particular in conveying information about procedures and opportunities. For example, a news piece by Le Boucher (2010) reports that the growing intricacies of the system to attend the elite *grandes écoles* in France favors insiders, i.e. in particular people whose relatives work at the ministry of education.[10]

I now formally discuss the consequences of favoritism for the political economy of spending, indebtedness, and crises.

[7] *This mechanism is absent from our theoretical analysis below, because by definitions crises occur when the government has exhausted its fiscal capacity and is only left with the option of reneging on its commitments over publicly provided goods.*

[8] *See Olken and Pande (2012) for a survey on the measurement of various forms of corruption.*

[9] *Gerber (1980) discusses how some members of the jewish community in Morocco could bribe officials to escape tax levies, which in turn undermined the community's ability to resist such hikes.*

[10] *Saint-Paul (2014) discusses a model where social connections gives one an edge in accessing rationed goods (such as public housing), so that policies of price controls that lead to such rationing are likely to be supported by people well endowed in social connections, despite that such policies are inefficient. Afonso et al. (2015) document the role of clientelism in the adjustment to the Greek fiscal crisis. See Robinson and Verdier (2013), for an analysis of clientelism.*

3. A model

In this section I sketch a formal model of fiscal policy under favoritism. I use it to analyze basic intuitions-the reader can refer to Saint-Paul et al. (2017) for a related model with formal results.[11]

My key assumptions are the following:

- Society precommits on an "entitlement" level of the publicly provided good, denoted by G. This means that any individual is entitled to consume G. For example, any person may have the right to use the local public library, to a certain number of years of education, to access day care, and so forth. However, the government can deny access to some individuals, which may be picked more or less randomly, but that is costly. This default on the government's commitment to the people may occur because a fiscal crisis may force a cut in expenditures, or, as discussed below, as an outcome of collective choice, in particular due to a 'populist' government allocating public spending in a discriminatory fashion.

- Society is partitioned into groups, indexed by a parameter λ. Groups with a higher value of λ have a better access to their entitlement. This access is summarized by a function $\psi(\phi, \lambda)$, assumed to be C^2, where ϕ is the aggregate probability of getting one's entitlement; that is, $\phi = \hat{G}/G$, where G is the entitlement level and $\hat{G}$ the actual spending on publicly provided goods. Assuming that there is a continuum of groups with λ uniformly distributed over [0, 1], we must have that

$$\int_0^1 \psi(\phi, \lambda) d\lambda \equiv \phi.$$

  It is natural to assume that no group becomes better-off in accessing the public good if it becomes harder to get on average, that is,

$$\psi_1 \geq 0.$$

  The assumption that access is easier for more highly ranked groups reads as

$$\psi_2 \geq 0. \tag{1}$$

  A special case is random rationing, i.e. $\psi(\phi, \lambda) = \phi,\ \forall \lambda$. In such a case every group has the same access to the publicly provided good. Otherwise, the process for allocating the public good is discriminatory over some range, implying from (1) that $[\psi(\phi, 1) > \psi(\phi, 0)]$ for a range of values of ϕ.

- As in Meltzer and Richard (1981), the publicly provided good has a redistributive dimension. In each group there are rich and poor people. Aggregate GDP is denoted by y, and is subject to random fluctuations – i.e. y is drawn from a distribution with support $[\underline{y}, \bar{y}]$ and density $f()$. The income of a poor (of any group) is βy, while the income of a rich is γy, where $\beta < 1 < \gamma$. For this to be consistent, the proportion of rich people in the population, θ, must be such that $(1-\theta)\beta + \theta\gamma = 1$, that is

$$\theta = \frac{1-\beta}{\gamma-\beta}.$$

[11] *Relative to that paper, the model presented here has a more flexible representation of favoritism and allows to analyze the interplay between favoritism and fiscal and economic condition when choosing between "populist" and "non-populist" parties. On the other hand, it delivers fewer clear-cut predictions.*

- Taxes are proportional to income. Let τ be the equilibrium tax rate. Then a poor consumes $\beta y(1-\tau)$ of the generic consumption good, while a rich consumes $\gamma y(1-\tau)$. Fiscal policy will be determined by a political equilibrium as in Meltzer and Richard. In this paper I do not prove any results regarding such an equilibrium. Instead I simply assume that there exists a decisive voter who is poor and belongs to some group λ.[12]

- Public spending may be financed by taxes or debt. I introduce debt as follows: Public debt is determined prior to the realization of the shock y and is used to finance a proportional tax credit. Therefore, denoting the debt level by D, each poor gets a credit equal to βD and each rich gets a credit equal to γD. I normalize the interest on debt to zero, so that the total tax receipts that are needed to finance the government's commitments simply equal $D+G$.

- A fiscal crisis may occur under poor macroeconomic conditions. I assume a maximum possible tax rate denoted by $\bar{\tau}$., and referred to below as *fiscal capacity*. More specifically,

  – If
$$y \geq \frac{D+G}{\bar{\tau}},$$
  the tax rate is set to
$$\tau = \frac{D+G}{y}$$
  and all agents get their entitlement level G.

  – If
$$y < \frac{D+G}{\bar{\tau}},$$
  aggregate expenditures are equal to
$$\hat{G} = \bar{\tau}y - D < G$$
  and access to the public good is rationed so that $\phi = \hat{G}/G < 1$. Furthermore, such rationing is costly. I assume it entails a disutility loss incurred by all agents and equal to
$$\delta = \varepsilon G(1-\phi).$$
  I shall assume that the marginal cost of rationing, ε, is smaller than β. This makes it potentially valuable for the decisive poor to issue debt at the cost of raising the likelihood of a crisis. I denote by $A = F\left(\frac{D+G}{\bar{\tau}}\right)$ the probability of a crisis.

- The utility of any given individual is given by
$$U(c, G, \psi, \delta) = c + \psi G - \delta,$$
  where c is his consumption of the private good and ψ his probability of accessing his entitlement.

  From there it is easy to see that absent a crisis, the utility of a poor of any group is equal to
$$u_P(y, G, D, \lambda) = \beta y + (1-\beta)G.$$

[12] *Again, see Saint-Paul et al. (2017) for a formal characterization of the equilibrium.*

In particular, it does not depend on D : one euro of tax credit is paid back as one extra euro of tax liability. On the other hand, if there is a crisis, this utility is equal to

$$u_P(y,G,D,\lambda)=\beta y(1-\bar{\tau})+\beta D+\psi(\frac{\bar{\tau}y-D}{G},\lambda)G-\varepsilon(G+D-\bar{\tau}y).$$

Fiscal policy consists of G, the entitlement level, and D, the debt level. I assume that these two quantities are decided by the decisive voter, who is poor and of type λ_d, prior to the realization of the shock y. Therefore the equilibrium values of D and G maximize

$$\begin{aligned}V(G,D) &= E_y u_P(y,G,D,\lambda_d)\\ &= \int_{\underline{y}}^{\frac{D+G}{\bar{\tau}}}\left(\beta y(1-\bar{\tau})+\beta D+\psi(\frac{\bar{\tau}y-D}{G},\lambda)G-\varepsilon(G+D-\bar{\tau}y)\right)dy\\ &\quad+\int_{\frac{D+G}{\bar{\tau}}}^{\bar{y}}(\beta y+(1-\beta)G)dy.\end{aligned}$$

We want to analyze how favoritism affects fiscal discipline, i.e. how it affects the equilibrium values of G and D. For this we can simply look at the marginal utility to the decisive voter of raising D and G.

Let us start with the incentives to issue debt, for any given G. Clearly, we have that

$$V_2(G,D)=\beta A-\int_{\underline{y}}^{\frac{D+G}{\bar{\tau}}}\left(\psi_1(\frac{\bar{\tau}y-D}{G},\lambda_d)+\varepsilon\right)f(y)dy.$$

All the terms in this expression come from states of fiscal crisis. This is because absent a crisis, Ricardian equivalence holds: an extra euro of tax credit financed by debt is simply matched by an extra euro of taxes. In a crisis, raising taxes is impossible because they are constrained by fiscal capacity. Therefore, the individual experiences a gain from the tax credit (the βA term), but that is financed by a reduced access to the public good (captured by the ψ_1 term). To this cost should be added the distortionary effects of rationing (the ε term).

Let us now discuss how the choice of debt is affected by favoritism. We start from the benchmark case of no favoritism; then $\psi(\phi,.)\equiv\phi$ and $\psi_1=1$, implying that $V_2=-A(1+\varepsilon-\beta)<0$. Absent favoritism, the decisive poor always loses from increasing debt. This is because in a crisis, a dollar of additional government liability is matched by a dollar of reduced expenditure. Since the poor only pay β dollars per dollar of extra taxes, their monetary gain from the increased debt is lower than the monetary value of their reduced consumption of the public good; and these losses are compounded by the distortions induced by rationing. Under proportional rationing, then, the decisive voter is averse to debt; more so, the greater the probability of a crisis A.

What happens, now, if the decisive voter is favored "at the margin" in accessing his entitlement? This means that his probability of being served, ψ , only falls by a small amount when ϕ falls, i.e. that ψ_1 is small. Intuitively, this will be the case if group λ_d is among the groups that are "served first", while other groups would bear most of the adjustment burden. Clearly, then, V_2 may be positive, more so, the smaller ψ_1 , i.e. the more the decisive group is protected from the burden of adjustment. The above formula also implies that the propensity to issue debt is larger, the larger β and the greater the probability of crisis A . The larger β, the greater the gains to the decisive group of voting for a tax credit in exchange for rationing G in times of crises. Since it is in crisis times that the gains are incurred, they are greater, the more likely the crisis is.

Let us now turn to the determination of G, the entitlement level of the public good. We have that[13]

$$V_1(G,D) = (1-A)(1-\beta) + \int_{\underline{y}}^{\frac{D+G}{\bar{\tau}}} \left(H(\frac{\bar{\tau} y - D}{G}, \lambda_d) - \varepsilon \right) f(y)dy, \tag{2}$$

where $H()$ is defined as

$$H(\phi, \lambda) \equiv \psi(\phi, \lambda) - \phi \psi_1(\phi, \lambda).$$

The first term in (2), $(1-A)(1-\beta)$ tells us that absent a crisis, there is a net gain to the poor from raising G, because they only pay β euros per euro spent. This is the standard Meltzer and Richard effect. The second term tells us that the marginal gain from increasing G in crisis states is equal to the probability of being served, ψ , from which one deducts the distortionary cost induced by the additional rationing, ε, as well as the average reduction in the amount of G consumed due to rationing, $\phi \psi_1$. To the extent that the decisive group is favored, ψ is large while ψ_1 is low; therefore we expect that $H > 0$, more so, the more favored the group (Note that $H = 0$ if $\psi \equiv \phi$). Altogether, this discussion implies that society will choose a greater entitlement level, the greater inequality, and the greater favoritism. Furthermore, if favoritism is very large, one will have that $H - \varepsilon > 1 - \beta$,, and V_1 and the optimal level of G will go up with the crisis probability A. On the other hand, if favoritism is not too strong, G will fall with the probability of crisis A .

At this stage, it is useful to summarize the predictions of this section:

- *For a given level of favoritism, debt will be higher, (i) the greater the probability of a crisis, (ii) the lower the inequality between the politically decisive voter and the average one. Furthermore, debt goes up with the degree of favoritism.*

- *For a given level of favoritism, public spending goes up with inequality. Furthermore, public spending goes up with the degree of favoritism. Finally, a rise in the crisis probability raises public spending if favoritism is very large, but reduces it if it is moderate.*

3.1 Endogenizing favoritism: the populist as a discriminator

So far, the discussion has assumed that favoritism is a structural property of the society under study. In reality, favoritism may be the outcome of collective choices – indeed, all political parties represent the interests of some specific groups of people. Indeed, much of the debate about populism, and many of the arguments of the so-called populist politicians, center around whether national citizens should have a better access to publicly provided goods than non-nationals. This example shows that people can choose some degree of discrimination by electing a populist or not, even though nationality or ethnicity are only one set of attributes along which one may discriminate.

The preceding section highlights the potential benefits of voting for a populist: his policies will grant privileged access to the favoured groups in a crisis. In this section I analyze the incentives for putting a populist in power. I assume people can choose between two kinds of politicians: a technocrat and a populist.

[13] *We ignore any resource cost of G, for simplicity. Of course, actual spending GA has a resource cost and must be financed by taxes; but G is the nominal entitlement level of citizens to the publicly provided good; it is a social contract, not an actual production activity.*
See Saint-Paul et al (2017) for a full analysis which embodies a convex resource cost for G and a complete characterization of the political equilibrium.

3.1.1 The technocrat

The technocrat does whatever it takes to balance the budget while providing the entitlement level G. Therefore, he will set $\tau = \bar{\tau}$ and $\hat{G} = \bar{\tau}y - D$ in a crisis, and $\hat{G} = G$, $\tau = (D + G)/y$, absent a crisis. The outcome is the one described in the preceding section, with $\psi \equiv \phi$.

3.1.2 The populist

The populist favors a specific group, denoted by λ_p. He always allocates the public good on the basis of favoritism, i.e. with access probability $\psi(\phi, \lambda)$. Furthermore, upon realization of the shock y, he picks the value of $\hat{G}$ which maximizes the utility of group λ_p. Potentially, this means that once the entitlement level G is set the populist may elect to ration access to the public good, even though there exists a feasible tax rate which guarantees universal access. If fiscal capacity is not binding, then $\tau = (D + \hat{G})/y = (D + \phi G)/y$, and the utility of group λ_p is given by

$$\beta y(1 - \frac{D + \phi G}{y}) + \psi(\phi, \lambda_p)G - \varepsilon(1 - \phi)G + \beta D.$$

Consequently, as long as fiscal capacity is not binding, the populist picks $\hat{G} = \phi^* G$ such that

$$\phi^* = \arg\max_{\phi} -(\beta - \varepsilon)\phi + \psi(\phi, \lambda_p).$$

If $\phi^* < 1$, access is rationed regardless of the realization of the macroeconomic shock y. This is because rationing allows to reduce taxes, which favors group λ_p as long as its rank is high enough, since its access to G is not much reduced as ϕ falls.

3.1.3 Choosing between the technocrat and the populist

I now discuss who gains and who loses from populism, depending on the current state of the economy, and conditional on the preset values of G and D. I focus on the case where $\phi^* < 1$ and distinguish between three possible outcomes:

- Normal times, where neither a technocrat nor a populist would be constrained by fiscal capacity, which occurs if

$$y \geq \frac{D + G}{\bar{\tau}}.$$

- Crisis, where the technocrat would be constrained by fiscal capacity and would have to implement rationing, while the populist can pursue his unconstrained policy. This occurs if

$$\frac{D + \phi^* G}{\bar{\tau}} \leq y < \frac{D + G}{\bar{\tau}}.$$

- Supercrisis, where fiscal capacity is binding for both kinds of participants, which occurs if

$$y < \frac{D + \phi^* G}{\bar{\tau}}.$$

In a supercrisis, both the technocrat and the populist implement the same rationing level, $\phi = \frac{\bar{\tau}y - D}{G} < \phi^*$.

I now compute the set of people who gain and lose from the populist being in power, for each of those environments.

Preferences for populism: Normal times In normal times, the utility of group λ from the technocrat being in office is given by

$$U_T(y, G, D, \lambda) = \beta y + (1-\beta)G;$$

if the populist is in power, utility is given by

$$U_P(y, G, D, \lambda) = \beta(y - \phi^* G) + \psi(\phi^*, \lambda)G - \varepsilon(1-\phi^*)G. \tag{3}$$

We note that D is absent from these formulas: absent a fiscal crisis, Ricardian equivalence holds. It is then clear that

$$U_P > U_T \Leftrightarrow \lambda > \lambda_N,$$

where λ_N is the solution to

$$\psi(\phi^*, \lambda_N) = 1 - (\beta - \varepsilon)(1 - \phi^*). \tag{4}$$

Not surprisingly, those who support the populist are those whose rank is above some critical λ_N, i.e. those who are favored by the allocation of public resources. Furthermore, $\partial \lambda_N / \partial \beta < 0$, meaning that when inequality is lower, the poor benefit more from the tax cuts implemented by the populist, which increases the support for the latter.

Crisis In a crisis, the utility of group λ from the technocrat has to be changed and is now given by

$$U_T(y, G, D, \lambda) = \beta y(1-\bar{\tau}) + \beta D + \bar{\tau}y - D - \varepsilon(G + D - \bar{\tau}y). \tag{5}$$

The term $\bar{\tau}y - D$ comes from the fact that $\psi(\phi, \lambda)G = \phi G = \hat{G} = \bar{\tau}y - D$. The formula for U_P is unchanged, and supporters of the populist are now ranked above some critical $\lambda_C(y)$ which is defined by

$$\psi(\phi^*, \lambda_C(y)) = (\beta - \varepsilon)\left(\phi^* - \frac{\bar{\tau}y}{G}\right) + \frac{\bar{\tau}y}{G} - (1 - \beta + \varepsilon)\frac{D}{G}. \tag{6}$$

From this formula, we can establish the following predictions:

First, as long as $D \geq 0$, $\lambda_C < \lambda_N$.[14] This means that the support for the populist is larger in times of crisis than in normal times. The technocrat somehow loses his comparative advantage in times of crisis, because he is forced to implement rationing, which the populist does anyway. As the marginal voter in normal times λ_N is relatively favored, in crisis times he switches to the populist who offers a better access to his entitlement, given that such access is rationed.

Second, $\partial \lambda_C / \partial D < 0$. An increase in the stock of debt raises the support for the populist. This is because the greater the stock of debt, the greater the degree of rationing that the technocrat implements in a crisis. As seen above, this harms the poor because, through debt, they only get β euros of tax rebate per euro in average reduction of their public good consumption. In contrast, the stock of debt has no effect on the populist's choice for ϕ, because he is not constrained by fiscal capacity. Thanks to his choice of rationing people, the populist has a greater fiscal margin

[14] *From (4),* $\psi(\phi^*, \lambda_N) - \psi(\phi^*, \lambda_C(y)) = (1 - \beta + \varepsilon)(\frac{D}{G}. + 1 - \frac{\bar{\tau}y}{G}) > 0$. *Since* $\psi_2 > 0$, it follows that $\lambda_N > \lambda_C(y)$.

of maneuver, and raises taxes when debt goes up, leaving ϕ^* and therefore ψ unchanged, which is a better adjustment strategy for the poor than what the technocrat does, i.e. reducing ϕ.

Third, the support for the populist is greater, the more severe the crisis (the lower y) and the lower the state's fiscal capacity (the lower $\bar{\tau}$). In both cases, the technocrat is constrained to increase the extent of rationing by reducing ϕ, while the populist raises τ while maintaining the same degree of access to the publicly provided good.

Supercrisis In a supercrisis, both politicians are constrained by fiscal capacity. As a result, (3) has to be replaced with

$$U_P(y,G,D,\lambda) = \beta y(1-\bar{\tau}) + \beta D + G\psi\left(\frac{\bar{\tau}y - D}{G}, \lambda\right) - \varepsilon(G + D - \bar{\tau}y). \tag{7}$$

Comparing (7) with (5), we find that people support the populist provided their rank is higher than $\lambda_S(y)$, where $\lambda_S(y)$ is given by

$$\psi\left(\frac{\bar{\tau}y - D}{G}, \lambda_S(y)\right) = \frac{\bar{\tau}y - D}{G}.$$

This formula is simple to understand. In a supercrisis, both the populist and the technocrat will spend the same amount on the public good, $\hat{G} = \bar{\tau}y - D$. Hence, they implement the same degree of rationing, which induces the same distortions. Under a technocrat each group is served its entitlement with the same probability $\phi = \hat{G}/G$. Under a populist, group λ is served with probability $\psi(\phi,\lambda)$. Group λ favors the populist if and only if it has a greater than average probability of being served under the populist, which is equivalent to $\lambda > \lambda_S$.

We also note that

$$\frac{d\lambda_S}{d\phi} = \frac{1-\psi_1}{\psi_2} \tag{8}$$

This quantity is positive iff $\psi_1 < 1$, i.e. if the pivotal group is marginally favored. If that is true, λ_S falls when ϕ falls, implying that the support for the populist will go up, the smaller the government's "fiscal space", i.e. the smaller y, the smaller $\bar{\tau}$, and the larger D. Also note that $\lambda_S(\frac{D+\phi^* G}{\bar{\tau}}) = \lambda_C(\frac{D+\phi^* G}{\bar{\tau}}) < \lambda_N$. By continuity, $\lambda_S < \lambda_N$ for y not too small. Furthermore, if in addirtion the RHS of (8) is always nonnegative for $\phi < \phi^*$, then since $\frac{d\lambda_S}{d\phi} > 0$, one has that $\lambda_S < \lambda_N$ for any $y \leq \frac{D+\phi^* G}{\bar{\tau}}$, i.e. in any supercrisis. We can conclude that in a supercrisis, as in a crisis, the support for populism is typically larger than in normal times.[15]

To summarize, this section has studied the endogenous political choice between a technocrat and a populist. Key results are summarized as follows:

- *The support for a populist government as opposed to a technocratic one, is larger, the greater the required degree of adjustment. This means that it is larger in crises than in normal times, and typically larger, the higher the inherited level of public debt, the lower the state's fiscal capacity, and the more adverse current macroeconomic conditions are.*

[15] *However, the opposite case is not entirely ruled out for severe supercrises if* $\frac{1-\psi_1}{\psi_2} < 0$ *over some range.*

4. Empirical evidence

I now provide some empirical evidence, based on two different approaches.

First I discuss the recent history of French pension reforms. I argue it highlights the mechanisms analyzed above. Essentially, supporters of unsustainable reductions in the retirement age had good reasons to anticipate that subsequent adjustments were likely to hit other social groups proportionally more.

Second, I provide evidence across a panel of countries that supports the above predictions. Unequal treatment from administrations is more likely to generate high debt, high public expenditures, and high deficits, as well as (indirectly through debt) sovereign default. Also, I study the determinants of populism in government and show that, consistent with the predictions in Section 3.1, adverse fiscal conditions such as high public debt, high deficits and low fiscal capacity are more likely to lead to a populist government. On the other hand, there is no robust evidence that adverse macroeconomic conditions have any effect on the likelihood of populism.

4.1 The French pension reform saga

While the French general government budget has never encountered a formal fiscal crisis which has led to cut in entitlements, the pension system traditionally has its own separate budget and its financial difficulties and the reforms that they have triggered do illustrate the mechanisms highlighted by the model.

In 1981, the left-wing Mitterrand administration was elected. A notable characteristic of the French left is that it is disproportionately supported by civil servants. The following Table, taken from Roubaud (1999), depicts the evolution of the vote for the main two French left-wing parties of the time, the Communists and the Socialists, for public vs. private sector employees. We see that civil servants vote for these parties much more than private sector employees; the difference is 10 to 15 points.

Table 1: Percent vote for the French Communist Party (PC) and the French Socialist Party (PS) in parliamentary elections.

	1973		1978		1986		1993		1997	
	PC	PS	PC	PS	PC	PS	PC	PS	PC	PS
Private	13.3	23.3	18.4	21.7	7.1	38.6	7.1	25.8	9.1	23.6
Public	18.3	32	23.2	28	9.3	46.9	10.1	32.5	11.4	30.4

Source: Roubaud (1999).

A key point in Mitterrand's electoral platform was the reduction in the retirement age from 65 to 60. This reform was implemented despite demographic projections that indicated such a measure was financially unsustainable in the long run, as pointed out by the 1991 *Livre Blanc* (white paper).[16] Another important feature of the French public pension system, is that it is split in different *régimes*. This means that workers are treated differently by the pension system depending on their industry, occupation, or type of labor contract. For example, civil servants are part of a different régime from private sector employees, while many publicly owned companies (in particular EDF, SNCF, RATP and the Banque de France) have their own regimes, called *régimes speciaux*. Regimes differ from each other in terms of (i) the age at which people may retire, (ii) the number of years of contribution necessary in order to retire, and (iii) the amount of the pension, in relation to the total amount that has been contributed. Given their entitlements, dif-

[16] *See French Office of the Prime Minister (1991).*

ferent regimes have different financing needs, and restoring fiscal balance is likely to involve cross-subsidies. Absent such cross-subsidies, some regimes would have to levy very large taxes upon their active members in order to fulfil their commitments to their pensioners. These inequities reflect the respective bargaining strength of the different social groups involved in the pension game. Publicly owned companies typically are strongly organized by the major labor union CGT, while this is less so for civil servants, and even less so for private sector employees, whose unionization rate is about half that of their public sector counterparts.[17] While this balance of power implies that some groups are favored on *average* (i.e. have a higher value of ψ in terms of the above model), it is also likely that they will be favored at the *margin* when the inevitable fiscal consolidation of the pension system is implemented (i.e. have a lower value of ψ_1). According to my analysis, these groups are more likely to favor unsustainable increases in the generosity of the system, as was the case for civil servants who supported Mitterrand in 1981 to a greater extent than private sector employees.

The first attempt to balance the accounts of the pension system came with the 1993 "Balladur" reform. While it did not formally overturn the reduction in the retirement age, it did make this entitlement much more difficult to obtain for a fraction of the population, namely private sector wage earners.[18] While having reached the retirement age is one necessary condition for becoming a pensioner, there is another one having to do with the duration of contributions. The 1993 reform raised the duration of contributions from 37.5 years to 40 years *for private sector employees only*, thus making it less likely that they be able to effectively retire at 60. In the public sector, the 37.5 year rule was left untouched. It took 10 more years for a subsequent reform, the 2003 "Fillon" reform, to align the required duration of contributions of the public sector to that of the private sector. The Balladur reform also toughened the conditions for retirement for private pensioners only in other dimensions: private sector pensions were now based on the average wage earned over the 25 best years, instead of the 10 best years,[19] they were indexed on the CPI instead of the average wage (in times of productivity growth, the latter grows faster than the former).

Not surprisingly, as illustrated on Table 2, the reform curbed expenditures on the *régime général*, while expenditures on the régime for civil servants continued to grow at a higher rate.

Table 2: Growth rate of expenditures on the three main regimes, 2000–2002.

Regime	Growth (%), 2000–2002
Private	6.2
Central government	7.5
Local government	13.5

Source: Conseil d'Orientation des Retraites (2004).

It is highly plausible that the right-wing Balladur administration shied away from a reform of public pensions because of the superior ability of public employees' unions to organize in order to block a reform[20] – as was painfully experienced by his successor Juppé, who had to withdraw a proposal for reforming the *régimes spéciaux* in the face of violent protests. In turn, consistently with the model, it was rational for public sector employees to support Mitterrand

[17] *See DARES (2016).*

[18] *See Observatoire des Retraites (2009).*

[19] *As of today, the civil servant's pensions is based on the last 6 months of wages, reflecting their highest level throughout their career, due to seniority.*

[20] *Given its ideological positioning, it is natural, as far as the pension issue is concerned, to interpret the right-wing government of Balladur as a "Technocrat" and Mitterrand as a "Populist". However we do observe that under pressure from the unions, Balladur failed to act as a technocrat and instead implemented a reform which increased inequities, at the expense of his own constituency.*

in the 1981 election, despite overwhelming evidence that the reduction in the retirement age was fiscally unsustainable; the civil servants knew that because of their greater clout in resisting reforms, any adjustment was likely to fall predominantly on private sector employees, who ended up effectively subsidizing the civil servants' superior entitlement.

4.2 Evidence from a panel of countries

I now provide some more formal statistical evidence in favor of this paper's key hypotheses.

My most central argument is that lack of equal treatment of citizens by the administration is likely to deliver fiscal indiscipline at the macroeconomic level. To test for this hypothesis, I use the institutional profiles database (IPD). It includes a large number of indicators of institutional quality across countries, in particular with respect to equality of treatment of citizens by the state.[21] Furthermore, the IPD survey is conducted every four years, implying that these indicators can be organized in a panel. In practice, the number of countries is quite small before 2012, so that I construct a panel with only two time observations by matching the 2012 and 2016 waves of IPD-the fixed effects estimates reported below are in effect difference-in-differences estimates. These two waves contain data on some 300 institutional indicators for 144 countries. For all these indicators, a higher number indicates better institutions. Therefore, for example, a country is better at collecting taxes, the higher its fiscal capacity indicators, and less corrupt, the higher its corruption indicators. One should keep this in mind when reading the estimates below.

I use five indicators to proxy for favoritism:

- Equality of treatment (A1032), which measures "equality of treatment of citizens in their relationship with the administration".
- "de facto equality of treatment of citizens by the public service" in the four specific areas of schooling (A9040), health (A9041), formalities (A9042) and access to public jobs (A9043).

For these indicators, I estimate their effect on a macro fiscal performance measure, controlling for time and country fixed effects, GDP per capita, as well as three IPD-based composite indicators of institutional quality that are likely to affect aggregate fiscal discipline:

- A measure of fiscal capacity, constructed as the first principal component of the following IPD fiscal efficiency indicators: state efficiency in collecting corporate taxes (A3030), income taxes (A3031), and taxes across the territory (A3032).
- A measure of conflict, constructed as the first principal component of the following IPD conflict indicators: ethnic/religious/regional (A2020), social (A2021), rural land (A2022), urban land (A2023), plus two measures of violent activities related to political organizations (A2040) and criminal organizations (A2041).
- A measure of corruption, equal to the first principal component of the IPD corruption indicators for small corruption (A3020), political corruption (A3021), corruption between administration and local firms (A3022) and corruption between administration and foreign firms (A3023). Arguably, this control variable can be interpreted as another measure of favoritism.

The macroeconomic variables are borrowed from the IMF's World Economic Outlook Data.

[21] *See https://www.tresor.economie.gouv.fr/Articles/2018/05/03/tresor-economics-no-221-institutions-and-development-insights-from-the-institutional-profiles-database-ipd*

4.2.1 The effect of favoritism on debt

I first look at the effect of favoritism on the level of public debt.

Table 3: Effect of equality of treatment on public debt

Controls			Equality of treatment
Fiscal capacity	Conflict	Corruption	A1032
			-2.7*
	-1.6***		-3.2**
1.4			-2.0
1.4	-1.6***	-2.3***	-4.3***
	0.1		-2.6*
		-2.3***	-4.2***
1.9*		2.6***	-3.7**
1.9*	0.07	-2.7***	-3.6**

*Dependent variable: Gross public debt relative to GDP. All regressions control for time and country fixed effects as well as GDP per capita in constant PPP terms.*** = significant at the 1% level; ** = significant at the 5% level; * = significant at the 10% level.*

Table 3 estimates the effect of equality of treatment on the ratio between gross public debt and GDP. For the main indicator A1032, in most specifications the effect is strongly significant with the expected sign: better equality of treatment reduces the debt/GDP ratio. The coefficients are virtually unchanged when one varies the set of controls. Note also that higher fiscal capacity has a positive effect on debt, suggesting it actually makes it easier for governments to borrow, in contrast to what one might believe based on the naive view that debt is an alternative to taxation. Corruption also raises debt levels in all specifications, consistent with our analysis if one is willing to interpret it as a form of favoritism. Finally conflict typically raises public debt (since a higher level of that variable indicates less conflict), in accordance with theories of divided government,[22] although the coefficient becomes essentially zero when corruption is controlled for.

In the Appendix I replace the favoritism measure by alternative ones, including all controls as in the last line of Table 3. The results are qualitatively identical when instead of equality of treatment A1032, I use either equality of access to public jobs (A9043) or of access to the administration (A9042). For the two other indicators – access to schooling and access to health – the coefficient on equality has the right sign but is not significant.[23]

[22] *Tabellini and Alesina (1990), Persson and Svensson (1989).*

[23] *On the other hand, the coefficients on fiscal capacity and corruption have the same sign as in the last line of Table 3 and are significant – See Table A1.*

4.2.2 The effect of favoritism on budget deficits

Alternatively, we can use net primary government lending as a fraction of GDP as a measure of fiscal discipline. The results (Table 4) essentially confirm those of Table 3. Equality of treatment A1032 has a significant positive effect on the primary budget surplus. Both conflict and corruption have a positive sign, meaning that more conflict and more corruption are conducive to budget deficits.

On the other hand, none of the specific indicators A9040-A9043 has any significant effect, although they all have the predicted positive sign (See Appendix).

Table 4: Effect of equality of treatment on primary budget surplus. Controls include time and country fixed effects as well as GDP per capita. Dependent variable: net primary government lending as a fraction of GDP.

Variable	Coefficient
Fiscal capacity	0.0
Conflict	0.7*
Corruption	1.3***
Equality of treatment (A1032)	2.6***

4.2.3 The effect of favoritism on public spending

Another prediction is that the decisive voter will tend to vote in favor of higher public expenditure, the greater the degree of favoritism. Such prediction is validated in Table 5, which regresses public spending as a share of GDP on our usual equal treatment indicators.

Table 5: Effect of equality of treatment on public expenditure. Same controls as for the preceding tables.

Variable	Coefficient
Fiscal capacity	0.31
Conflict	-0.6**
Corruption	-0.5*
Equality of treatment (A1032)	-1.8***

Again, equality of treatment has a significant negative effect on spending, as predicted by the model. This only holds, however, for the general indicator A1032, the other ones are again insignificant (See Appendix).

4.2.4 Favoritism and fiscal crises

Does favoritism make fiscal crises more likely to occur? To test for this prediction, I need a measure of sovereign default. To do this, I import data from the Bank of Canada CRAG database (See Beers and Mavalwalla, 2017). This panel of countries provides estimates of the total dollar amount owed to institutional creditors such as the IMF and the Paris Club. I construct a default indicator as the ratio between the amount owed the following year and the outstanding level of public debt (gotten from the WEO database). This, in principle, is an indicator of default in the year following the observation, although due to rescheduling and rollover of delinquent debt there is much serial correlation in this default variable. One other drawback is that since the second time observation of the IPD is 2016 and the CRAG database does not have data posterior to that date, while the 2016 data themselves have many missing values, I can no longer exploit the panel dimension of IPD and can only run a cross-sectional regression for 2012.[24] Given these shortcomings, we expect poorer quality results than in the preceding regressions. The estimates are reported in Table 6.

Table 6: Effect of equality of treatment on subsequent sovereign default rate.

Controls			Equality of treatment	Debt/GPD ratio
Fiscal capacity	Conflict	Corruption	A1032	
			-0.22	0.02***
	0.0		-0.22	0.02***
-0.23**			-0.12	0.02***
		0.1	-0.18	0.02***
-0.23**	-0.03		-0.13	0.02***
	-0.02	0.1	-0.19	0.02***
-0.22**		0.02	-0.11	0.02***
-0.23**	-0.04	0.03	-0.13	0.02***

Dependent variable: subsequent delinquent amount/outstanding public debt (Source: CRAG database). All regressions are cross-sectional for year 2012. They all control for GDP per capita.

Table 6 only provides mild support for a direct effect of favoritism on the likelihood of fiscal crises. While the equality measure A1032 always has the correct sign, it is never significant at the 15% level or less. Note however that the coefficient on equality captures its effect on default while *controlling for public debt*. In all specifications, the outstanding level of public debt has (unsurprisingly) a strong positive effect on default (See Table 6, column 4). Since, on the other hand, we have documented a significant effect of favoritism on public debt, clearly favoritism makes crises more likely through that channel. But the evidence in favor of a direct effect of its own is much weaker.

[24] *Instead of using subsequent default as our dependent variable, I could use contemporaneous default, which given the availability of CRAG for 2016, would in principle allow us again to use the panel dimension. However, the estimates would be polluted by the endogenous effect of default itself on favoritism. In any case, there are so many missing data for 2016 that running this exercise does not deliver any useful result.*

4.2.5 IV estimates

One issue with the estimates of Tables 3 to 5 is that the equal access measure may be correlated with the error term. For example, one may plausibly argue that high public debt, or more generally high fiscal strain, breeds favoritism as an outcome of the uneven effects across citizens of the public administration's attempts to save money. To control for this potential source of bias I instrument the (preferred) A1032 equality of treatment indicator with a set of variables from the IPD database which capture deeper institutional characteristics of a country that are unlikely to be affected by current budgetary developments and, to varying degrees, are arguably correlated with effective equality of treatment.[25] In addition to my equality of treatment variable, I also instrument for the conflict and corruption indicators, that may arguably suffer from the same endogeneity bias as the equality of treatment indicator.

The IV estimates of the effect of favoritism are reported in Table 7. For the sake of comparison I also report the OLS estimates from Tables 3–5.

Table 7: Instrumental variables estimates of the preferred regressions in Tables 1 to 3.

Dependent Variable (%GDP)	Debt		Net gvt lending		Public spending	
Covariates	OLS	IV	OLS	IV	OLS	IV
Fiscal capacity	1.9*	2.3*	0.006	-0.5	0.31	0.6*
Conflict	-0.07	1.4	0.71*	0.9	-0.6**	-0.7
Corruption	-2.7***	-6.1*	1.3***	0.5	-0.5*	0.4
Equality of treatment A1032	-3.6**	-6.2*	2.6***	-0.1	-1.8***	1.0

The picture that emerges from Table 7 is mixed. While conclusions are essentially unchanged for the determinants of debt (favoritism comes with the expected sign and is significant, as well as corruption; fiscal capacity again tends to raise debt), the IV estimates, in contrast to their OLS counterparts, do not show evidence of any significant effects of either the equality of treatment variable or the main controls of interest on either net lending or public spending. The evidence therefore seems less robust for the effects of favoritism on these two measures of fiscal performance than for public debt.

4.2.6 Explaining populism

In this section I investigate the prediction outlined above that adverse fiscal and macroeconomic conditions are conducive to populist governments. While the definition of populism is controversial, in the context of this theory it is natural to define a populist party as one whose platform would favor some social groups at the expense of others, for groups defined by characteristics other than income: regional, ethnic, religious, etc.

[25] *These instruments are: electoral freedom (A1000), regularity of electoral processes (A1001), representativeness of institutions (A1002), efficiency of control institutions (A1003), freedom of association (A1030), freedom of reunion (A1031), intensity of counterpowers (A1020), national participation (A1021), local participation (A1022).*

To construct such an indicator, I use the Interamerican Development Bank's *Database of Political Indicators* (DPI), which is a panel of countries for which a number of political variables are provided.[26] In particular, indicator variables capture whether a party is either (i) nationalist, (ii) rural, (iii) regional or (iv) religious. I define any such party as "populist". While the definition of populism is much open to debate, this particular one, among the variables available in DPI, is the one most consistent with this paper's definition of populism.

This leads me to construct four alternative measures of populism, based on the DPI indicators: 1. A dummy equal to one if the executive belongs to a populist party, in the sense I just defined; 2. A dummy equal to one if the ruling coalition party is populist; 3. That dummy multiplied by the fraction of parliamentary seats held by the main coalition party, and 4. That dummy multiplied by the fraction of votes obtained by the main coalition party. The last two indicators weigh the presence of a populist party in the government by power, measured as its relative importance among either MPs or voters.

Next, I match the DPI with my WEO database of macro outcomes to estimate a fixed effects regression of my populism indicators on lagged macro variables. In particular, based on the above theory, we expect adverse macroeconomic conditions such as a low output gap to raise the likelihood of a populist govenment. Similarly, adverse fiscal conditions such as high debt or large deficits should also lead to a greater likelihood of populism.

Table 8: Macroeconomic determinants of populism, fixed effects estimates. Dependent variable: (1)=populist executive dummy, (2)=populist main coalition party dummy, (3)=populist main coalition party dummy times fraction of seats held by this party, (4)= populist main coalition party dummy times fraction of votes obtained by this party. Coefficients multiplied by 100 except on GDP per capita. Time dummies included. † significant at the 15% level.

	(1)	(2)	(3)	(4)
A. Lagged dependent variable not included				
GDP per capita (-1) real PPP constant USD	4.5	2.8	-0.6	-0.4
Inflation (-1)	-0.1	-0.2	-0.06	-0.11
Output Gap (-1)	-1.0	-1.1†	-0.3	-0.4
Unemployment Rate (-1)	-0.5	-0.7	-0.5*	-0.6**
Gross debt/GDP ratio (-1)	0.1	0.1	0.08***	0.1**
Primary Gvt net lending/GDP (-1)	-0.3	-0.04	-0.08	-0.1
Gvt spending/GDP (-1)	0.5	0.4	0.15	0.1
Current account/GDP (-1)	1.0**	0.3	0.04	-0.7
B. Lagged dependent variable included				
GDP per capita (-1) real PPP constant USD	2.5	1.5	0.2	0.46
Inflation (-1)	0.15	0.05	0.02	0.003
Output Gap (-1)	0.1	0.2	0.08	0.02
Unemployment Rate (-1)	0.6	0.5	0.2	0.2
Gross debt/GDP ratio (-1)	0.1*	0.1*	0.04*	0.04*
Primary Gvt net lending/GDP (-1)	-0.9***	-0.7*	-0.2†	-0.2*
Gvt spending/GDP (-1)	-0.6*	-0.6t	-0.2*	-0.25*
Current account/GDP (-1)	0.4t	0.1	0.03	0.01

[26] *See https://publications.iadb.org/en/publication/12390/database-politicalinstitutions-2015-dpi2015*

The results are reported in Table 8. In the first panel I report 4 regressions where the lagged dependent variable is not included. All explanatory variables are lagged: hopefully the coefficients capture the causal effect of preexisting macro and fiscal conditions on the nature of government (populist vs. technocrat). The only salient finding is that a higher level of public debt raises the likelihood of a populist government – the coefficient is always positive and statistically significant in two specifications. The output gap and unemployment rates appear to have a negative, sometimes significant, effect but these estimates will turn out to be less robust than the estimates on fiscal variables.

Despite that explanatory variables are lagged, these estimates may reflect the effect of populism on macro and fiscal performance as much as the converse. One way to alleviate this problem is to add the lagged dependent variable as a regressor (the effect of an incumbent populist government on macro and fiscal performance at $t-1$ would then be reflected in the correlation between the lagged dependent variable and the other covariates, as opposed to a correlation between the latter and the error term).

The results where the lagged dependent variable is included are reported in the bottom panel of Table 8. Macro variables are no longer significant. On the other hand, fiscal variables become more significant. Debt now has a positive significant effect on populism in all specifications, while the budget surplus has a negative significant effect in all specifications. Finally, under inclusion of the lagged dependent variable, we see that a higher level of spending reduces the likelihood of populism.

Given the presence of the lagged dependent variable, it makes sense to compare these estimates with GMM ones.[27] This is done in Table 9. If anything, the estimates improve. In all specifications, debt has a positive effect on populism, while budget surpluses and government spending have a negative effect. The coefficients are always significant. Furthermore, inflation now makes populism less likely.

[27] *Arellano and Bond, 1991; Arellano and Bover, 1995.*

Table 9: Dynamic panel data estimation of the determinants of populism. Significance levels computed using robust p-values.

	(1)	(2)	(3)	(4)
A. GMM estimates				
GDP per capita (-1) real PPP constant USD	4.1t	4.4t	0.6	0.9
Inflation (-1)	-0.9*	-0.9**	-0.2*	-0.3**
Output Gap (-1)	-0.4	-0.4	-0.1	-0.3
Unemployment Rate (-1)	0.6	-0.2	0.0	0.0
Gross debt/GDP ratio (-1)	0.7**	0.9***	0.2**	0.2**
Primary Gvt net lending/GDP (-1)	-2.6**	-1.9**	-0.8***	-0.7***
Gvt spending/GDP (-1)	-2.4*	-2.4***	-0.8***	-0.8***
Current account/GDP (-1)	0.0	-0.2	-0.1	-0.1
B. System estimates				
GDP per capita (-1) real PPP constant USD	4.0t	3.9*	0.8	1.0†
Inflation (-1)	-0.6†	-0.7**	-0.15*	-0.2**
Output Gap (-1)	-0.4	0.1	0.0	-0.1
Unemployment Rate (-1)	1.1**	0.7	0.2†	0.2
Gross debt/GDP ratio (-1)	0.3†	0.5***	0.1**	0.1**
Primary Gvt net lending/GDP (-1)	-1.5**	-1.2*	-0.6***	-0.5***
Gvt spending/GDP (-1)	-1.0*	-1.3***	-0.5***	-0.5***
Current account/GDP (-1)	0.4	0.2	0.0	0.0

The finding that greater government spending reduces the likelihood of a populist government is not so suprising, in light of the fact that debt and deficits are controlled for. Controlling for net lending, greater spending mean greater revenues, hence a greater fiscal capacity. This confirms my prediction above that greater fiscal capacity reduces the support for the populist. As for inflation, it may seem surprising that higher inflation reduces the likelihood of a populist government. To the extent that this is a sign of fiscal stress, one would expect the opposite. But the coefficients may be viewed as consistent with the theory if one considers inflation as a source of tax revenues which, everything else equal, reduces future levels of required fiscal adjustment, thus affecting the trade-off between a technocrat and a populist in favor of the former.

5. Conclusion

In this paper I have analyzed the connections between inequality of treatment of citizens, on the one hand, and policies that make fiscal crises more likely, on the other hand. Empirical evidence suggests that inequality of treatment is associated with higher levels of public debt, public spending, and public deficits, consistent with the theory – although the evidence seems less robust for spending and deficits than for debt. Furthermore, a “populist” platform (defined as more likely to discriminate between groups) is more likely to conquer power, the greater the degree of required fiscal adjustment. This is consistent with the prediction that fiscal crises alter the trade-off between electing a technocrat vs. a populist in favor of the latter. Finally, the theory sheds light on the reason why a priori unsound policies may nevertheless be implemented, as exemplified by the French 1981 reduction in the retirement age. □

APPENDIX

Table A1: Regression results with different equal treatment indicators. Same specification as Table 3. Dependent variable: Debt/GDP ratio.

Equal treatment in	Fiscal	Confl.	Corruption	Eq. tr.
Public jobs (A9043)	2.0*	0.1	-2.5***	-3.5**
Administration (A9042)	2.2**	0.0	-2.3***	-3.2*
Public health (A9041)	1.8†	0.1	-2.5***	-0.1
Education (A9040)	1.9*	0.1	-2.5***	-1.0

Table A2: Effect of indicators A9040–9043. Same specification as Table 4. Dependent variable: Government net primary surplus/GDP ratio.

Equal treatment in	Fiscal	Confl.	Corruption	Eq. tr.
Public jobs (A9043)	-0.1	0.7*	1.1*	1.51
Administration (A9042)	0.0	0.7*	1.1***	0.6
Public health (A9041)	0.1	0.7*	1.1**	0.5
Education (A9040)	0.1	0.7*	1.2***	0.7

Table A3: Effect of indicators A9040–9043. Same specification as Table 5. Dependent variable: Government spending/GDP.

Equal treatment in	Fiscal	Confl.	Corruption	Eq. tr.
Public jobs (A9043)	0.2	-0.6**	-0.5†	0.2
Administration (A9042)	0.2	-0.7**	-0.4†	0.3
Public health (A9041)	0.3	-0.6**	-0.4†	-0.2
Education (A9040)	0.3	-0.6**	-0.4†	-0.5

References

Acemoglu, D., Egorov G. and Sonin, K. (2013). "A political theory of populism," *Quarterly Journal of Economics*, pp. 771–805.

Afonso, A., Zartaloudis, S. and Papadopoulos, Y. (2015). "How party linkages shape austerity politics: clientelism and fiscal adjustment in Greece and Portugal during the eurozone crisis," *Journal of European Public Policy*, vol. 22(3), pp. 315–334.

Alesina, Alberto, and Allan Drazen (1991) "Why are stabilizations delayed?" *American Economic Review* 81(5):1170–1188.

Arellano, M., and S. Bond. 1991. "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations". *Review of Economic Studies* 58: 277–297.

Arellano, M., and O. Bover. 1995. "Another look at the instrumental variable estimation of error-components models". *Journal of Econometrics* 68: 29–51.

Banerjee, Abhijit and Rohni Pande (2007), "Parochial Politics: Ethnic Preferences and Politician Corruption", MIT working paper.

Beers, David and Jamshid Mavalwalla (2017), *Database of Sovereign Defaults*, 2017, Bank of Canada technical report #101, https://www.bankofcanada.ca/wp-content/uploads/2016/06/r101-revised-june2017.pdf

Conseil d'Orientation des Retraites (2004), *Retraites : les reformes en France et a l'etranger, le droit a l'information*, Paris: La Documentation Francaise.

http://www.cor-retraites.fr/IMG/pdf/doc-1705.pdf

Cukierman, Alex (2003), *Central Bank Strategy, Credibility cf Independence – Theory cf Evidence*, MIT Press

DARES (2016), "La syndicalisation en France", DARES Analyses, 25, https://www.fonction-publique.gouv.fr/files/files/statistiques/Hors_collection/Dares-2016-025.pdf

Dornbusch, R. and Edwards, S. (1991), eds. *The Macroeconomics of Populism in Latin America*, Chicago, IL: University of Chicago Press.

Ferrière, Axelle (2015), "Sovereign Default, Inequality, and Progressive Taxation", working paper, European University Institute.

Franck, R. and Rainer, I. (2012). "Does the leader's ethnicity matter? Ethnic favoritism, education and health in sub-Saharan Africa," *American Political Science Review*, vol. 106(2), pp. 294–325.

French Office of the Prime Minister (1991), *Livre Blanc sur les Retraites*, Paris: La Documentation Francaise, http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/134000051.pdf

Gerber, Jane (1980), *Jewish Society in Fez 1450–1700*: Studies in Communal and Economic Life, Leiden: E. J. Brill

Grim, B. J. and Finke, R. (2006). "International religion indexes: government regulation, government favoritism, and social regulation of religion," *Interdisciplinary Journal of Research on Religion*, 2 (1), pp. 1–40.

Le Boucher, Eric (2010), "Le scandale de la generation X", *Slate*.

Meltzer, A. and Richard, S. (1981). "A rational theory of the size of government," *Journal of Political Economy*, vol. 89(5), pp. 914–927.

Observatoire des retraites (2009), *Les chiffres de la retraite*, Paris: Observatoire des retraites, http://www.observatoire-retraites.org/uploads/tx_orpublications/L0RC5.pdf

Olken, Benjamin and Rohini Pande (2012), "Corruption in developing countries", MIT, mimeo

Persson, T., Roland, G. and Tabellini, G. (2000). "Comparative politics and public finance," *Journal of Political Economy*, vol. 108(6), pp. 1121–1161.

Persson, T. and Svensson, L. E. O. (1989). "Why a stubborn conservative would run a deficit: policy with time-inconsistent preferences," *Quarterly Journal of Economics*, vol. 104(2), pp. 325–345.

Robinson, J. and Verdier, T. (2013). "The political economy of clientelism," *Scandinavian Journal of Economics*, vol. 115(2), pp. 260–291.

Rouban, Luc (1999), "Les Attitudes politiques des fonctionnaires : vingt ans d'evolution", *Les cahiers du CEVIPOF*, Mai 1999, 24, http://www.cevipof.com/fichier/p_publication/446/publication_pdf_cahierducevipof24.pdf Saint-Paul, Gilles (2014), "La rigidite comme paradigme socio-economique", *L'Actualite Economique*, 90 (4), pp. 265–287

Saint-Paul, Gilles, Davide Ticchi and Andrea Vindigni (2017), "Engineering Crises: Favoritism and Strategic Fiscal Indiscipline", CEPR Working Paper

Tabellini, G. and Alesina A. (1990). "Voting on the budget deficit," *American Economic Review*, vol. 80, pp. 37–49.

A researcher's guide to the Swedish compulsory school reform*

Helena Holmlund
IFAU – Institute for the Evaluation of Labour Market and Education Policy
UCLS
Ifo Institute

Abstract

This paper demonstrates how a natural experiment in education can be used to estimate causal effects. The Swedish compulsory school reform extended basic education gradually across cohorts and municipalities, allowing for a difference-in-differences analysis. The paper summarizes the literature using this reform and shows that it provided individuals from low socio-economic backgrounds with better opportunities in life. Not only did they attain higher levels of education – they also earned higher earnings, were less likely to participate in crime, and more likely to run for office.

JEL-codes: I24; I26; I28

* *I wish to thank two anonymous referees, Anders Björklund, Mikael Lindahl, Olmo Silva and participants at the CEE group meeting at LSE for helpful comments. Thanks also to Anders Björklund, Valter Hultén and Mikael Lindahl for help with the coding of the reform. Parts of the research has been financed by the Swedish Council for Working Life and Social Research, and their support is gratefully acknowledged.* *IFAU, Kyrkogårdsgatan 6, 751 20 Uppsala, Sweden. Email: helena.holmlund@ifau.uu.se.*

1. Introduction

In the 25 years following the Second World War, many western European countries undertook major educational reforms with the main purpose of extending compulsory education. The Nordic countries, continental Europe and the United Kingdom, with different traditions of education policy, were all part of this widespread expansion (Viarengo 2007). The strong economic growth in the post war era created a demand for a higher skilled workforce, and in some countries, for example Sweden, there was a strong push for reforming the education system in order to increase equality of opportunity. The European experience was also a reflection of an earlier development in the United States, where compulsory school attendance and child labour laws were enacted throughout the states in the early decades of the 20th century. These compulsory schooling reforms have spurred an enormous interest in applied economics; reforms that because of their design offer the promise of estimating causal effects of extending basic education on a range of outcomes such as earnings, health, crime, and intergenerational effects on the education of offspring in the next generation.

Sweden extended compulsory education gradually across the country, starting in the late 1940s. The reform was implemented in different municipalities at different points in time, meaning that for a given birth cohort some individuals went through the old two-tier selective system where basic education ended after 7 or 8 years, and others went through the new school, comprising of one or two more years in a comprehensive system. Similar gradual expansion paths were also adopted in the other Nordic countries, and the design of these reforms has resulted in a large number of studies that exploit variation across regions and over time as a source of quasi-random variation in both length of compulsory schooling and/or educational tracking (seminal papers are Meghir and Palme 2005 for Sweden, Black et al. 2005 for Norway, and Pekkarinen et al. 2006 for Finland).[1] Closely related to these papers is the U.S. literature, which with a similar methodological approach studies compulsory school leaving ages across U.S. states (see for example Lochner and Moretti 2004 and Llleras-Muney 2005).

Today, 15 years after Meghir and Palme's paper was first published, there is a large body of research based on the Swedish compulsory school reform; the combined use of Swedish register data and difference-in-differences analysis exploiting the gradual nature of reform implementation has given rise to many good publications. This paper summarizes the existing papers up to date and discusses the insights from the literature in light of the political background of the reform. In addition, the paper offers a documentation of the reform data collection previously presented in Holmlund (2007), which has been to the benefit of many of the papers cited here. More specifically, I discuss how information on reform implementation can be linked to different sources of data, I present balancing tests to examine if reform exposure is conditionally correlated with individuals' observed characteristics and report estimates of the effect of the reform on educational outcomes. Using two independent measures of reform assignment, I also run IV regressions that take into account measurement error bias and bound the reform "first stage" estimates.

So, what have we learned? First, as I show in the empirical part of this paper, the difference-in-difference approach is successful since both balancing tests and tests for pre-reform parallel trends suggest that the underlying assumption of parallel trends is satisfied. Second, the reform had larger effects on educational attainment among individuals from lower SES backgrounds, thus contributing to intergenerational mobility. Finally, summarizing the literature today, I conclude that an impressive number of good publications have emerged from combining register data and the reform. We have learned that the extension of basic education has had positive effects on cognitive skills – but that non-cognitive characteristics among children from high SES backgrounds might have been negatively affected. We can also conclude that the reform reduced criminal involvement, and that individuals from working class backgrounds became more likely to engage politically by running for office. Studies on mortality and health show diverging results, but the overall impression is that there are no effects on health outcomes. However, the reform seems to have affected financial decision-making. Lastly, the reform has proven to affect outcomes in the next generation: there are spillover effects to the skills of targeted individuals' children.

[1] *The Danish school reform was not subject to gradual implementation and as such has not been used as widely to estimate causal effects. See Arendt (2008) for a description of Danish reforms and an application.*

The remainder of the paper unfolds as follows: section 2 presents a general discussion on estimating returns to education, section 3 covers institutional details regarding the Swedish compulsory school reform, and sections 4 and 5 detail reform assignment and data sources. Section 6 presents balancing tests and estimates of reform effects on educational outcomes. Section 7 summarizes the literature using the Swedish reform up to date, and finally, section 8 offers conclusions. For more detail, Appendix A contains a documentation of sources used to determine reform status.

2. Returns to education

There is a long tradition in empirical labour economics of studying returns to education. The early literature focused on the pecuniary returns – the percentage wage gain from one more year of schooling. Later, this literature was extended to focus on non-pecuniary returns to education, such as effects of education on health, crime and fertility. These studies all face the same challenge, that is, how to control for unobserved ability or other unobserved factors that are correlated with education and also with the outcome of interest. For example, is the wage premium from schooling truly an effect of education in itself, or does it purely reflect the fact that more able or more motivated workers, earning higher wages, also choose a higher level of education? Moreover, estimating the effect of education on health, how do we take into account that individuals in poor health might not find it worthwhile investing in education because their health condition implies lower returns to their investment (an example of reverse causality)? Health and education can also be correlated because of discount rate bias: individuals with a high discount rate might invest neither in human capital nor health. And similarly, with fertility as with many other potential outcomes, how can we as researchers control for unobserved preferences that jointly determine both education and fertility outcomes?

The remedy in many studies of these issues has been either to control for ability by using samples of identical twins (Ashenfelter and Krueger 1994, Behrman and Rosenzweig 1999), or to make use of some exogenous source of variation in education, typically in the form of a natural experiment. Natural experiments offer variation in some treatment, in this case compulsory education legislation, for individuals that otherwise can be assumed to be identical; natural experiments allow us to come around the problem of education being correlated with individual characteristics such as ability or motivation. There are a number of well-known examples, apart from the Nordic papers cited above, and the Swedish papers cited in section 7. Angrist and Krueger (1991) use the fact that a fixed school-leaving age in the US allows students to drop out from school earlier if they are born early in the year, that is, the length of compulsory education varies with month of birth.[2] Acemoglu and Angrist (2000), Lochner and Moretti (2004), Lleras-Muney (2005), Oreopoulos et al. (2006) all use variation across US states in compulsory schooling laws: variation across states is introduced by both compulsory attendance and child labour laws. Currie and Moretti (2003) account for endogeneity of schooling by using variation induced by college openings. Chevalier (2004) studies a reform in the UK and Maurin and McNally (2005) use variation in schooling introduced by the 1968 revolts in Paris. The outcomes in the studies mentioned above all range from the pecuniary return to education, to effects on mortality, crime, birth outcomes and the education of the children (the intergenerational effect of education).

3. The Swedish reform

The Swedish educational reform is carefully described in the work by Marklund (1980, 1981). Detailed information can also be found in a report by the National Board of Education (1960). The following brief description builds on these sources, which are recommended for further details on the topic.

Prior to the school reform, pupils in Sweden went through grades 1 to 4 or 1 to 6 in a common school (*folkskolan*). In either fourth or sixth grade, more able students were selected (based on past performance) for the five or three/four-year

[2] Later research has invalidated this approach by showing that month of birth in itself is directly related to educational outcomes, through e.g. maturity at school starting-age, or through the benefits of relative age in the classroom (see e.g. Bound and Jaeger 1996 for a critical discussion).

long junior-secondary school (*realskolan*). Remaining students stayed in the common school until compulsory education was completed. In most cases, compulsory education comprised seven years, but in some municipalities, mainly the big cities, the minimum was eight years. The system resembled the traditional European model with early selection, parallel school forms and a small tertiary sector (Erikson and Jonsson 1996).

In 1946, the social-democratic government appointed a parliamentary committee (*1946 års skolkommission*) which was given the task to analyze the Swedish school system and to develop proposals and guiding principles for a non-selective compulsory school. The main purpose with such a change was to postpone the tracking decision to higher grades, in an effort to increase equality of opportunity.[3] Two years after the appointment of the committee, in 1948, the committee released its proposals. The main suggestion was to introduce a nine-year compulsory school, where pupils were kept together in common classes longer than in the earlier school system. As a compromise between the opponents of early tracking and its advocates, the committee proposed tracking in 9th grade; pupils would follow either a vocational track, a general track, or a theoretical track preparing for upper-secondary school. The 9th grade streaming was later abandoned in favor of a completely comprehensive system.

Erikson and Jonsson (1996) argue that more than in most other Western countries, school reforms in Sweden have been characterized by the specific aim of reducing social and educational inequalities. Early selection was considered a hurdle for children from low socioeconomic backgrounds to access secondary education, and with a comprehensive system the idea was to provide equal opportunities for all children, regardless of family background. Naturally, the reform also had its critics and the question of tracking became the key controversy around the reform, with the right-wing party in opposition of late selection.

To evaluate the appropriateness and whether the proposed nine-year comprehensive school would serve its purpose, in 1949 the committee suggested that an "experiment" would take place, where during an assessment period some municipalities and schools would implement the new school system such that the results could be scrutinized before further decisions were made.

The assessment programme came to start in 1949/1950, this year under the surveillance of the parliamentary committee. In 1950, the Swedish parliament committed to the introduction of a nine-year comprehensive school and approved of the idea of a trial period at the outset of the reform. When the formal decision was made in 1950, the National Board of Education (*Skolöverstyrelsen*) took over the administration of the reform.

The new comprehensive school was to be introduced throughout a whole municipality, or in certain schools within a municipality. Following the 1948 proposal of the parliamentary committee, a number of municipalities had declared interest in reforming their comprehensive schools. For this reason, 264 municipalities (out of around 1000) were asked if they were willing to introduce the nine-year school immediately or within a few years. The municipalities that were approached had either shown interest in the reform or expanded their junior secondary school to four years. 144 municipalities showed interest in the reform. 14 municipalities were selected for the first year of the assessment (1949/50), all of those were required to have an eight-year comprehensive school already.

The following years, the National Board of Education continued with the implementation of the reform. Year by year, more municipalities joined the reform assessment programme. Municipalities that wanted to take part in the reform were asked to report on their population growth, on the local demand for education, tax revenues and local school situation. For example, the availability of teachers, the number of required teachers for the nine-year comprehensive school, and the available school premises were explored. The National Board of Education took these municipality characteristics into account when deciding on their participation. In general, implementation of the reform started in grades 1 and 5, the following year covering grades 1, 2, 5 and 6 and so on. From 1958 the reform was introduced in grades 1–5 already from the starting year.

[3] *The large baby boom cohorts that passed through the education system during this period, and higher demand for junior secondary schooling overall, are also likely drivers behind reforming the education system.*

Apart from extending compulsory education from seven (or in some cases eight) years to nine years, and to postpone tracking, the educational reform was also pedagogical and affected the curriculum somewhat. The main change of the curriculum was that English was introduced in 5th grade in the new comprehensive school, while this was not necessarily a compulsory subject in the old school system. The school starting age was set at the year the child turned seven in both the old system and the new comprehensive school.

The assessment period was also accompanied by financial support to families and to municipalities that implemented the reform. A universal child allowance was introduced in 1948 and implied support for children until the age of 16. In reform municipalities, a means-tested scholarship compensated families for foregone earnings from keeping their children longer in school. Municipalities were compensated for the increased costs following the expansion of education. The state provided funds targeted at the new comprehensive school, one example is complete funding of teacher salaries for grades 7–9, in the years 1952–1955.

In 1962, the parliament came to a final decision to permanently introduce the nine-year school throughout the country. At this point, the implementation came to be a matter for each municipality; by 1969 they were obliged to have the new comprehensive school running. Since the timing was much in the hands of each municipality, the implementation was far from a randomized experiment, but nevertheless provides a source of variation in schooling laws that may be fruitfully explored by the empirical researcher.

4. Linking individuals to treatment

Since the educational reform provides a potential source of quasi-random variation in education, I take a closer look at the available data. For the quantitative researcher, knowledge about which municipalities implemented the reform, and which birth cohorts were affected, is of particular importance. Below, I list three different data sources available to study reform effects on individual outcomes.

1. The IS data

It is possible to use the IS (individual statistics) data, from the Institute of Education at Gothenburg University (Härnqvist 2000). The data stem from surveys, conducted in 6th grade, of around 10 percent of the cohorts born in 1948 and 1953. When these data were collected, information on type of school (the old *folkskola* or the new nine-year comprehensive school) that each individual attended was recorded, based on information provided by the local school. Register information on adult earnings and other register-based information can be matched to individuals in the data. This is the data set explored in Meghir and Palme's (2005) work on the Swedish compulsory school reform.

2. The Swedish Level of Living Survey

The Swedish Level of Living Surveys, based on random samples of the Swedish adult population, have been conducted in 1968, 1974, 1981, 1991, 2000 and 2010 (Erikson and Åberg 1984). The surveys ask specifically whether an individual went through the old system or the new nine-year school. These data have been used by Jasmina Spasojevic (2010) in her work on the effects of education on health.

3. Register data from Statistics Sweden

The Swedish administrative registers do not contain information on whether individuals in the affected cohorts went through the old or the new school system. With help from other sources (described in Appendix A) it is however possible to deduct when and for which grades each municipality introduced the new comprehensive school and based on this information one can assign reform status to the individuals in a data set extracted from registers. Censuses can be used to track in which municipality an individual lived at the time of compulsory education. With this information it is possible to attach a reform indicator to each individual based on year of birth and municipality of residence, maintaining the assumption that individuals are in the right grade according to their age. In some cases, it is also necessary with more detailed information on in which parish or school district the individual went to school, since the reform was sometimes introduced in parts of a municipality in different years. Any given dataset with information on birth year and municipality/parish of residence can assign reform participation to the cohorts that were subject to the education reform.

There are two possible ways to construct a reform coding that can be matched to individual-level register data. In the remainder of the paper I will label them coding 1 (based on documentation) and coding 2 (deduced from register data).

The sources of information necessary to construct coding 1 of the reform implementation are the following:

- Marklund (1981) and the National Board of Education (1953–1962). These sources document the assessment programme when the reform was gradually introduced across the country, and they include lists of which municipalities implemented the reform each year. In the latter publication it is also possible to see which grades that were affected in a particular municipality. These sources cover the assessment period and only allow coding of the cohorts born 1938–1949.

- The Educational Bureau (Undervisningsbyrån) (1960–1964) and Statistics Sweden (1968–1969). From municipality-level tables of the number of pupils in each grade in the old and new school system, it is possible to deduct when the reform was implemented, and the remaining cohorts can be coded.

Register data sets with large sample sizes allow for an alternative procedure to assign reform status. Coding 2 is simply obtained by splitting the sample, using the first part of the sample to identify treatment status by empirically observing when the minimum level of education (by municipality/birth year) jumps up from folkskola (the old compulsory minimum) to grundskola (the new minimum), and using the second part of the sample to estimate effects of treatment on outcomes.

Appendix A explains in detail how the different sources have been used to create coding 1 and coding 2, and also highlights some of the difficulties relating to the coding of some municipalities, where the reform was not implemented universally at one point in time. In the remainder of this paper I discuss and demonstrate the reform, and explore several important data issues, using a data set compiled from Swedish registers. Coding 1 is available from the author for researchers who wish to use it.

5. Data

The empirical analyses are based on data from Swedish administrative registers, available for researchers through Statistics Sweden. The population of interest is defined as cohorts born in Sweden between 1945 and 1955.[4] By linking several registers and censuses through personal identifiers, individuals can be linked to their parents, and to information on both family background and long-term outcomes.

Family background. Family background is characterized by father's education level and father's earnings. Father's education level is derived from the 1970 census and is defined as a dummy for high education, which represents any education above compulsory level. Father's earnings is a measure of average earnings over the years 1968, 1971 and 1973, percentile ranked within father's own birth cohort. These variables are used in a balancing test to investigate whether treatment is correlated with pre-determined characteristics.

[4] *This population is chosen because the birth years span the vast part of the new school expansion, and because it is possible to assign reform status to these cohorts.*

Educational outcomes. I study educational outcomes through information on highest completed level of education in the education register from 1995. Highest level of education has been translated into years of education by assigning the expected number of years to each level.[5] I also study the probability of attaining any post-compulsory education beyond the new compulsory minimum. To arrive at this measure, I use observed levels in the education registers, which implies education at secondary or tertiary level.

Assigning reform status. Reform status is assigned to individuals based on birth year and municipality of residence. The cleanest way to assign treatment is to use information on pre-treatment location, which in my data is available as municipality of birth, derived from the parish where the child was registered at birth. However, for cohorts born until 1946, the parish of birth that was reported refers to the location of the hospital in which they were born (Skatteverket 2007). At this time, most births did take place out of the home, and a majority of all municipalities did not have their own maternity ward. Thus, the information on parish of birth cannot be used to identify treatment status for early cohorts in the sample.

The alternative (used in this paper) is to extract census information on home municipality in 1960 and 1965, approximately at age 10–15. This arguably assigns reform status to individuals based on where they lived while going to school but could be considered as endogenous to reform exposure if families move in response to the reform.

After excluding individuals with missing observations on either education or municipality of residence in 1960/1965, I arrive at 1,182,063 observations. Treatment is identified for 91 percent (using coding 1 based on documentation). Descriptive statistics of the sample are presented in Figure 1.

Table 1. Summary statistics

Variable	Nr of obs	Mean	Std. Dev.
Years of schooling	1,021,996	11.46	2.52
Any post-compulsory schooling	1,021,996	0.75	0.43
Reform exposure	1,021,996	0.48	0.50
Woman	1,021,996	0.49	0.50
High educated father	767,567	0.31	0.46
Income rank father	966,031	52.31	27.86
Birth year	1,021,996	1949.89	3.18

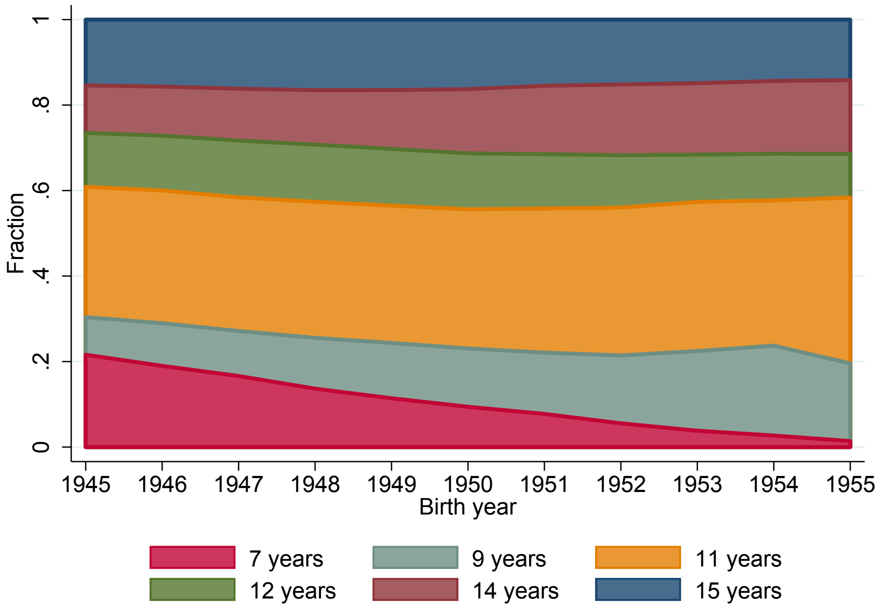
[5] *Fisher et al. (2018) show that deriving years of schooling from education levels can understate reform effects, since some compliers to the reform will be coded with "too much" education using this method. Using alternative data sources, they find that the reform effect on years of schooling is 76 percent larger when taking into account compliers that are not observed when using education levels to identify years of schooling.*

6. An analysis of the Swedish compulsory school reform

The compulsory schooling reform affected cohorts whose education levels were on the rise across the board – both at the low and the high end of the distribution. Besides the focus on raising the minimum level, the study grant system for higher education was reformed in 1965, and several new tertiary education institutions were opened in the 1960s and 70s. Accordingly, Figure 1 shows that the number of years of schooling were increasing, both as a consequence of a higher share of individuals attaining 9 years, and higher shares both at upper-secondary (11–12 years) and post-secondary (14 + years) levels.

Figure 2 shows the share of individuals treated by the reform, for the two different sources of reform assignment described in section 4. The different coding schemes follow each other closely in terms of the share of individuals in each birth cohort that is affected.[6] It also shows that for the cohorts depicted in the figure, the increase over time is fairly linear. When reaching the 1955 cohort, almost 100 percent of individuals have been assigned treatment according to the two different coding schemes.

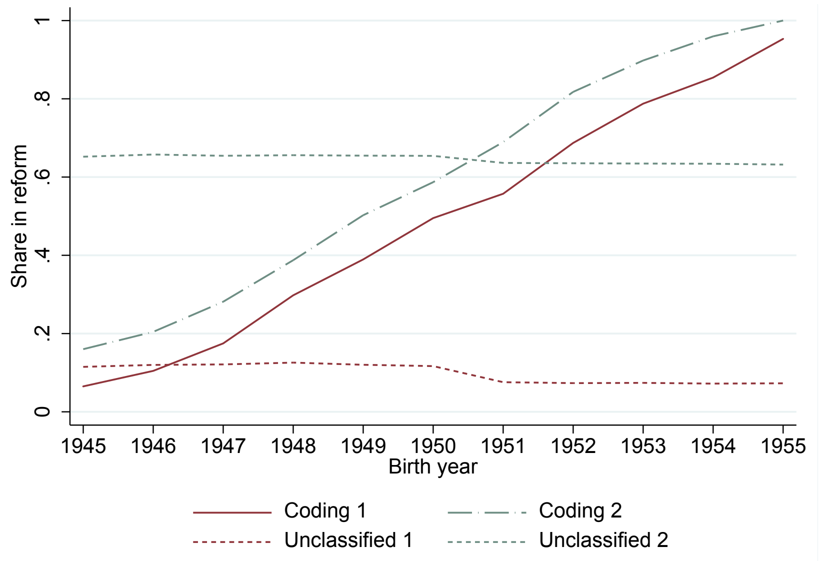
Figure 1. Share of population at different levels of schooling



Note: Years of schooling assigned according to highest level in the 1995 education register. 15 years refers to 15 years or more.

[6] *The large share with unclassified treatment status for coding 2 is explained by the fact that 60 percent of the sample has been used to identify treatment status by municipality/birth year, and 40 percent of the sample is used in estimation.*

Figure 2. Share of population treated by the reform



Note: Coding 1 refers to assignment based on documentation, coding 2 refers to reform assignment based on observed minimum levels by cohort and municipality (using 60 percent of the sample to identify treatment status, 40 percent for descriptives and estimation). Reform status has been assigned to individuals based on birth cohort and municipality of residence observed in the censuses 1960 (cohorts 1945–1950) and 1965 (cohorts 1951–1955).

6.1 Reform effects on schooling outcomes

The gradual implementation implies that a staggered difference-in-differences design is a natural starting point for estimating effects of reform exposure on outcomes. Consider the following baseline specification:

$$y_{icm} = \alpha + \beta T_{cm} + \gamma_c + \delta_m + \varepsilon_{icm} \quad (1)$$

where y_{icm} is the outcome of individual i, belonging to cohort c and municipality m. T_{cm} is a treatment indicator that takes the value of 1 if a cohort and municipality is treated, and 0 otherwise. The equation also includes cohort- and municipality-specific fixed effects (γ_c and δ_m). The parameter of interest is β, which gives us the ITT (intention-to-treat) parameter of reform exposure on the outcome of interest. In essence, the method compares the difference in outcomes over time – before and after treatment in treated regions – to the same time difference in untreated regions or in regions treated at a different point in time. The staggered differences-in-differences specification has been common in many empirical applications and reform evaluations but has recently received attention from a methodological point of view: new econometrics papers demonstrate that the method can lead to biased estimates if heterogeneous treatment effects are present. I return to this topic in section 6.6 where I briefly discuss the methodological problems in more detail and perform a sensitivity tests suggested in the recent literature.

The difference-in-differences design relies on the identifying assumption of parallel trends between treated and control units in the absence of any intervention. This assumption cannot be tested, but there are two standard ways to assess the credibility of the assumption. First, a balancing test regressing pre-determined characteristics on treatment and region- and cohort controls sheds light on whether there are differential compositional trends in terms of observed characteristics in treated and control regions. Second, a specification estimating pre-reform effects illustrates whether the assumption holds in the pre-reform period. Before turning to the main effects of the reform on educational outcomes, I present these two analyses in turn.

6.2 Balancing tests

Table 2 begins by presenting balancing tests relating reform exposure to two pre-determined characteristics capturing the socioeconomic background of individuals in the population. The first panel presents results using coding 1, while the second panel corresponds to coding 2. First, columns 1 and 2 present the relationship between treatment and father's education (a dummy for any post-compulsory education). Column 1 presents the baseline specification, while column 2 presents a specification including linear time trends interacted with indicators for implementation year (a simplification of a specification including municipality-specific linear trends). Columns 3 and 4 present the corresponding estimates for father's income percentile. The overall impression is that the difference-in-differences specification is successful in handling any form of sorting into treatment based on observable characteristics; the coefficients are precisely estimated and close to zero.

Table 2. Balancing tests

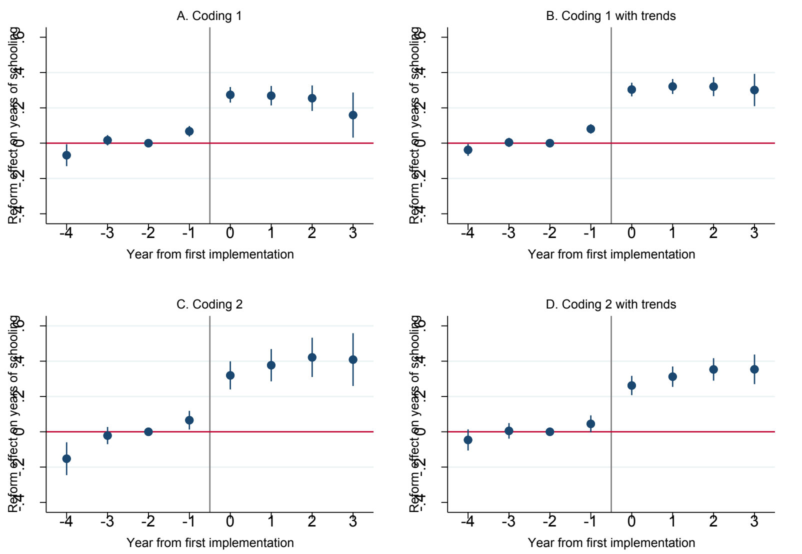
	(1)	(2)	(3)	(4)
	Father high education		*Father's income percentile*	
	A. Balancing test reform coding 1			
Reform exposure coding 1	-0.001	0.002	0.022	0.085
	(0.002)	(0.002)	(0.113)	(0.116)
Observations	767,567	767,567	966,031	966,031
R-squared	0.002	0.002	0.000	0.000
Number of municipalities	1,020	1,020	1,020	1,020
Outcome mean	0.310	0.310	52.32	52.32
	B. Balancing test reform coding 2			
Reform exposure coding 2	0.005	-0.004	0.587	-0.095
	(0.007)	(0.003)	(0.432)	(0.163)
Observations	302,131	302,131	380,059	380,059
R-squared	0.002	0.002	0.000	0.000
Number of municipalities	986	986	986	986
Outcome mean	0.312	0.312	52.38	52.38
Muni f.e.	Yes	Yes	Yes	Yes
Cohort f.e.	Yes	Yes	Yes	Yes
Implement-spec trends	No	Yes	No	Yes

*Note: Robust standard errors in parentheses clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1. The sample size is reduced in panel B since coding 2 is defined for a subset of the population (here 40 percent).*

6.3 Event-study analysis

The event-study design expands the difference-in-differences specification by introducing treatment dummies interacted with time since first treatment year. By doing so, we are able to both address the pre-reform parallel trends assumption and depict the time pattern of treatment effects in one regression. Figure 3 presents the results, where t-2 is the excluded time category. The figures on the left show the baseline specification, and those on the right the extended specification including trends. The four sub-figures show a similar pattern: there is a clear reform effect starting in the first implementation year (year 0), of about 0.3 years of schooling. The figures also show small positive and significant pre-reform estimates in year t-1. One possible explanation to this is the ambiguity in terms of documented starting year, where some municipalities phased in the reform over two years and as such a fraction of students were actually treated in the year before the first indicated start year. Moreover, in relation to the baseline year, treated areas have negative estimates in t-4, i.e., treated areas had a slower growth in education compared to non-treated areas four years before the reform was implemented. When controlling for implementation-year specific trends, these negative estimates move closer to zero. It is also worthwhile pointing out that in the event-study analysis, coefficients estimated a few years before (after) first implementation year are identified using subsamples of late (early) implementers. In terms of the parallel trends assumption, using the specification with trends seems to be preferable, and acknowledging that the t-1 effect is expected the diff-in-diff specification performs well. In fact, many studies have chosen to drop t-1 from the analysis to arrive at a sharper pre and post distinction. I follow this procedure in the remainder of the paper.

Figure 3. Event-study analysis of reform exposure on years of schooling



Note: Event-study estimates of reform exposure interacted with time since first treated cohort. t-2 is the left-out time category. Spikes show 95 percent confidence intervals, clustered at the municipality level.

6.4 Reform effects on years of schooling and post-compulsory schooling

The vast majority of studies of effects of the compulsory schooling reform take an interest in secondary outcomes beyond education, that come as a result of being exposed to the reform. Effects on e.g. income, health or crime can operate through longer education, but other direct channels are also possible. Quality of education, peer composition, tracking and changes to the curriculum could directly affect future outcomes. The effect of treatment, i.e. reform exposure, on length of education however serves as a proxy for a "first stage" in terms of schooling – it tells us something about how large the intervention is, for whom it bites the most, and helps to generate hypotheses regarding externalities on other outcomes. One key question also relates to whether the reform had effects beyond the new compulsory minimum.

Table 3, panel A, presents baseline estimates of the reform on years of education. In this table, results are based on coding 1, and the last pre-reform cohort has been excluded from the regressions. Column 1 shows the estimate excluding trends and indicates that the reform increased years of schooling by 0.28 years. Column 2 including trends gives an estimate of 0.31 years, or 12 percent of the standard deviation of years of schooling. Having established that the reform implied an increase in years of education, it is worth taking a closer look at the dynamics of this effect. Did the reform have differential impacts depending on socioeconomic background? Did it solely add two years of education for those at the bottom of the distribution, or did it also induce a shift beyond the new compulsory minimum? Such a spill-over effect could occur for example if the pre-reform early selection was unfavorable to talented children from disadvantaged backgrounds, which in the new school system possibly got the chance to move on further in the education system.

Columns 3 and 4 show effects on years of schooling for children with high and low educated fathers, respectively. As expected, the effect is much larger (about three times as large) among children with low educated fathers. Most students with highly educated fathers already attended education beyond compulsory level before the reform, implying that the reform had less bite in this group.

Table 3. Reform effects on educational outcomes

	(1)	(2)	(3)	(4)
	All	*All*	*High edu father*	*Low edu father*
		A. Years of schooling		
Reform exposure coding 1	0.280***	0.309***	0.115***	0.361***
	(0.026)	(0.016)	(0.020)	(0.019)
R-squared	0.013	0.014	0.002	0.021
Control group outcome mean	11.09	11.09	12.43	10.69
		B. 2 year upper secondary or more		
Reform exposure coding 1	0.013***	0.018***	0.009***	0.018***
	(0.003)	(0.002)	(0.003)	(0.003)
R-squared	0.014	0.014	0.004	0.018
Control group outcome mean	0.708	0.708	0.867	0.665
Number of municipalities	1,020	1,020	1,020	1,020
Muni f.e.	Yes	Yes	Yes	Yes
Cohort f.e.	Yes	Yes	Yes	Yes
Implement-spec trends	No	Yes	Yes	Yes
Observations	939,130	939,130	219,104	485,595

*Note: Robust standard errors in parentheses clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1*

Panel B shows estimates of the reform on a dummy indicating completion of two-year upper-secondary school or more.[7] The results indicate that a small spill-over effect is present: a 0.013–0.018 higher probability to complete two years of upper-secondary education or more. As a point of comparison, the overall probability to attend two-year secondary school or any education beyond that is 0.71. Although a small effect, this indicates that reform exposure did push some individuals to higher levels beyond the expected minimum. The spill-over effect is twice as large when comparing children with low and high educated fathers (columns 3 and 4), which indicates that the shift to a comprehensive system served one of its purposes – to increase equality of opportunity.

6.5 Measurement error bias and bounding the estimates

Reform assignment comes with measurement error – in some cases it is an approximation of the starting year of the reform. As described above, some municipalities kept parallel school systems which means that there is no possibility to find a clear-cut starting point. Hence, the coding of the reform does in some cases represent an average or the majority in a given municipality and birth cohort, which will introduce measurement error in the reform indicator. Another aspect is that even though implementation might have been extensive, there was room for single individuals to apply for an exemption. We also need to assume that pupils are in the expected grade according to their age; if grade repetition or skipping a grade was a prevalent phenomenon among the affected cohorts, this is also one source of measurement error to keep in mind. To better understand the consequences of measurement error in regression analysis based on the reform, I now turn to examining the quality of reform indicators.

As a starting point to a reliability analysis of the reform coding, I acknowledge that since reform participation is a binary indicator variable, the measurement error is not classical. That is, the measurement error is correlated with the true underlying variable (Aigner 1973). The formula describing attenuation bias in the case of classical measurement error must now be modified to represent the case of non-classical measurement error.

We would like to estimate reform effects on an outcome y in the following way:

(2) $$y = \alpha + \beta r^* + \varepsilon$$

Where r_{icm}^* denotes the true (unobserved) reform status of an individual i., belonging to cohort c., going to school in municipality m. In the data we observe two measures of the reform measured with error (omitting the subscripts for simplicity and following the notation in Kane et al. 1999):

(4) $$E(r_1|r^*, r_2, y) = \pi_{10} + \pi_{11} r^* \qquad \text{Coding 1}$$

(4) $$E(r_2|r^*, r_1, y) = \pi_{20} + \pi_{21} r^* \qquad \text{Coding 2}$$

Given the true reform participation r^*, I assume that the observed variables r_1 and r_2 are independent of each other and of y. In order for the measurement error to be classical, the further assumptions $\pi_{11} = \pi_{21} = 1$, and $\pi_{10} = \pi_{20} = 0$ must be satisfied. With a binary indicator variable these assumptions do not hold and the measurement error is correlated with the true underlying variable. We have that $\pi_{11} < 1, \pi_{21} < 1$ and $\pi_{10} > 0, \pi_{20} > 0$.

Following Aigner (1973) and Kane et al. (1999) the probability limit of β in the case of measurement error in a binary variable can be derived as follows:

(5) $$p\lim \beta_{OLS} = \beta[1 - P(r^* = 1|r = 0) - P(r^* = 0|r = 1)]$$

[7] *Two-year secondary school is the lowest post-compulsory degree and refers to vocational post-compulsory education.*

Just as in the case of classical measurement error, the OLS estimate is biased towards zero and the estimated effect is attenuated. In the case of classical measurement error, a standard remedy to inconsistencies in OLS parameters has been to use an instrumental variables strategy. With two independent measures of the variable of interest, two-stages-least squares when one measure is used as an instrument for the other produces consistent coefficients. When measurement error is non-classical, however, an IV strategy is not likely to produce consistent estimates. Nevertheless, an IV estimate can be informative, since it turns out that with non-classical measurement error it will be upward biased (Kane et al. 1999):

$$(6) \qquad p\lim \beta_{2SLS} = \beta \frac{1}{\pi_{11}}$$

where r_1 has been instrumented in the first stage using r_2 as an instrument. We see that only in the case of classical measurement error $(\pi_{11} = 1)$ 2SLS produces consistent estimates, and with measurement error in the categorical variable $(\pi_{11} < 1)$ β_{2SLS} will be upward biased.

Thus, with measurement error in the binary indicator variable for reform participation, it turns out that both the OLS and the IV estimate (using two reform codings and instrumenting one with the other) are inconsistent, one downwards and the other upwards. Therefore, the two estimates provide a lower and an upper bound for the true parameter, and we are able to narrow down the range of possible true effects. Taking the (downward biased) estimates presented in Table 3 as a benchmark, Table 4 presents the corresponding IV estimates, where coding 2 has been used as an instrument for coding 1. Columns 1 and 3 present the OLS specifications including trends, and columns 2 and 4 the corresponding IV-2SLS estimates. The differences between the OLS and the IV estimates are relatively large; the latter are about twice as large as the OLS estimates. The take-away from this analysis is therefore that the true parameter estimate of the effect of the reform on years of schooling is within the range 0.32–0.58 years. The effect on post-compulsory education can be bounded to an increase in the range 0.021–0.046 percentage points.

Table 4. OLS and IV estimates bounding the effect of the school reform on educational outcomes

	(1)	(2)	(3)	(4)
	OLS	IV	OLS	IV
	Years of schooling		**2 year upper secondary or more**	
Reform exposure coding 1	0.316***	0.580***	0.021***	0.046***
	(0.023)	(0.082)	(0.004)	(0.010)
Observations	368,480	368,480	368,480	368,480
R-squared	0.014		0.014	
Number municipalities	983	983	983	983
Muni f.e.	Yes	Yes	Yes	Yes
Cohort f.e.	Yes	Yes	Yes	Yes
Implement-spec trends	Yes	Yes	Yes	Yes
Control group outcome mean	11.09	11.09	0.708	0.708

*Note: Robust standard errors in parentheses clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1. In columns 2 and 4, coding 1 is instrumented with coding 2. The sample size is reduced since coding 2 is defined for a subset of the population (here 40 percent).*

6.6 Can we trust staggered differences-in-differences? A robustness analysis

A number of recent papers highlight problems with a pooled difference-in-difference estimator when units are treated at different points in time, and additionally when all units are treated at some point. de Chaisemartin and D'Hautfoeuille (2020) and Goodman-Bacon (2018) show that linear regression with time- and group-fixed effects estimate weighted sums of average treatment effects in each group and time period. Treatment dummies assume that the treatment effect is constant while event-time treatment effect heterogeneity is a possibility. Borusyak and Jaravel (2018) show that short-run effects are overweighed, and long-run effects are negatively weighted. Since weights are not equal across all LATES that feed into the average, treatment effect heterogeneity might lead to a biased average effect. de Chaisemartin and D'Hautfoeuille show that weights might even be negative and that the estimate therefore can be of the wrong sign. Negative weights appear because the phase-in of treatment implies that "late" time periods can have a predicted treatment probability above 1.

To begin with, the event-study graphs in section 6.2 above show relative stability of treatment effects over time since treatment, which implies that even if long-run effects have lower weight, the pooled estimate is likely to be representative of the average effect.

Second, de Chaisemartin and D'Hautfoeuille (2020) provide a stata package to test the presence of negative weights in the context of group- and time-fixed effects models. I apply this test to the setting in this paper, i.e. the staggered difference-in-differences model estimating effects of the compulsory school reform. Using the package that checks for negative weights [twowayfeweights], I find that for all 8279 LATE:s for each combination of group and time period in my data, weights are positive. As such, the traditional difference-in-differences estimate is likely to be representative of the true average in this setting, even though treatment effect heterogeneity at the group level cannot be ruled out and might also lead to a biased average effect.

7. Summing up 15 years of research – Evidence based on the Swedish compulsory school reform

This section summarizes the papers exploiting the Swedish reform up to date. This turns out to be an impressive list of publications, published in many good journals. The papers can roughly be categorized by type of outcome; skills, income, health, crime, political decision making, financial literacy and intergenerational effects. Below I present the papers using this categorization. Some papers estimate the reduced form of the reform, while others use it as an instrument for years of education. I will return to this distinction below.

Skills. One of the basic questions in labour economics asks whether returns to education reflect higher skills and productivity that come with higher education, or if education purely works as a signal of inherent ability. Before turning to evidence on reform effects on income and other outcomes, I therefore start by presenting evidence on how the reform impacted individual skills. Lager et al. (2017) study how the reform affected mens' cognitive and non-cognitive skills using data from the Swedish military enlistment. The results indicate that the reform raised cognitive skills by 5 percent of a standard deviation on average, and effects were larger among sons of fathers in low SES occupations (farmers, unqualified manual workers), and inexistent among sons of skilled workers and professionals. This result squares well with those presented in 6.4 above, which show that the reform had a larger impact on children from low SES backgrounds. The results on non-cognitive skills (emotional control) instead show negative effects on average (of 3 percent of a standard deviation), and when splitting by family background it becomes clear that the effect should be attributed to children from higher social classes. One possible interpretation of this result is that high SES sons fared worse in terms of non-cognitive skills in the comprehensive system, when exposed to a broader peer group. This finding is also in line with evidence from the Finnish reform, which indicates negative effects from late tracking on mental health for women from highly educated families (Böckerman et al. 2019).

The effects on skills and later-life outcomes may not only be explained by direct exposure to the reform and skills acquired by staying longer in school. As demonstrated by Koerselman (2013), curriculum tracking can have incentive effects for students already before the point of tracking: it creates incentives for students to work harder to be admitted to the desired track. Koerselman (2013) finds evidence of tracking effects in the U.K. – i.e. that students in the tracked system have higher test scores before the point of tracking compared to the non-tracked system. Using the Swedish compulsory school reform for a similar analysis however shows no incentive effects for the tracked system.

Finally, the reform might have affected skills through other mechanisms, for example through education spillovers to other individuals, or by "protecting" low-performing individuals from dropping out and giving them a second chance. Adermon (2013) studies sibling spillovers and finds that an older sibling exposed by the reform did not induce younger siblings to stay on longer in school. Fredriksson and Öckert (2013) study effects of school starting age and find that the educational achievement gap by school starting age is larger in the pre-reform early tracking system than post reform.

Income. Meghir and Palme (2005) presented the first reform estimates on earnings. They found a small average effect on earnings of 1.4 percent, but effects varied by parental background. Children with low educated fathers gained 3–7 percent, where the largest estimate refers to high-ability girls, and children with high educated fathers had non-trivial negative earnings effects. In light of the results by Lager et al. (2017) described above, and the literature documenting how important non-cognitive skills are for long-term outcomes, the negative earnings effects among advantaged children could be explained by a loss of socio-emotional skills rather than by a loss of cognitive skills.

More recently, Fischer et al. (2018) re-estimate the effects of the comprehensive school reform. While Meghir and Palme's original sample was based on random samples of two cohorts, Fischer et al. use population data covering cohorts born 1938–1954 and as such cover almost the whole implementation period. The results confirm a small positive earnings effect of the reform on average, but do not lead to the same conclusions when it comes to differences by parental background: Fischer et al. find positive effects for both children of low- and high-skilled workers.

Health. More educated individuals have better health – but the causal relationship is debated. Motivated by the quest to establish whether there is a causal link between education and health several papers study reform effects on health-related outcomes.

Spasojevic (2010) builds on a relatively small survey data set and uses the reform as an instrument for years of education, focusing on outcomes measuring self-reported health (BMI and a health index combining information about both minor and severe conditions). Spasojevic's analysis finds that one more year of education leads to a lower ill-health index and a higher probability to have a healthy BMI. Lager and Torssander (2012) study reform effects on mortality – they find that reform exposure reduced mortality risks after the age of 40. Meghir et al. (2018) study mortality, hospitalization and drug prescriptions, but find that the reform did not affect any of these outcomes. Finally, Fisher et al. (2019) study reform effects on mortality, self-reported bad health, and smoking and similarly find no effects on these health outcomes. A possible explanation as to why the results on health and mortality differ between studies is that the two latest studies use larger samples and a longer follow-up periods.

Finally, Palme and Simeonova (2015) study a more specific health outcome: incidence and mortality of breast cancer – a form of cancer positively associated with education and labelled a "welfare disease". They find that the reform increased the risk of women being diagnosed with breast cancer, and lead to an elevated probability of death from breast cancer.

Crime. Individuals engaging in criminal activity are on average low educated and a key policy question is whether educational interventions can help to protect individuals from a criminal career. Education can raise the returns to legal activities, can lead to a different peer group, and also protects individuals directly through the incapacitation effect while still in education.

Hjalmarsson et al. (2015) study the reform effect through years of schooling on convictions and incarcerations. One additional year of schooling induced by the reform decreased the likelihood of conviction by 6.7 percent, and incarceration by 15.5 percent among men. Meghir et al. (2012) study reform effects on crime both in the treated generation and among their offspring – i.e., the intergenerational effect of the reform on crime. The study finds that the policy reduced crime rates both for the targeted generation and among their children. Reform exposure reduced the probability of ever being convicted by 5 percent among men born 1954–1955. Sons of exposed men were also affected and had a lower probability of ever being convicted.

Political participation. Lindgren et al. (2017) take an interest in the social divide in political participation and ask whether the comprehensive schooling reform was an effective policy to raise political involvement among individuals from low SES backgrounds. The study focuses on political candidacy: it uses information on all nominated and elected candidates in six parliamentary, county council and municipal elections in Sweden between 1991 and 2010. The reform significantly increased the probability of political candidacy among individuals from working class backgrounds – the impact of family background on the likelihood of seeking public office was reduced by up to 40 percent.

Financial decision-making. More educated individuals participate in financial markets – even after controlling for wealth or income. Black et al. (2018) ask whether there is a causal link between education and investment behavior, and whether policies that increase educational attainment also change people's investment behavior. The study uses the reform as an instrument for years of education to analyze the effect of one more year of schooling on stock market participation. The results show that one more year of education increases stock market participation among men by 2 percentage points (over a baseline of 42 percent), and a suggestive channel explaining this result is greater financial wealth. There is no evidence that reform-induced changes in education affected women's financial decision-making, nor that there were spillover effects to children.

Girishina (2019) studies the effects of education on wealth. Using the reform as an instrument for education she finds that one more year of schooling increases the value of an individual's total assets, which is consistent with Black et al.'s conclusion.

Intergenerational spillovers. Strong intergenerational correlations are observed on a range of outcomes, such as education, earnings, health and criminal involvement. These correlations can be the result of direct causal influences from parents to children, through investment behavior or role model mechanisms, but are also explained by other underlying characteristics shared by parents and children that give rise to correlations in outcomes. Because of the difficulty in distinguishing between direct causal spillovers and correlations, the compulsory school reform has offered an opportunity to learn something about intergenerational spillovers of exposure to comprehensive schooling on offspring's outcomes.

Holmlund et al. (2011) study the effects of parents' schooling on children's schooling. The intergenerational correlations, regressing years of schooling of children on years of schooling of fathers/mothers, yields estimates at 0.23 and 0.28 years for fathers and mothers, respectively. Using the reform as an instrument for education, they find that one more year of schooling among parents leads to about 0.1 more years in the next generation, which is much lower than the correlations in the data. Also Sikhova (2019) shows that the reform affected intergenerational transmissions of human capital.

Lundborg et al. (2014) study intergenerational effects on sons' cognitive and non-cognitive outcomes, as measured at military enlistment at the age of 18. The study uses an IV approach and finds positive effects of maternal education on sons' cognitive ability and a global health index, but no effects of father's education. The effect magnitudes indicate that one more year of maternal schooling increases both cognitive ability and global health by about 0.1 of a standard deviation.

Lundborg and Majlesi (2018) consider that intergenerational transmissions also can run from child to parent. Using the reform as an instrument for education, they study how education in the treated generation affects health outcomes among elderly parents. If more education provides children with better resources to care for their elderly parents, it is possible that parental health will be affected. On average, the reform-induced changes in years of schooling do not affect parents' longevity, but there is some heterogeneity suggesting that low educated fathers are positively affected by their offspring's education – they exhibit higher survival at age 70–80 when their children are exposed to longer compulsory education.

Nybom and Stuhler (2016) study trends in intergenerational mobility and ask how such trends should be interpreted: changes in mobility can be the result of current policies, but can also be explained by events in the past that have dynamic effects on mobility over several generations. As an example, a shift towards a more meritocratic society increases mobility when highly skilled children from poor families are rewarded for their skills rather than their background. In the next generation however, it is possible that mobility declines again, since skills are passed on from parents to their children. To illustrate their structural model, Nybom and Stuhler show that individuals exposed to the Swedish compulsory school reform exhibit higher intergenerational mobility, but that the reform decreased mobility in the next generation.

7.1 IV and the exclusion restriction

The reform extended compulsory education, but also implied changes to the curriculum, affected tracking, and consequently the peer group composition. The effects of the reform on future outcomes should therefore be interpreted as the total (net) effect, incorporating all possible channels through which reform exposure affected outcomes. Despite this, a number of the papers mentioned above use the reform as an instrument for years of education (see e.g. Spasojevic 2010, Lundborg et al. 2014, Hjalmarsson et al. 2015, Black et al. 2018, Lundborg and Majlesi 2018), although the exclusion restriction is unlikely to hold. If we are primarily interested in understanding the effects of the Swedish reform on a variety of outcomes, the reduced form vs. IV distinction may not be so problematic – we can in most cases back out the reduced form, and the main question is to identify whether there is a significant effect, its sign, and which subgroups are mostly affected. Using the reform as an instrument for years of schooling is more problematic when we start interpreting the effect as caused by years only, and when comparing the estimates with those in the literature on returns to schooling. The estimates based on the Swedish reform are unlikely to be comparable to those of other studies both because the exclusion restriction is unlikely to hold, and because the compliers come from the lower tail of the education distribution. If there are heterogeneous treatment effects, it is not clear that we can compare studies using compulsory schooling reforms with those using e.g. college openings as instruments for years of schooling.

8. Conclusions

This paper demonstrates how a natural experiment can be used to estimate causal effects. The Swedish compulsory school reform was implemented gradually across cohorts and municipalities, which under certain assumptions makes it possible to estimate causal effects of the reform. Many papers have used differences-in-differences specifications to estimate effects of the reform on a range of outcomes, showing that extending basic education provided children from low socioeconomic background with better opportunities in life. Not only did they attain higher levels of education – they also earned higher earnings, were less likely to participate in crime, and more likely to run for office. But there is also some evidence in support of the notion that a comprehensive system could be harmful to high ability children. Meghir and Palme (2005) found negative earnings effects for high SES groups, although this result is not replicated by Fisher et al. (2018). Evidence both from Sweden (Lager et al. 2017) and Finland (Böckerman et al. 2019) indicate negative effects of comprehensive schooling on non-cognitive skills for children from higher socioeconomic backgrounds. The reform had winners and losers, but it undoubtedly reduced inequality.

Earlier academic work has argued that the reform was a failure, on the account that recruitment to tertiary education dropped among youth with working-class backgrounds for cohorts affected by the reform (Rothstein 1996).[8] But this conclusion does not account for the possibility that the counterfactual outcome could have been even less encouraging

[8] *Earlier work with a different conclusion regarding the reform outcome is Erikson (1996) who found that the introduction of the comprehensive school coincided with reduced social inequality in education.*

from an equality point of view and highlights the importance of rigorous evaluation methods. As I have shown in the empirical part of this paper, the reform increased the probability that children stayed in school beyond compulsory level, and this effect is twice as large for children from low educated families compared to high educated families. The reform also reduced the gap in cognitive skills between children of different social background. In this light, the reform must be considered successful in reaching its objective of reducing social inequalities. □

References

Acemoglu, Daron and Joshua D. Angrist (2000), "How Large are Human-Capital Externalities? Evidence from Compulsory Schooling Laws", in Bernake, Ben and Kenneth Rogoff (eds.), NBER Macroeconomics Annal 2000, Volume 15.

Adermon, Adrian (2013), "Sibling Spillovers in Education", in "Essays on the Transmission of Human Capital and the Impact of Technological Change", Doctoral thesis, Uppsala university.

Aigner, Dennis J. (1973), "Regression with a Binary Independent Variable Subject to Errors of Observation", *Journal of Econometrics* 1: 49–60.

Angrist, Joshua D. and Alan B. Krueger (1991), "Does Compulsory School Attendance Affect Schooling and Earnings?", *Quarterly Journal of Economics* CVI(4): 979–1014.

Arendt, Jacob Nielsen (2005), "Does Education Cause Better Health? A Panel Data Analysis Using School Reforms for Identification", *Economics of Education Review*, 24(2), pp. 149–160.

Ashenfelter, Orley and Alan B. Krueger (1994), "Estimates of the Returns to Schooling from a New Sample of Twins", *American Economic Review* 84(5): 1157–1173.

Behrman, Jere R. and Mark R. Rosenweig (1992), "Does Increasing Women's Schooling Raise the Schooling of the Next Generation?", *American Economic Review* 92(1): 323–334.

Black, Sandra E., Paul J. Devereux and Kjell G. Salvanes (2005), "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital", *American Economic Review* 95(1): 437–449.

Black, Sandra E., Paul J. Devereux, Petter Lundborg and Kaveh Majlesi (2018), "Learning to Take Risks? The Effect of Education on Risk-Taking in Financial Markets", *Review of Finance*, 22(3), pp. 951–975.

Borusyak, Kirill and Xavier Jaravel (2018), "Revisiting Event Study Designs, with an Application to the Estimation of the Marginal Propensity to Consume", unpublished manuscript.

Bound, John and David A. Jaeger (1996), "On the Validity of Season of Birth as an Instrument in Wage Equations: A Comment on Angrist & Krueger's "Does Compulsory School Attendance Affect Schooling and Earnings?", NBER working paper 5835.

Böckerman, Petri, Mika Haapanen, Christopher Jepsen and Alexandra Roulet (2019), "School Tracking and Mental Health", IZA Discussion paper No. 12733.

Chevalier, Arnaud (2004), "Parental Education and Child's Education: A Natural Experiment", IZA Discussion Paper No. 1153.

Currie, Janet and Enrico Moretti (2003), "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings", *The Quarterly Journal of Economics* 118(4): 1495–1532.

de Chaisemartin, Clément and Xavier D'Hautfoeuille (2020), "Two-way Fixed Effects Estimators with Heterogeneous Treatment Effects", forthcoming, *American Economic Review*.

Educational Bureau (1960-1964), Tables of pupils in the compulsory school.

Erikson, Robert and Rune Åberg, eds. (1984), *Välfärd i förändring. Levnadsvillkor i Sverige 1968-1981*. Stockholm. (English translation: W*elfare in Transition*. Oxford: Clarendon Press, 1987.)

Erikson, Robert and Jan O. Jonsson (1996), "Introduction. Explaining Class Inequality in Education: The Swedish Test Case", in Erikson, Robert. and Jan. O. Jonsson (eds.), Can Education Be Equalized?, Westview Press.

Erikson, R. (1996), "Explaining Change in Educational Inequality – Economic Security and School Reforms", in Erikson, Robert and Jan O. Jonsson (eds.), Can Education Be Equalized?, Westview Press.

Fischer, Martin, Gawain Heckley, Martin Karlsson and Therese Nilsson (2018), "Did Sweden's Comprehensive School reform reduce Inequalities in Earnings?", in Fisher, Martin, *The Long-Term Effects of Education on health and Labor Market Outcomes – Evidence from Historical School Reforms in Sweden and Germany*, PhD dissertation, University of Duisburg-Essen.

Fischer, Martin, Ulf-G Gerdtham, Gawain Heckley, Martin Karlsson, Gustav Kjellsson and Therese Nillson (2019), "Education and Health: Long-run Effects of peers, Tracking and Years", IFN Working Paper No. 1300.

Fredriksson, Peter and Björn Öckert (2014), "Life-Cycle Effects of Age at School Start", *The Economic Journal*, 124, September, pp. 911–1004.

Girshina, Anastasia (2019), "Wealth, Savings, and Returns Over the Life-Cycle: The Role of Education", Job Market Paper, Stockholm School of Economics.

Goodman-Bacon, Andrew (2018), "Difference-in-differences with Variation in Treatment Timing", NBER working paper 25018.

Hjalmarsson, Randi, Helena Holmlund and Matthew J. Lindquist (2015), "The Effect of Education on Criminal Convictions and Incarceration: Causal Evidence from Micro-data", *The Economic Journal*, 125 (September), pp. 1290–1326.

Holmlund, Helena (2007), "A Researcher's Guide to the Swedish Compulsory School Reform", SOFI working paper 9/2007.

Holmlund, Helena, Mikael Lindahl and Erik Plug (2011), "The Causal Effect of Parents' Schooling on Children's Schooling: A Comparison of Estimation Methods", *Journal of Economic Literature*, 49(3), pp. 615–651.

Härnqvist, Kjell (2000), "Evaluation through Follow-up. A Longitdinal Program for Studying Education and Career Development", in C.-G. Janson, ed., *Seven Swedish longitudinal studies in behavioral science*. Stockholm, Forskningsrådsnämnden.

Kane, Thomas J., Cecilia Elena Rouse and Douglas Staiger (1999), "Estimating Returns to Schooling When Schooling is Misreported", Working Paper 419, Industrial Relations Section, Princeton University.

Koerselman, Kristian (2013), "Incentives from Curriculum Tracking", *Economics of Education Review*, 32, pp. 140–150.

Lager, Anton and Jenny Torssander (2013), "Causal Effect of Education on Mortality in a Quasi-Experiment on 1.2 million Swedes", PNAS, 109(22), pp. 8461-8466.

Lager, Anton, Dominika Seblova, Daniel Falkstedt and Martin Lövdén (2017), "Cognitive and emotional outcomes after prolongued eduation: a quasi-experminent on 320 182 Swedish boys", International Journal of Epidemiology, 46(1), pp. 303–311.

Lindgren, Karl-Oskar, Sven Oskarsson and Christopher T. Dawes (2017), "Can Political Inequalities Be Educated Away? Evidence from a Large-Scale Reform", *American Journal of Political Science*, 61(1), pp. 222–236.

Lleras-Muney, Adriana (2005), "The Relationship between Education and Adult Mortality in the U.S.", *Review of Economic Studies* 72(1): 189–221.

Lochner, Lance and Enrico Moretti (2004), "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests and Self-Reports", *American Economic Review* 94(1): 155–189.

Lundborg, Petter, Anton Nilsson and Dan-Olof Rooth (2014), "Parental Education and Offpring Outcome: Evidence from the Swedish Compulsory School Reform", *American Economic Journal: Applied Economics*, 6(1), pp. 253–278.

Lundborg, Petter and Kaveh Majlesi (2018), "Intergenerational Transmission of Human Capital: Is it a One-Way Street?", *Journal of Health Economics*, 57, pp. 206–220.

Marklund, Sixten (1980), "Från reform till reform: Skolsverige 1950-1975, Del 1, 1950 års reformbeslut", Skolöverstyrelsen och Liber Utbildnings Förlaget.

Marklund, Sixten (1981), "Från reform till reform: Skolsverige 1950-1975, Del 2, Försöksverksamheten", Skolöverstyrelsen och Liber Utbildnings Förlaget.

Maurin, Eric and Sandra McNally (2005), "Vive La Révolution! Long Term Returns of 1968 to the Angry Students", IZA Discussion Paper No. 1504.

Meghir, Costas and Mårten Palme (2005), "Educational Reform, Ability and Family Background", *American Economic Review* 95(1): 414–424.

Meghir, Costas, Mårten Palme and Marieke Schnabel (2012), "The Effect of Education Policy on Crime: An Intergenerational Perspective", NBER Working paper 18145.

Meghir, Costas, Palme, Mårten and Emilia Simeonova (2018), "Education and Mortality; Evidence from a Social Experiment", *American Economic Journal: Applied Economics*, 10(2), pp. 234–256.

National Board of Education (Skolöverstyrelsen) (1954–1962), "Redogörelse för försöksverksamhet med enhetsskola", Aktuellt från Skolöverstyrelsen.

National Board of Education (Skolöverstyrelsen) (1960), "Enhetsskolan under 10 år, Kort redogörelse för försöksverksamheten läsåren 1949/50 – 1958/59", Kungliga skolöverstyrelsens skriftserie 46.

Nybom, Martin and Jan Stuhler (2016), "Interpreting Trends in Intergenerational Mobility", manuscript. Previous version published as SOFI Working paper 3/2014.

Oreopoulos, Philip, Marianne E. Page and Ann Huff Stevens (2006), "The Intergenerational Effects of Compulsory Schooling", *Journal of Labor Economics* 24(4): 729–760.

Palme, Mårten and Emilia Simeonova (2015), "Does Women's Education Affect Breast Canceer Risk and Survival? Evidence from a population Based Social Experiment in Education", *Journal of Health Economics*, 42, pp. 115–124.

Pekkarinen, Tuomas, Roope Uusitalo and Sari Pekkala (2006), "Education Policy and Intergenerational Income Mobility: Evidence from the Finnish Comprehensive School Reform", IZA Discussion Paper No. 2204.

Rothstein, Bo. (1998), "The Social Democratic State: Swedish Model and the Bureaucratic Problem of Social Reforms", University of Pittsburgh Press.

Sikhova, Aiday (2019), "Better Parents or Richer Parents: understanding Intergenerational Transmission of Human Capital", manuscript.

Skatteverket (2007), "Sveriges församlingar genom tiderna", http://www.skatteverket.se/folkbokforing/sverigesforsamlingargenomtiderna.4.18e1b10334ebe8bc80003817.html.

Spasojevic, Jasmina (2010), "Effects of Education on Adult Health in Sweden: Results from a Natural Experiment", *Contributions to Economic Analysis*, vol 290, Chapter 9.

Statistics Sweden (1968–1969), Tables of pupils in the compulsory school, Statistiska meddelanden U1968:2, U1969:5.

Viarengo, Martina (2007), "An Historical Analysis of the Expansion of Compulsory Schooling in Europe after the Second World War", Working paper 97/07, Department of Economic History, London School of Economics.

APPENDIX

A.1 Reform coding for register data

There are two independent ways to obtain a reform code to attach to register data, one is to use available documentation on when the reform was in place, the other is to deduce it from a register-based data set of large sample size.

A.1.1 Reform coding based on documentation [CODING 1]

To obtain a complete code of the implementation of the nine-year comprehensive school, I have, with help from Anders Björklund, Valter Hultén and Mikael Lindahl, used two main sources:

a. Marklund and the National Board of Education

Marklund (1981) provides a list with the quantitative development of the reform from 1949/50 to 1960/61. This documentation states which municipalities (or which parts of a municipality) that entered the assessment programme each year. However, Marklund (1981) does not list which grades that were exposed to the reform, in each municipality and each year. This information is available in the yearly publications that the National Board of Education published during the course of the trial period (Aktuellt från Skolöverstyrelsen 1953-1962). These publications summarized many of the aspects of the ongoing educational development, one of which was the participating municipalities and also which grades that were subject to the reform. These publications together cover the years 1951/52 to 1960/61. It is noteworthy that these two sources, Marklund and the Board of Education reports, in general coincide in terms of municipalities listed. There is one difference in that the yearly publications from the Board of Education list a few more municipalities as participating in the reform than what is mentioned in Marklund.

In the guidelines for the reform assessment, it was stated that only pupils in grades 1 through 5 would be subject to any changes. Therefore the above information from Marklund and the Board of Education, that covers almost the whole assessment period with the last year being 1960/61, makes it possible to assign whether individuals born in 1938 to 1949 were subject to the reform or not (the 1938 cohort was the first one to be affected, 1949 is the cohort of 5th graders in 1960/61). From 1961/62 these sources do not tell us what is going on, but we know for sure that pupils in 6th grade and above should not be subject to any changes. Therefore, we can code the cohort of 6th graders in 1961, and older pupils (the 1949 cohort and older), but for younger cohorts there could be changes from 1961 an onwards that are not captured by these sources.

b. The Educational Bureau (1960–1964) and Statistics Sweden (1968–1969)

When the final decision about the complete introduction of the new school was taken, in 1962, the experimental period also came to an end. Now, municipalities were required to implement the reform, but a transition period allowed them to postpone the implementation, however no longer than until 1969. Thus, also in the early 1960s there is some variation in reform implementation, affecting cohorts born from 1950 an onwards. Marklund (1981) and the publications from the Board of Education were mainly concerned with the assessment programme, and thus they do not document reform implementation in the 1960s. To use the variation in compulsory schooling legislation for the 1950s cohorts, it is possible to trace reform implementation in the early 1960s from municipality tables from Statistics Sweden (1968, 1969) and from the Educational Bureau (Undervisningsbyrån) (1960, 1961/62, 1963/64). For each municipality, the tables from the Educational Bureau give the number of pupils in each grade in both the old school (folkskolan) and the new nine-year comprehensive school (grundskolan). From such a table it is possible to see in which grade and year the implementation took place. See the examples following below.

Example 1

Municipality m in year t, cohort size is around 500.

Grade	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Type of school	Nr of pupils								
Folkskola (old)	0	500	500	500	0	500	500	0	0
Grundskola (new)	500	0	0	0	500	0	0	0	0

Municipality m in year t+1

Grade	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Type of school	Nr of pupils								
Folkskola (old)	0	0	500	500	0	0	500	0	0
Grundskola (new)	500	500	0	0	500	500	0	0	0

Municipality m in year t+2

Grade	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Type of school	Nr of pupils								
Folkskola (old)	0	0	0	500	0	0	0	0	0
Grundskola (new)	500	500	500	0	500	500	500	0	0

Municipality m in year t+3

Grade	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Type of school	Nr of pupils								
Folkskola (old)	0	0	0	0	0	0	0	0	0
Grundskola (new)	500	500	500	500	500	500	500	500	0

From these tables it is possible to conclude that the cohort of 5th graders in year t, that is, the cohort born in t-11, is the first cohort in municipality m, to be affected by the reform. All younger birth cohorts were also affected (since even if you were in grades 2–4 in year t, you would eventually reach grade 5 and thus be phased into the new school).

Example 1 is a stylized example, in reality the tables year-by-year might look either as in example 2 or 3 below.

Example 2

Municipality m in year t

Grade	1	2	3	4	5	6	7	8	9
Type of school	Nr of pupils								
Folkskola (old)	0	500	500	500	250	500	500	0	0
Grundskola (new)	500	0	0	0	250	0	0	0	0

Municipality m in year t+1

Grade	1	2	3	4	5	6	7	8	9
Type of school	Nr of pupils								
Folkskola (old)	0	0	500	500	0	250	500	0	0
Grundskola (new)	500	500	0	0	500	250	0	0	0

In this case, it is not clear which cohort that should be assigned as the first reform cohort. Is it the cohort in 5th grade in year t or in t+1? In these cases the reform implementation has been set to start when at least half of a cohort is facing the reform. However, it is clear that the coding here will introduce some measurement error.

Example 3

Municipality m in year t

Grade	1	2	3	4	5	6	7	8	9
Type of school	Nr of pupils								
Folkskola (old)	0	500	500	500	0	500	500	0	0
Grundskola (new)	500	0	0	0	500	0	0	0	0

Municipality m in year t+1

Grade	1	2	3	4	5	6	7	8	9
Type of school	Nr of pupils								
Folkskola (old)	0	0	500	500	0	0	0	0	0
Grundskola (new)	500	500	0	0	500	500	500	0	0

This example shows that the tables are not always coherent between years. In year t, it looks like the first cohort is the fifth graders in t, whereas in year t+1, it seems like the first cohort is the one of 7th graders in t+1. In these cases, the information on which cohort entered first is taken from the last table that reveals a shift between the old and the new school (in the light of the example above, it would be the 7th graders in year t+1, that is the cohort born in t+1-13).

Note that the first table from the Educational Bureau is from 1960/61, which means that municipalities that introduced the reform very early cannot be coded using this second source of information. That is, in the case all pupils in grade 1 through 9 were already in the new school in 1960, it is not possible to see when the shift took place. In those cases, we rely solely on the first source (Marklund). Luckily, there is some overlap between the two sources: for 158 municipalities I have obtained a coding from both Marklund and the Bureau. In 9 out of 158 cases, the coding differs between these sources, and in those cases, I have used Marklund.

In some cases we know from Marklund that a municipality introduced the reform in different parts of the municipality at different points in time. If these were early implementers, the statistical tables do not reveal when the majority of the pupils in a municipality were shifted into the new system. In that case, Marklund states which school district, or which schools within a municipality that introduced the new school, it is possible to assign these schools to a sub-region of the municipality (a parish). There are however, a few municipalities where we know that the reform was introduced gradually, but there is no information on which schools or which part of the municipality. These municipalities cannot be coded and must be dropped from the sample: Hälsingborg, Jönköping, Linköping, Skellefteå, Sundbyberg and Södertälje.

The three big cities, Stockholm, Göteborg and Malmö, are also problematic to code. They implemented the reform at different points in time in different parts of their municipalities, and the coding has been constructed as follows (note that information on parish is necessary)[9]:

Stockholm. From the statistical tables from the Educational Bureau, it is clear that in 1962, the whole cohort of 8th graders (the 1948 cohort) was shifted into the new comprehensive school. However, reform implementation had started gradually earlier, at first in the southern suburbs of Stockholm. Based on information on parish of residence, the south suburbs can be dropped, and the change that affects (approximately all the rest) of Stockholm can be coded and the first cohort affected is set to 1948.

Göteborg. The first cohort where all pupils are in the new school is the 1950 cohort. Early implementing parishes are dropped.

Malmö. The first cohort where all pupils are in the new school is the 1949 cohort. Early implementing parishes are dropped.

The procedure outlined above allows me to find the first cohort affected by the reform in almost all municipalities. Some could not be coded due to ambiguity as to which part of the municipality implemented the reform (mentioned above). Yet another three municipalities could not be coded, simply because they did not show up in the statistical records: Fjälkinge, Svarteborg and Sörbygden.

A.1.2 Reform coding deducted from large-sample register data [CODING 2]

With a large enough sample, it is possible to adopt the following strategy to find out when a municipality implemented the reform: split the sample (randomly) in two parts (in this paper 60/40 percent, sampling by municipality to maintain proportions at the group level). I use the 60 percent sample to identify the reform date, and 40 percent for estimation. Within the 60 percent sample, I drop all individuals with education higher than the new compulsory minimum (grundskola), using the information on completed education levels in Statistics Sweden's education register. Now we are left with only observations of the old minimum (folkskola) and the new minimum (grundskola). Assign a dummy equal to one for the new comprehensive school. Collapse this data by birth cohort and municipality, and look at the average of the comprehensive school dummy for each cell. With a clean-cut implementation, we should observe that within a municipality, the average shifts from 0 to 1 between two specific cohorts, and this is when the reform is implemented. In reality, the cohort-to-cohort changes are not always so clean, and one can assign a reform to cohorts where the cohort/municipality-average of the compulsory school dummy is >=0.5.

[9] *Assigning the reform based on information of parish of residence is an approximation, where I map a given school (or school district) to the parish/es in which it lies.*

With this procedure it might be the case that you assign the reform to cohort t, but in cohort t+1 the dummy average is <0.5 and for cohort t+2 it shifts back to >=0.5. In the empirical part of this paper I have, in the case I observe more than one shift, assigned the reform to the shift that never moves back below 0.5.

Note also that some municipalities did not implement the reform uniformly within itself; most notably this was the case in the big cities. To arrive at a cleaner definition I therefore drop "early implementing" parishes in the big cities, and other unclear parishes/municipalities where it is known that implementation was gradual (see coding 1 above).

A global CO_2 price – Necessary and sufficient

John Hassler
Institute for International Economic Studies (IIES)
at Stockholm University

Abstract

During the last 10 years, I have spent most of my research time on the economics of climate change. Basically all of it has been done together with my colleagues Per Krusell at IIES and Conny Olovsson at Sveriges Riksbank. Being a truly cross-disciplinary field, the close interaction with many natural scientists, in particular Jonas Nycander at the department of meteorology at Stockholm University has been an absolute necessity. In this article, based on a talk at the Finnish Economic Association annual conference in February 2020, I summarize what we have learned over the years. Hopefully it can be of value to other researchers and policy makers. In any case, I am convinced that economics is key for understanding what to do about global warming. The key conclusion is that a global agreement on a (minimum) price on fossil carbon emission is necessary, sufficient and efficient solution to limiting climate change.

The natural science background

Without the inflow of sunlight, life on earth as we know it would be impossible. On average over time and geographic space, earth receives an inflow of energy from the sun of 340 W/m^2. For earth not to accumulate heat, an equal amount has to flow out from earth into space. We call the account of these aggregate flows *earth's energy budget*. When the budget is balanced, in- and outflows are equal, and no heat is accumulated.

The inflow of energy is largely in the form of visible light. Apart from a third of the outflow (being direct reflection of visible light), the outflow is instead largely in the form of infrared radiation. Sunlight passes easily through the atmosphere, but this is not the case for infrared radiation since the atmosphere contains greenhouse gases which trap the infrared radiation. The most important greenhouse gases are water vapor, carbon dioxide (CO_2) and methane. Human activities have increased the amount of CO_2 and methane in the atmosphere. This creates a surplus in the energy budget by reducing the outflow.[1]

A surplus in the energy budget leads to accumulation of energy in the form of heat – the temperature increases. As the temperature increases, more energy is emitted from earth to space and eventually, budget balance is restored, but at a higher global temperature.

Modern models incorporating both the carbon circulation[2] and how changes in the energy budget affect the climate have been shown to imply that the global temperature increase is proportional to the accumulated amount of fossil carbon that is emitted since we started emitting (Matthews et al., 2009). A key feature behind this result is that a substantial share of carbon stays in the atmosphere for thousands of years. Methane, on the other hand, while being a highly potent greenhouse gas, only stays in the atmosphere for a short time (most is gone after a few decades). Thus, in contrast to CO_2, the warming effect of methane depends on the flow of emissions, not the accumulated amount.

Although the models agree on the proportionality between accumulated emissions and the increase in temperature, they do not agree on the proportionality factor (sometimes called CCR – Carbon Climate Response). A key explanation for this disagreement is that science is still unsure of how cloud formation is affected by emissions from fossil fuel. Changes in when and where clouds are formed are important for whether the direct effect of CO_2 on the energy budget is dampened or reinforced. Due to this (and other) feedback mechanisms that are difficult to quantify, there is a large range of uncertainty regarding the proportionality factor between accumulated emissions and global warming. UN's climate panel IPCC suggests a likely interval of 0.8 – 2.5 degrees C per TtC[3]. This is a very large degree of uncertainty. To see this, note that we have so far globally emitted close to 0.6 TtC since we started some 150 years ago. If the proportionality factor is 0.8, we can emit three times as much as we have done so far before reaching an increase of 2 degrees global warming. At the current rate of emission of approximately 0.01 TtC per year, this would take a couple of hundred years. If, on the other hand, the proportionality factor is in the upper end of the interval, we can only emit 0.2 TtC more if we want to stay below 2 degrees. This amount would be emitted in 20 years at the current emission rate.

The increase in the global mean temperature is a key summary measure of climate change. However, climate change is obviously enormously multi-faceted. The same is true of the consequences for human welfare. The direct effect of climate change might be small or even positive in some parts of the world, for example in the Nordic countries. In other parts of the world, often densely populated, climate change may have catastrophic consequences. Also the uncertainty about the consequences of climate change and the possibility to adapt to it is very large and in the same order of magnitude as the

---

[1] *More CO_2 in the atmosphere implies that the altitude at which heat radiation can "escape" the atmosphere is pushed outwards towards colder layers of the atmosphere. Since the amount of energy radiated depends on the temperature, less energy is emitted. More CO_2 thus works like putting on a thicker blanket on earth. The first to quantify the relation between CO_2 in the atmosphere and temperature was Svante Arrhenius (Arrhenius, 1896).*

[2] *The circulation of carbon between different carbon reservoirs (carbon sinks) such as the atmosphere, the biosphere and the oceans.*

[3] *Terraton carbon, i.e., 1000 billion ton carbon. Since burning one ton of carbon produces 3.67 tons of CO_2, these numbers can easily be expressed in CO_2 units.*

natural science uncertainty about the CCR. The death of our civilization as a consequence of climate change is science fiction, but cannot be ruled out on logical or scientific grounds.

Economics and economists are needed

Most researchers who are active in the area of climate change are natural scientists. However, economists are needed to provide key answers to the questions about how to reduce emissions. If it were the case that the reasonable response to the climate change problems was to impose a global, total and immediate ban on fossil fuel, economists would not be required. Arguably, however, such a medicine would kill the patient and cannot be prescribed. Instead fossil fuel needs to be phased out over time and perhaps at different speeds in different parts of the economy and differentially for different fossil fuels. Doing this by central planning where the emissions of each emitter is prescribed by a global emission agency is not practical, to say the least.

Instead we need to realize that emissions are the consequences of economic activities such as consumption, investment, and production. Behind these activities are humans and firms that make decisions, largely on markets. How such decisions are made and how they can be affected by various policies is what economists study.

A key economic lesson going back to the work of Pigou exactly 100 years ago is that markets will not deliver good outcomes when there are externalities (Pigou, 1920). With this we mean a situation where an activity on a market has direct consequences for other parties than the ones involved in the transaction behind the activity. Emission of CO_2 is a perfect example of such an externality. When I ride my motorcycle and burn the gasoline I bought at the gas station, the emitted CO_2 quickly (time scale of weeks) spreads in the global atmosphere and affects the climate everywhere and for a very long time. These effects are not part of the price I pay for the gasoline unless there is policy that makes me pay for them. Without a price that makes me pay for the externalities, the market fails to deliver the socially right outcome. I use too much gasoline.

An alternative way of describing the market failure is to note that the atmosphere's capacity to absorb CO_2 is a resource in limited supply. If this resource is free to use for everyone, we will get overuse in the same way as free access to the fish in the ocean or trees in the forest leads to overuse. Economist call this phenomenon the Tragedy of the Commons.

It is sometimes claimed that the causes of our problems with climate change is economic and/or population growth. Is this correct? Well both yes and no. No since the root cause of the problem is the lack of a price on emissions – the atmosphere capacity to absorb CO_2 is freely up for grabs. Yes since growth exacerbates the negative consequences of the absence of a price on emissions. Again, this is very similar to the case of fish in the oceans. We get overfishing if everyone is allowed to fish as much as they want without a price or a fishing quota. But the amount of overfishing certainly increases with growth in technology (larger boats), in population and GDP. However, the solution to the problem is not to go back to small fishing boats bound to the coastlines. It is to regulate the fishing industry with prices or quotas. The same thing is true for CO_2 emissions.

Integrated assessment models

One hundred years ago, Arthur Pigou provided the key conceptual insight required to understand what to do with the externalities – a price equal to externality must be imposed on the agents who control the activity leading to it. However ingenious Pigou's idea was, it is not sufficient to give policy advice. For this, we need quantitative answers about how high the price should be. To provide incentives for policy makers to follow our advice, we need to show what happens under different, also suboptimal, polices. For this, we need what is called *Integrated Assessment Models*, IAM's. William Nordhaus received the Nobel Economics Prize in 2018 for being the first to construct such models.

IAM's consist of three modules.

1. A carbon circulation module that describes the circulation of carbon between different carbon sinks such as the atmosphere, the biosphere and the oceans.

2. A climate module. This is built around the energy budget described above and describes what happens with various aspects of the climate over time when there is a budget imbalance.

3. A global economy module where production, consumption and emissions are determined and climate damages incurred.

The modules are linked in the following way; in the economy module a dynamic path of emissions of CO_2 are determined together with a number of other variables like income, production, consumption and investment. Emissions enter as an input to the carbon circulation module. There, a path of atmospheric CO_2 concentrations is determined. This becomes the input into the energy budget in the climate module in which the dynamics of the climate is determined. Finally, to close the loop, the climate is affecting the economy that incurs various damages caused by climate change.

All IAMs are built in this way, but they of course differ with respect to e.g., the degree of complexity of the various modules. Regardless of this, all modules need to be consistent with observations and our understanding of how the world works. This is of course as important when it comes to the economy as it is for the natural science modules.

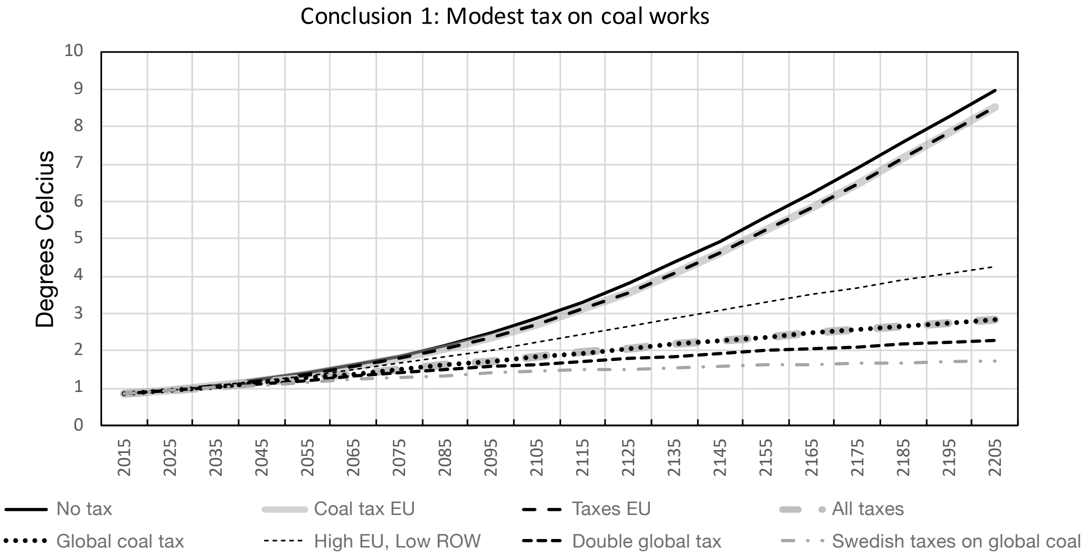
In previous work (Golosov et al., 2014) we constructed an IAM based on the fundamental contributions by Nordhaus (Nordhaus, 1994). We used it in particular to provide a simple formula for the optimal price on emissions. Key factors that determine the optimal level are how sensitive the climate is to emissions, how long carbon stays in the atmosphere and how damaging climate change is for human welfare. As described above we unfortunately have quite limited knowledge of these factors. Another important factor in determining the optimal level of the tax is how much we discount the welfare of future generations. Here, the ambiguity is due to the fact that we all can have different opinions on how to do this discounting.

In more recent work (Hassler et al., 2018 and 2020), we have used the model to address more positive questions. In those papers, three conclusions are drawn.

First, already a modest global price on emissions is effective in curbing climate change. Already an emission price of around 20 dollar per ton of CO_2 has strong effects on emissions and is in the main calibration sufficient to keep the temperature under 2 degrees during this century. Since a liter of gasoline produces around 2.3 kg of CO_2 an emission price of 20 dollar per ton amounts to only 4.6 euro cents per liter of gasoline. In figure 1, we show the effect on global warming over time for different emission prices. The main case is the 20 dollar price. Note, however, that in all simulations the price grows over time in at the same rate as global GDP. In Golosov at al. (2014), we show that this is optimal.

The lion's share of the effect on emissions in the model does not come from reductions in oil consumption but from coal. In sharp contrast to conventional oil, the market price of coal is close to the cost of extraction and the long-run elasticity of supply is high. Thus, even a small tax on emissions can make a lot of coal based energy unprofitable. This is fortunate, because also in contrast to conventional oil and gas, coal reserves are huge and most of it must stay in ground for any reasonable climate target to be met. Conventional oil instead, exists in limited supply and the model result clearly indicates that it is likely that it is socially optimal to use it also when the climate externalities are taken into account. I will return to this issue below.

Figure 1. Global warming for different emission prices.



An important qualification needs to be made here, however. For a modest tax to be effective, it has to be global. If, for example, China is allowed to continue using unpriced coal it becomes practically impossible to reach any climate goals. Even if the rest of the world implements a tax 20 times higher than the 20 dollar per ton discussed above it cannot compensate for the high emissions of China. Similar results accrue if fast growing regions like India or Africa are left out of an agreement on carbon pricing.

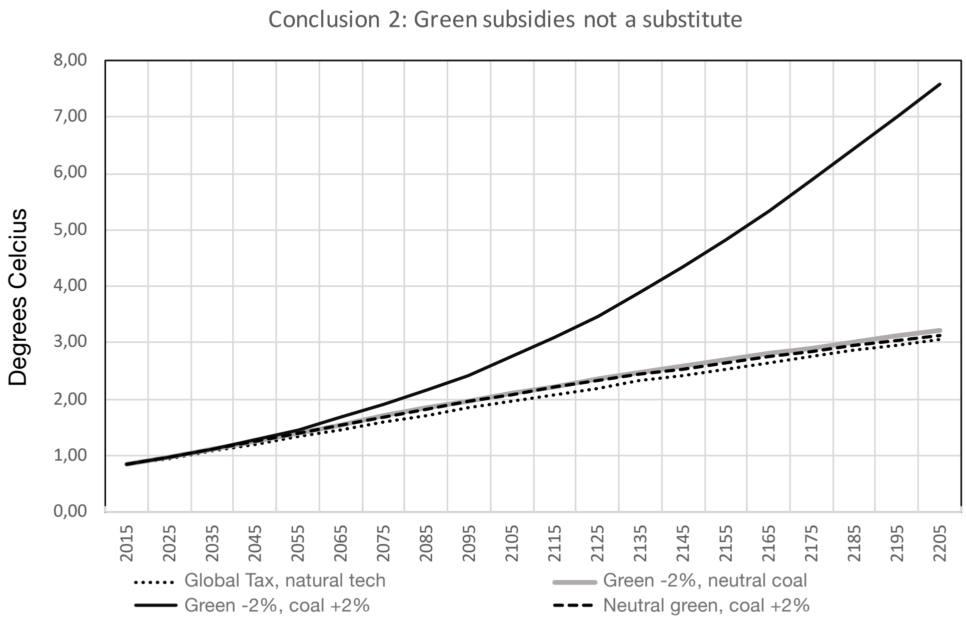
The second conclusion we draw is that subsidies to the development of cheap green energy is not likely to be an effective substitute for carbon pricing. In Figure 2, we show the consequences of a few experiments. The first is to subsidize technological improvements of green energy so that its price falls by 2% per year. At the same one manages to make the price of coal increase by 2% per year, either by a tax or by stopping technological advances in the coal industry. The consequence of this experiment is represented by the green line in Figure 2. As we see, global warming is very similar to the case of a moderate global tax represented by the dashed red curve (being the same as blue curve in figure 1). The second experiment is to only make green energy cheaper. This is represented by the thin black curve. As we see, this does not help the climate at all.

The reason for the result that cheaper green energy is unable to limit climate change is that the aggregate substitutability between different sources of energy is fairly low. Based on empirical surveys we set the elasticity of substitution to 0.95. This implies that cheaper green energy leads to more use of it, but it does not reduce the consumption of fossil based energy. So far, this is what we have seen globally. The use of green energy increases fast, but does not seem to drive out the use of fossil energy unless it is taxed.

There are many reasons for the, perhaps surprising, finding that the elasticity is not higher. One is that that green energy in the form of wind and sun, in contrast to fossil based energy, is non-controllable. This implies that the larger the share of wind and sun in the energy mix, the more negative becomes the correlation between price and production. The price is low when the sun shines and the wind blows. Therefore, a higher share of wind and sun reduces its profitability relative to controllable alternatives. Clearly, technical development in storage and demand flexibility may lead to higher substitutability. The hope that this will be sufficient for cheap green energy to drive out fossil energy without taxes seems fragile.

The third and final conclusion addresses the huge uncertainty described above. A traditional calculation of the optimal tax requires that a probability distribution is assigned to uncertain parameters like the response of the climate to accu-

Figure 2. Global warming for different price paths of green and fossil energy.



mulated carbon emissions. Natural science cannot provide reliably such distributions since the uncertainty comes from the fact that different models yield different results and no one knows which is the right model.

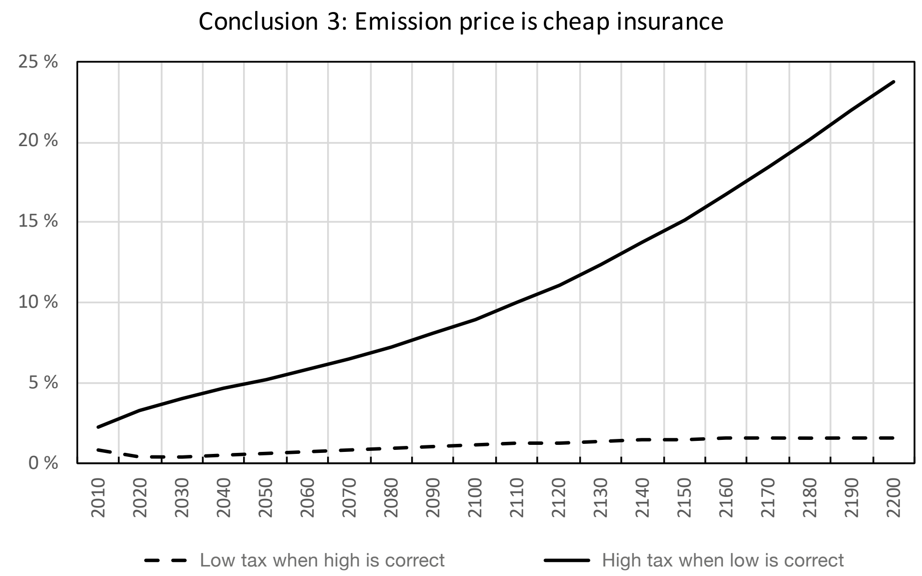
However, we argue that IAM's can provide valuable information also in a situation of such Knightian uncertainty. To illustrate this we note that given the large uncertainty, any chosen policy will ex-post turn out to be sub-optimal with probability one. But, all policy mistakes are not equally costly and the model can be used to evaluate this. A good policy recommendation in a situation of large uncertainty is to choose a policy that is robust, i.e., it is producing outcomes that are relatively insensitive to the things one is uncertain about.

Our simple application of this idea is to calculate the consequences of two policy mistakes. The first is to hope for the best and set a low carbon price that is optimal if the climate sensitivities to emission are low and climate damages are small. The policy mistake is realized ex post, when it turns out that climate sensitivities are high and climate damages large so that a high carbon price should have been chosen. The second policy mistake is the complete opposite. A high carbon price is chosen but ex-post it is realized that this was unnecessary since the climate sensitivity and climate damages are small.

To operationalize these ideas, we set the low climate sensitivity to the value at the lower end of the interval given by the climate panel IPCC.[4] Similarly, we set the high climate sensitivity to the value at the high end. Furthermore, we use a survey (Nordhaus and Moffat, 2017) of studies on global climate damages to get a similar range of the likely degree of damage sensitivity to climate change (see Hassler et al., 2018 for details). We use the end-values in this interval to generate a high and low economic sensitivity. In the best of cases, the climate sensitivity is low and the economic damage sensitivity is also low. If this turned out to be right, the optimal tax is low, according to our calculations only 6.9 USD per ton carbon (1.9 USD per ton CO_2). In the opposite case, high climate sensitivity and high economic damage sensitivity, the optimal tax becomes 264 USD per ton carbon (72 USD per ton CO_2). The first policy mistake is now to set the tax to 6.9 USD per ton carbon, while parameters are such that 264 USD is optimal. The second is to set it to 264 USD per ton carbon, while the right tax is 6.9.

[4] We use another measure of climate sensitivity than the one discussed above, namely the increase in the global mean temperature associated with a doubling of the atmospheric CO_2-concentration. IPCC's likely range for this value is 1.5 to 4.5 degrees C. See Hassler et al. (2018) for details.

Figure 3. Costs in terms of lost consumption for two types of policy mistakes



What we find is that these two types of policy mistakes have very different costs. The first, setting a low carbon price while a high turned out to have been right has much larger negative consequences than the opposite. Figure 3 depicts the cost of the two policy mistakes over time measured as a share of aggregate global consumption. As we see the costs are very different.

Note also that this is the cost of the policy mistake. In addition, welfare is of course much lower if the sensitivities are high for a given policy. The conclusion is therefore that an ambitious climate policy is a good insurance. It does not cost much if it turns out you don't need it but it is very good to have if you do. It should be noted, however, that the argument relies on using a global carbon price as the climate policy. One can certainly think of climate policies that achieve the aim of climate neutrality in very expensive ways. In this case, the precautionary argument for the policy does not fly.

Timing of fossil fuel phase-out

In the longer run, say during this century, all fossil fuels needs to be phased out. In this section I will discuss the question which order fuels should be phased out and how we should over time allocate a given amount of emissions across different fuels. A way to analyze this question is to think about what would happen to the use of different fuels if a tax that corrects the emission externality were to be introduced. Fuels that remain profitable when the correct tax is introduced are by construction socially valuable to use. Their private values are larger than the damages they incur. The opposite is true for fuels that cannot bear the correct tax without becoming unprofitable.

As discussed above, the level of the optimal tax on carbon emissions depends on highly uncertain parameters and is therefore very hard to pin down. However, the analysis does not require a precise value of the tax. Coal for heat and electricity production is unprofitable also under very modest taxes. The current price of emissions allowances in the EU-ETS, around 25 euros per ton of CO_2 equivalent to approximately 6 euro cent per liter gasoline, makes coal unprofitable. Even in the US, the coal industry is unprofitable. Since 2011, the Dow Jones U.S. Coal index is down 99%. Despite the supposedly coal-friendly policies of Donald Trump, the index fell by 70% during 2019 and another 70% this year.

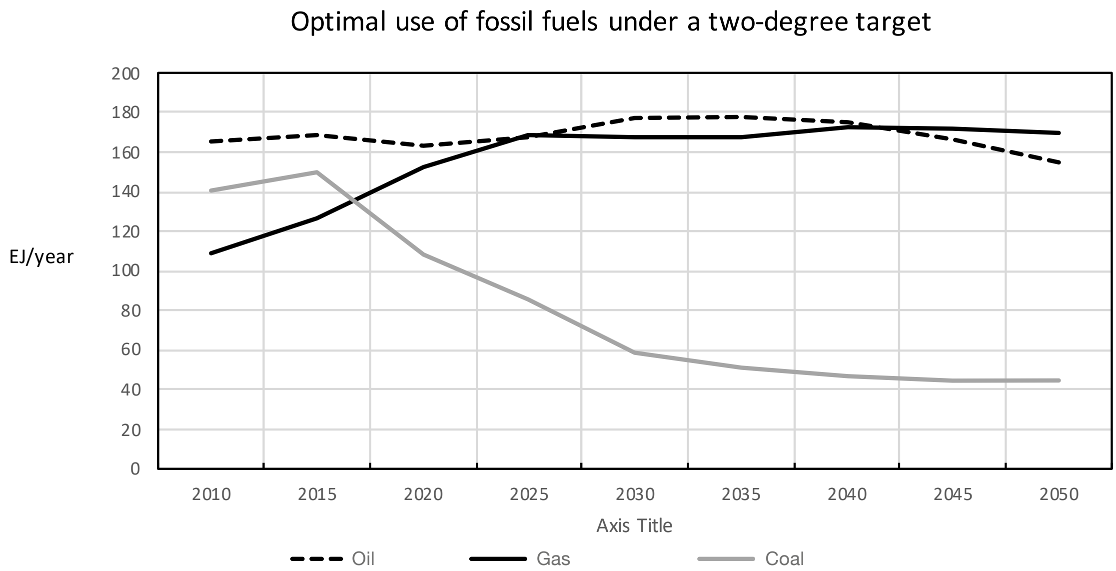
Conventional oil and gas are different. Even the high taxation of fossil fuels for transportation in Western Europe is clearly not making the sale of diesel and gasoline unprofitable. It certainly affects consumption and spurs the development of alternative green technologies. However, it is likely that conventional oil production will be profitable also with a carbon tax in line with the Swedish applied globally.[5]

Unconventional sources of oil and gas (e.g., from fracking, arctic reserves and tar-sand) are more sensitive to taxation. It is likely that a tax like the Swedish, and perhaps also lower rates, would make them unprofitable and take away the incentives to develop new techniques for using currently unprofitable sources of fossil fuel.

The conclusion from this somewhat informal analysis is that coal is the fossil fuel that should be phased out first and that most of the remaining reserves of it should stay in ground. Conventional oil and gas should be used, probably until we run out of it.[6] Unconventional reserves should stay in ground and new technologies for using them should not be developed.

My simple analysis can be done more elaborately. An example is McGlade and Ekins (2015). They calculate the optimal differential phase-out of coal, gas and oil under an emission budget that (with some probability) keeps global warming below 2 degrees C. The result of their analysis is depicted in figure 4.

Figure 4. Optimal phase-out of oil, natural gas and coal.



Conclusion and policy advice

Let me end the discussion by drawing some conclusions relevant for policy makers based on the research done by me and my colleagues, as well as many others.

1. A global agreement on a (minimum) price on fossil carbon emission is necessary. In fact it is also likely to be sufficient to curb climate change. So far, international negotiations have focused on country-specific quotas. The issue of an

[5] As a back-of-envelope calculation, we can note that a barrel of oil contains around 115 kg of carbon, producing 420 kg of CO_2 when burnt. The Swedish carbon tax of around 100 euros per ton CO_2 implies a tax in the order of 40 euros per barrel. This is sizeable, but in the same order of magnitude as the cost advantage of conventional oil over more unconventional sources.

[6] Who should use the conventional oil and gas is another question. On equity grounds, one could argue that we in the west should leave the oil for consumers in China, India and Africa.

agreement on a price has not seriously been at the negotiation table. Clearly, coming to such an agreement is not easy, but other ways are not likely to be easier.

How the price is implemented, with a tax or tradable emission allowances is not important to agree on. Neither is what is done with the revenue from the system. Our emission trading system EU-ETS is after the latest reforms in 2018 a show case. It shows that also a large region of quite heterogeneous countries can come to an agreement on an effective climate policy based on a price of emissions.

2. Coal is the main climate issue, neither oil nor gas. Unless we manage to convince the Chinese to phase out their coal dependence and make sure India and Africa don't follow the same coal intensive development path, we cannot solve the climate problem.

3. National climate policies must be designed with a global perspective in mind. Many well-intended national climate policies only move emissions to other countries. Such polices are fruitless and risk taking the focus away from the global issue. Certainly, rich countries like Sweden and Finland can affect others by being front-runners. A clever and cost-effective transition to climate neutrality that can be used in other countries is likely to be spur more followers than a costly one.

4. Subsidies to green technology may be a valuable complement to pricing of emission by facilitating the change, but it is not a substitute. Subsidies to technologies that can be used in other countries to de-carbonize are valuable. However, subsidies should not be used for technologies that are not scalable and only helpful for reaching national emission targets.

5. Limiting climate change to say, 2-2.5 degrees Celsius does not have to be expensive and should likely not lead to large global damages although regional climate damages can be devastating. Ill designed and uncoordinated climate policy may be excessively costly. □

References

Arrhenius, Svante, (1896), On the Influence of Carbonic Acid in the Air upon the Temperature of the Ground. Philosophical Magazine and Journal of Science, 41:5, p237–276.

Golosov, M., P. Hassler, P. Krusell and A. Tsyvinski (2014), Optimal taxes on fossil fuel in general equilibrium. *Econometrica*, 82:1 p. 41–88.

Hassler, J., P. Krusell and C. Olovsson, (2018), The Consequences of Uncertainty: Climate Sensitivity and Economic Sensitivity to the Climate, *Annual Review of Economics*, 10, pp. 189–205, 2018.

Hassler, J., P. Krusell, C. Olovsson and M. Reiter (2020), "On the effectiveness of climate policies", IIES WP.

IPCC (2013), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T. F., m.fl. (red.)]. https://www.ipcc.ch/report/ar5/wg1/.

McGlade, C. and P. Ekins, (2015), The geographical distribution of fossil fuels unused when limiting global warming to 2 °C, *Nature* 517, January.

Matthews, H. Damon, Nathan P. Gillet, Peter A. Stott and Kirsten Zickfeld (2009), "The proportionality of global warming to cumulative carbon emissions", Nature, Vol 459 June, p.829–33.

Nordhaus, W. D. (1994), Managing the Global Commons: The Economics of Climate Change. Cambridge, MIT Press.

Nordhaus, W. and A. Moffat (2017), A survey of global impacts if climate change: replications, survey methods and a statistical analysis. NBER Working Paper 23646.

Pigou, A. C., (1920), "The Economics of Welfare," Macmillan, London.

Taxes, benefits and labour force participation: A survey of the quasi-experimental literature

Jacob Lundberg[1] and John Norell[2*]

Abstract

We review the literature that uses quasi-experimental methods to estimate the elasticity of labour force participation with respect to the financial gain from work. We find a wide range of elasticities, with an average of 0.36. 27 out of 35 papers find elasticities larger than 0.1, providing strong evidence that individuals respond to incentives on the extensive margin of labour supply. Elasticities are larger for women, and have declined over time.

Keywords: participation elasticity, quasi-experimental methods, labour supply, extensive margin
JEL codes: H24, J22

1 Corresponding author. Ratio Institute, PO Box 3203, 103 64 Stockholm, Sweden. Email: jacob.lundberg@ratio.se.

2 Department of Economics, Stockholm University, 106 91 Stockholm, Sweden. Email: john.norell@ne.su.se.

* We would like to thank Bryan Caplan, Eva Uddén Sonnegård and two anonymous referees for helpful comments, as well as participants at a Ratio seminar on 24 August 2018.

1. Introduction

The effect of taxes and benefits on labour supply is a central topic in economics. Working entails loss of leisure, which individuals trade off against the monetary reward from working. Therefore, financial incentives matter for the labour supply decisions of individuals. The labour supply decision consists of the intensive and extensive margins. The intensive margin is the choice of hours of work, given that the individual is working. In this case, the marginal tax rate is what matters. The extensive margin, which is the subject of this paper, concerns the choice of working or staying out of the labour force completely, i.e., the participation decision.

Both taxes and social benefits matter for the participation decision and these are summarized by the participation tax rate. Theoretically, a higher participation tax rate leads to a lower rate of labour force participation and employment. The strength of this response is measured by the participation elasticity, which shows the percentage increase in labour force participation when the financial gain from work (the difference between after-tax wages and out-of-work benefits) increases by one percent.

This paper surveys the literature that uses quasi-experimental methods to estimate the participation elasticity.[3] Quasi-experimental methods use natural experiments and non-structural econometric methods, such as difference-in-differences, regression discontinuity design and instrumental variables, to estimate plausibly causal elasticities.

There is a larger and older structural labour supply literature which we do not survey, as it has been extensively reviewed several times before.[4] Researchers have increasingly moved to using quasi-experimental methods because the older literature, which typically uses cross-sectional data, suffers from potential internal validity problems: Strong assumptions are needed in order to give the estimated elasticities a causal interpretation. People with higher (potential) after-tax wages participate in the labour force to a greater extent, but this may be caused by unobserved tastes for work. If highly motivated people are more inclined to work and also have higher earnings potential (conditional on observed covariates), elasticity estimates from cross-sectional studies will be biased upwards.

While studies using a structural approach are often criticized on the issue of internal validity, quasi-experimental studies may be weaker in external validity – the concern that results from any given research design only applies to the specific context of that study. The way to deal with this issue, argue Angrist & Pische (2009), is the accumulation of more evidence from different contexts, so that more general conclusions can be drawn, with some claim to external validity. This critique highlights the need for literature reviews, such as this one, in order to find policy-relevant parameter estimates.

We are aware of only one previous survey of the quasi-experimental extensive margin literature: Chetty et al. (2013), which cites 15 papers.[5] We improve upon Chetty et al. (2013) by including a larger number of papers (35) and more recently published studies. In addition, we are stricter in what we deem to be quasi-experimental methods. We also only include papers that use policy as identifying variation (excluding papers that rely on before-tax wage trends, for example). This leads to our exclusion of eight of the studies in Chetty et al. (2013). Our motivations for each paper are stated in the appendix.

Of the 35 papers, about half use difference-in-differences methodology, for example comparing mothers benefiting from an in-work tax credit reform (such as the American earned income tax credit) to women without children. Most of the other papers use panel regression techniques such as fixed effects or correlated random effects (CRE).

---

[3] *Research using quasi-experimental methods to estimate the elasticity on the intensive margin have already been reviewed previously by Saez, Slemrod & Giertz (2012) and Neisser (2017).*

[4] *E.g., Meghir & Phillips (2010), Keane (2011) and McClelland & Mok (2012). Bargain & Peichl (2016) include a small number of quasi-experimental studies in their survey, though they do not make a clear distinction in their analysis.*

[5] *Bettendorf, Folmer & Jongen (2014) provide a more limited survey of the effects of in-work tax credits, citing eight papers. Hotz & Scholz (2003) survey the American EITC literature.*

The studies find a wide range of elasticities. Six papers find elasticities close to zero, while four arrive at elasticities around 1 or larger. The average elasticity is 0.36. Groups with low employment rates tend to exhibit larger participation responses. This is expected, since it implies a larger number of individuals who could potentially enter the labour market. As many papers focus on these groups, the policy-relevant full-population elasticity is likely lower than 0.36; we assess it to be in the range 0.1–0.2.

We find that elasticities have declined over time, possibly due to increased female labour force participation. Americans seem to respond more strongly to incentives than Europeans. We also conclude that papers that use a difference-in-differences methodology find larger elasticities.

This paper is structured as follows. The next section reviews the theory underlying the participation elasticity, clarifying its definition. In section 3, we review the quasi-experimental literature on labour participation responses to financial incentives. In section 4, we conduct a meta-analysis of the elasticity estimates. Section 5 concludes.

2. Theory

As there are different definitions of the participation elasticity used in the literature, it is important to be clear about the definition we are using and the theoretical basis behind it.

In this section, we set up a simple theoretical model. We assume that individuals choose between working for a gross wage Y or not participating in the labour force, receiving a benefit B. Those who work pay an income tax denoted T. Utility is equal to disposable income, except that working individuals also incur a fixed cost of work q (expressed in money terms). This captures all the monetary and nonmonetary costs associated with working: loss of leisure, commuting and childcare costs etc.[6]

Individuals are identical in all dimensions except q. We can think of our model as applying to a subset of the labour force, such as low-income single mothers, where incomes and tax rates are similar. We hold Y constant, thus abstracting from the intensive margin.

Utility for workers is given by $u_w = Y - T - q$ and for non-workers by $u_{nw} = B$. Individuals work if

$$u_w > u_{nw} \Leftrightarrow Y - T - q > B \Leftrightarrow Y - T - B > q.$$

Thus, the individuals who work are those for whom the financial gain from work, $Y - T - B$, exceeds the fixed cost of work, q. As individuals only are heterogenous in q, the distribution of fixed costs of work will determine the rate of labour force participation. We denote the probability density function $f(q)$. The rate of labour force participation can be expressed

$$L = \int_0^{Y-T-B} f(q)dq.$$

[6] *The concept of fixed costs of work was introduced by Cogan (1981). It explains why individuals not only adjust their labour supply at the (intensive) margin, but also switch from not working at all to working a significant number of hours.*

If the financial gain increases – due to lower taxes or benefits – some individuals will be incentivized to enter the labour force. To be precise, how many individuals will start working is determined by the density of fixed costs of work evaluated at the financial gain from work, i.e., those who are at the margin of entering employment:

$$\frac{dL}{d(Y-T-B)} = f(Y-T-B).$$

In the literature, the strength of this response is typically measured by the elasticity of labour force participation with respect to the financial gain from work, which can be expressed

$$\varepsilon_P = \frac{dL/L}{d(Y-T-B)/(Y-T-B)} = f(Y-T-B)\frac{Y-T-B}{L},$$

where the second step is the general definition and the third step applies in the context of our theoretical model. Our derivations illustrate that the participation elasticity is a population-level, not individual-level, parameter.

Using this elasticity definition has several advantages. First, it is now the most common elasticity definition in the literature. Thus comparison between countries, time periods and reforms is simplified. Second, the participation elasticity is a crucial parameter in optimal income tax models (e.g., Saez, 2002), which weigh the disincentive effects of income taxation against the benefits of redistribution. Third, quantifying the participation elasticity allows for ex-ante evaluation of policy reforms (see, e.g., Lundberg, 2017).

A related concept is the participation tax rate. This shows the percentage of the gross wage that an individual has to pay to the government in the form of income tax and foregone benefits: $\tau = (T + B)/Y$. The participation net-of-tax rate ($100\% - \tau$) is the financial gain from work expressed as a proportion of the gross wage.[7] So if the financial gain from work increases by a given percentage, so will the participation net-of-tax rate. Therefore the participation elasticity can also be termed the elasticity of participation with respect to the participation net-of-tax rate.

Some papers use a different elasticity definition: the elasticity of participation with respect to the net wage (or, equivalently, the average net-of-tax rate), denoted

$$\varepsilon_Y = \frac{dL/L}{d(Y-T)/(Y-T)}.$$

The two elasticity definitions are equivalent if out-of-work benefits are zero. For, e.g., secondary earners, this may not be far from the truth, so the choice of elasticity concept will not matter much. Note that $\varepsilon_Y \geq \varepsilon_p$ for a given size of the derivative. If benefits are significant, the choice of elasticity definition can yield very different estimates. For example, if benefits correspond to three quarters of after-tax labour income, the elasticities will differ by a factor of four.[8] As we include papers using both elasticity definitions, the exact elasticity concept used should be considered before drawing conclusions.

Theoretically, the participation elasticity is concerned with the impact on labour force participation, but the empirical literature typically analyzes employment because it is easier to observe at the individual level. Employment is usually what is relevant for policy purposes. In addition, when analyzing the full population, the only reasonable assumption is that Say's law will hold – supply will create its own demand such that increased labour force participation will translate into higher employment. Simply scaling up the labour market should not change any fundamentals, such as the unemployment rate, in the long run.

---

[7] *The participation net-of-tax rate can be expressed* $1-\tau = 1-\frac{T+B}{Y} = \frac{Y-T-B}{Y}$. *The numerator in the last step is the financial gain from work.*

[8] *Some algebraic manipulation reveals that* $\varepsilon_Y = \varepsilon_P(Y-T)/(Y-T-B)$. *Setting* $B = 0.75(Y-T)$ *yields* $\varepsilon_Y = 4\varepsilon_P$.

Table 1. Summary of participation elasticity estimates from quasi-experimental studies

Study	Country	Years	Identifying variation	Method	Sample	Elasticity
Alpert & Powell (2014)	US	1999–2009	Bush tax cuts	FE, IV	(a) women, (b) men, aged 55–71	0.75 (a) 0.55 (b)†
Bastian (2020)	US	1975	EITC introduction	DD	Mothers	0.58
Brown (2013)	US	1999	Pension reform	DB	California teachers near retirement	0.04
Chetty, Friedman & Saez (2013)	US	2000–2005	Geographical variation in EITC knowledge	IV	EITC-eligible parents	0.19†
Eissa & Hoynes (2004)	US	1984–1996	EITC expansions	Group IV	Married couples (aged 25–54) with children: women (a), men (b)	0.27 (a) 0.03 (b)†
Eissa (1995)	US	1987	Tax Reform Act of 1986	DD	High-income married women	0.4–0.6†
Eissa (1996)	US	1982	Economic Recovery Tax Act of 1981	DD	Married women aged 19–64	0.33–0.91†
French & Song (2014)	US	1990–1999	Random assignment of disability insurance judges	IV	Disability insurance applicants	1.53
Gelber & Mitchell (2012)	US	1975–2004	“variation across individuals and time in national and state policy changes”	FE	Singles aged 25–55: (a) women, (b) men	0.41 (a) -0.04 (b)†
Gelber et al. (2017)	US	1978–1987	Social Security earnings test	RKD	Retirees born 1918–1923: (a) all, (b) men, (c) women	0.49 (a) 0.25 (b) 0.49 (c)†
Hotz, Mullin & Scholz (2002)	US	1987–1998	EITC expansions	DD	California AFDC recipients	0.97–1.69†
Kumar & Liang (2016)	US	1998–2006	Over-time variation in tax rates and wages	CRE	Married women	0.35†
Kumar (2016)	US	1979–2007	Over-time variation in tax rates and wages	CRE IV	Married women	0.56†
Lin & Tong (2017)	US	2000–2009	Bush tax cuts, Obama recovery package	IV/IV-FD	Married couples aged 25–54: (a) men, (b) women	0.03/-0.01 (a) 0.10/0.08 (b)†
McClelland, Mok & Pierce (2014)	US	1999–2010	Bush tax cuts, state tax reforms	IV	(a) women, (b) men, born 1948–1978	0.02 (a) 0.004 (b)†
Milligan & Stabile (2007)	Canada	1998	Provincial variation in interaction between social assistance and National Child Benefit	DD	Single mothers aged 18–50	0.96
Bartels & Shupe (2018)	several	2008–2014	policy changes affecting demographic groups differently	Group IV	(a) women, (b) men, aged 25–54	0.14 (a) 0.08 (b)
Jäntti, Pirttilä & Selin (2015)	several	1970–2010	”compare otherwise similar groups of individuals who have been affected differently by tax reforms”	Group IV	Individuals aged 25–64	0.01
Blundell, Bozio & Laroque (2011)	UK	1978–2007	“differential changes across gender and education”	Control function	Individuals aged 34–54: (a) women, (b) men	0.34 (a) 0.25 (b)†
Meghir & Phillips (2010)	UK	1994–2004	Regional variation in housing benefit over time	IV	(a) single men, (b) married men, aged 22–59, low education	0.27 (a) 0.53 (b)†
Bettendorf, Folmer & Jongen (2014)	Nether-lands	2002	Reform of single parent tax credit	(a) DD, (b) RD	Single mothers	–0.02 (a) –0.02 (b)
Bastani, Moberg & Selin (2020)	Sweden	1997	Housing benefit reform	DD	Married low-income women	0.13
Laun (2017)	Sweden	2007	EITC and payroll tax cut for older workers	DD	65-year-olds	0.22
Selin (2014)	Sweden	1971	Abolition of joint taxation of spouses	DD	Married women	1
Kosonen (2014)	Finland	1994–2005	Municipal variation in Home Care Allowance	DD	Mothers	0.83
Martinez, Saez & Siegenthaler (2018)	Switzerland	1997–2003	Swiss tax holiday	FE	20–60-year-olds	0
Sigurdsson (2019)	Iceland	1987	Icelandic tax holiday	DD	16–70-year-olds	0.1
Stefansson (2019)	Iceland	1987	Icelandic tax holiday	DD	16–67-year-olds	0

Study	Country	Years	Identifying variation	Method	Sample	Elasticity
*Papers where the elasticity is calculated by other authors**:						
Eissa & Liebman (1996)	US	1987	Tax Reform Act of 1986	DD	Single mothers	0.3
Meyer & Rosenbaum (2001)	US	1984–1996	Tax reforms 1984–1996	DD	Single mothers	0.43
Card & Hyslop (2005)	Canada	1992–1995	Self-Sufficiency Project	RCT	Single parents	0.38†
Blundell, Brewer & Shephard (2005)	UK	1999	WFTC	DD	Single mothers	0.45
Francesconi & van der Klaauw (2007)	UK	1999	WFTC	DD	Single mothers	0.6
Gregg & Harkness (2003)	UK	1999	WFTC	DD	Single mothers	0.61
Leigh (2007)	UK	1999	WFTC	DD	Single mothers	0.07

Abbreviations:

EITC – earned income tax credit

WFTC – working families tax credit

DD – difference-in-differences

IV – instrumental variables

FE – fixed effects

FD – first differences

CRE – correlated random effects

RD – regression discontinuity

RKD – regression kink design

RCT – randomized controlled trial

DB – difference-in-bunching

* Chetty et al. (2013) or Bettendorf, Folmer & Jongen (2014). See text for details.

† The elasticity is expressed with respect to net wages instead of the financial gain from work.

3. A review of the quasi-experimental literature

Since the 1990s, an increasing number of papers use quasi-experimental methods to identify economic parameters, including the effect of taxes and benefits on labour supply. These papers, made possible by improved data access and econometric innovation, are primarily concerned with finding elasticity estimates that plausibly can be given a causal interpretation. The literature is called quasi-experimental because it strives to come as close as possible to the ideal of a randomized experiment. Because such experiments are rare in the social sciences, the literature uses real-world features, like reforms affecting groups differently, to estimate responses to policy changes. The methods most commonly used are difference-in-differences (DD), instrumental variables and regression discontinuity (see Angrist & Pischke, 2009, for a general description). Some papers use panel data with individual or group fixed effects. This is similar to DD in that it uses changes over time within individuals or groups to identify an elasticity.

We have identified 35 papers that use quasi-experimental methods to identify participation responses; see table 1.[9] We have included all papers that we could find that fulfil our basic criterion – estimating a participation elasticity with quasi-experimental methods, using tax or benefit policy changes as the identifying variation. We include both journal articles and working papers.

As explained in the theory section, the ideal elasticity concept is the elasticity of labour force participation with respect to the financial gain from work. However, many papers instead report the elasticity with respect to the net wage, perhaps due to transfers being unobserved. As this is fairly common, we include such papers as well, although the estimates may be biased due to out-of-work benefits being omitted. Also recall that the elasticity with respect to the net wage is always larger than the elasticity with respect to the financial gain for a given participation response. Therefore the elasticity definition should always be considered when drawing conclusions from the papers. Studies that use a different elasticity definition than the two mentioned are not included in this survey as these estimated elasticities are not directly comparable.[10]

In seven cases, the paper does not itself report a participation elasticity. Instead, we report elasticities calculated by other authors (Bettendorf, Folmer & Jongen, 2014, or Chetty et al., 2013) using information in the papers.

The papers are summarized below. We group them by country and introduce the various econometric methods throughout the text. We start with the American literature, which is by far the largest and most diverse.

United States

The earliest papers in the quasi-experimental extensive margin literature use difference-in-differences methodology to estimate how American tax reforms affected labour force participation, especially among women. Of particular interest is the earned income tax credit (EITC), which is targeted at low-income workers with children and was increased several times during the 1980s and 90s. Hotz & Scholz (2003) survey the literature that estimate extensive margin responses to the EITC. One such paper is Eissa & Liebman (1996), who estimate that single mothers increased their labour force participation by 2.8 percentage points following the expansion of the EITC after the Tax Reform Act of 1986 compared to childless single women and controlling for demographic characteristics. Chetty et al. (2013) calculate that this implies a participation elasticity with respect to the financial gain from work of 0.3.[11]

[9] *We found the papers by searching for "participation elasticity" and "extensive margin elasticity" together with "labour supply" on Google Scholar, and from references in other papers.*

[10] *Some examples of such papers are Autor et al. (2016), Bargain & Doorley (2011), Carbonnier (2008), Jonassen (2013), Fadlon & Nielsen (2015), Gruber (2000), Koning & van Sonsbeek (2017) and Kostol & Mogstad (2014). Most of these papers focus on groups with weak labour force attachment and find quite sizeable responses, in line with the papers included in our analysis.*

[11] *Hotz & Scholz (2003) calculate a participation elasticity of 1.16 from the same paper. The difference is due to two factors. First, Hotz & Scholz (2003) define the elasticity with respect to after-tax wages while Chetty et al. (2013) define it with respect to the financial gain from work. Second, Chetty et al. (2013) use a $1,000 tax cut in the denominator while Hotz & Scholz (2003) use $500. We choose to report the more conservative estimate.*

Hotz & Scholz (2003) report elasticities calculated from an unpublished study (Hotz, Mullin & Scholz, 2002) that uses data from California to analyze the 1990s EITC expansion. Making use of the fact that the expansion increased the return to work more for families with two or more children compared to one-child families, they estimate a participation elasticity between 0.97 and 1.69 depending on base year.

Bastian (2020) analyzes the introduction of the EITC in 1975. He shows that the employment rate of mothers increased significantly after 1975 compared to women without children, corresponding to a participation elasticity of 0.58 for a representative mother.

The EITC literature has recently been criticized by Kleven (2019), who shows that only the 1993 EITC expansion is clearly associated with an increase in the employment rate of single mothers. Further, Kleven argues that the employment increases of the 1990s can be better explained with macroeconomic conditions and welfare reform.

Eissa (1995) estimated a participation elasticity with respect to net wages of 0.4–0.6 for high-income women by examining the Tax Reform Act of 1986, using lower-income women as the control group. However, Liebman & Saez (2006) criticize this approach, arguing that lower-income women cannot serve as a control group and showing that the estimated effect (and therefore elasticity) varies greatly depending on which reference years are chosen.

In a related study, Eissa (1996) investigates the 1981 Kemp–Roth tax cut (the Economic Recovery Tax Act) using the same methodology. Comparing women married to husbands earning more than $50,000 to those whose husbands earned $30,000–50,000, she arrived at elasticities ranging from 0.33 to 0.91 depending on how the control group is formed and whether education-specific time trends are included. Because married couples are taxed jointly in the United States, the husband's income affects the wife's participation tax rate. As the tax cut flattened the tax structure, reducing marginal tax rates more for high-income couples, the fact that high-income women increased their labour force participation is evidence of their responding to the greater incentives for work.

Eissa & Hoynes (2004) use a repeated cross-section (the Current Population Survey) to examine how married Americans responded to tax reforms, notably several EITC expansions, over the period 1984–1996. Utilizing differences across demographic characteristics (such as number of children), they estimate participation elasticities of 0.27 for women and 0.03 for men.

Meyer & Rosenbaum (2001) use the same dataset and analyze the same time period, but instead focus on single women. They find an elasticity of participation with respect to gross wages of 1.07. However, as pointed out by Chetty et al. (2013), the elasticity should be expressed with respect to the increase in net earnings. They recalculate the elasticity to be 0.43.

Chetty, Friedman & Saez (2013) analyze the effects of the EITC using a different approach. They note that the EITC needs to be claimed by the taxpayer on the tax return and that take-up is not perfect. Further, they find evidence of substantial geographical variation in EITC knowledge across the United States. They do this by noting how many self-employment EITC filers – who have some freedom in how much income to report – locate exactly at the beginning of the plateau where the EITC is maximized, so-called bunching. If many small-business owners in a particular area bunch at this kink point, this indicates relatively widespread knowledge about the EITC. Thus having constructed an instrument for EITC take-up, the authors proceed to estimate a participation elasticity of 0.19.

More recently, it has become easier for researchers to use panel data of individuals to estimate labour supply elasticities. Using panel data can potentially alleviate the problem of unobserved individual heterogeneity by including individual fixed effects (FE) in the regression, implying that only within-individual variation over time is used to identify the elasticity.

One such paper is Gelber & Mitchell (2012), which examines the participation decisions of unmarried prime-age Americans during 1975–2004. The fact that they include individual fixed effects implies that the variation used is tax reforms that affected individuals differently. They find that a one percent increase in net wages raises the labour force participation of single women by 0.43 percent. In alternative specifications, the elasticity varies between 0.26 and 0.75. However, for men the elasticity is slightly negative.

There are econometric difficulties (the incidental parameters problem) associated with nonlinear fixed effects models – such as probit, often used to model labour force participation – when the number of time periods is relatively small. A common technique for avoiding this is correlated random effects, CRE. This can be described as being in between random effects and fixed effects. CRE requires a few additional assumptions about individual heterogeneity.

Kumar (2016) uses the Panel Study of Income Dynamics to study how married women responded to tax reforms (as well as variation in wages) over the period 1979–2007. He reports results for both CRE and FE, as well as pooled panels without controls for unobserved individual heterogeneity. CRE is his preferred specification, but the FE and pooled regressions yield elasticities of a similar magnitude. However, it makes a big difference whether the endogeneity of after-tax hourly wages is accounted for. In Kumar's preferred specification, this is done by using lagged demographic variables as instruments. The participation elasticity thus estimated is 0.56 in a lifecycle model and 0.46 in a static model.

In a similar paper, Kumar & Liang (2016) study the same sample, also focusing on married women. However, instead of estimating one elasticity for the entire time period, like Kumar (2016), they look for evidence of changing elasticities over time. In the CRE specification, they find an elasticity of 0.53 in the first period, 1980–1984, increasing to 0.83 in 1984–1989. After that the elasticities are lower, around 0.35.

One strand of the literature has taken inspiration from the new tax responsiveness literature (in particular Gruber & Saez, 2002) which estimates the intensive margin elasticity on individual panel data using quasi-experimental methods. An econometric problem when estimating this elasticity is that when the income tax is progressive, the marginal tax rate will depend on taxable income, causing endogeneity. Gruber & Saez (2002) handle this problem by instrumenting for the current year marginal tax rate with last year's income and marginal tax rate.

Alpert & Powell (2014) implement this so-called simulated instruments methodology to examine how the 2001 and 2003 Bush tax cuts affected the labour supply of workers aged 50 or older, who may be on the margin of retirement. They find relatively high participation elasticities: 0.75 for women and 0.55 for men.

McClelland, Mok & Pierce (2014) study the same time period and use the same methodology, but instead look at secondary earners within prime-age married couples. They find very little evidence of participation responses to the Bush tax cuts, estimating elasticities close to zero (0.03 at most). In the main analysis, they control for individual heterogeneity using correlated random effects. They report results for a fixed effects model as a robustness check, but the magnitude of elasticities is similar.

In a very similar paper, Lin & Tong (2017) study the same group using the same reforms as identifying variation, but use a larger sample and a slightly different method. They also find small elasticities, very near 0 for men and at most 0.1 for women.

Gelber et al. (2017) examine the labour supply of Americans in their 60s using a feature of the old-age part of the Social Security system, the annual earnings test. For every dollar a retiree's earnings exceeds $17,000, retirement benefits decrease by 50 cents. The authors show that labour force participation among retirees is increasing with prior earnings, but that the relationship has a noticeably smaller slope after the earnings test threshold. This is evidence of older workers with relatively high incomes dropping out of the labour force as a result of the Social Security annual earnings test. The method that uses the change in the slope of the treatment variable for identification is called regression kink design. Gelber et al. (2017) arrive at a participation elasticity of 0.49.

French & Song (2014) analyze a different part of the Social Security system – disability insurance. Americans who apply for disability benefit from the Social Security Administration but are denied can appeal to an administrative court. Assignment of cases to judges is essentially random, and judges vary considerably in their willingness to grant an appellant disability benefit. This can be used to estimate the effects of disability insurance on labour supply. The authors find that the effects are very large: labour force participation falls by 26 percentage points after disability benefit has been granted, corresponding to an elasticity of 1.53. The elasticity is lower for older and college-educated individuals.

Brown (2013) looks at the retirement behaviour of California public school teachers. She uses a difference-in-bunching design. As Kleven (2016) explains, this is a method that takes advantage of a change in the size of a kink or notch in the taxation policy. By observing bunching around the discontinuity before and after the policy change, a labour supply elasticity can be calculated. Retired teachers receive a higher benefit the more years that they work. Brown uses two nonlinearities in the determination of retirement benefits for identification: First, after a certain age the benefit amount increases by less for each year. Second, teachers with 30 years of service receive a retirement bonus. She shows that teachers adjust their behaviour very little in response to these discontinuities, which implies an elasticity close to zero.

Canada

Milligan & Stabile (2007) analyze an EITC-type programme, the National Child Benefit, introduced in Canada in 1998. Variation across provinces, as well as the fact that the benefit amount depends on the number of children, is used to estimate the effect on labour force participation. They find that single mothers responded strongly to the increased incentives for work, arriving at a participation elasticity of 0.96.

In the 1990s, Canada ran a large-scale randomized trial of work incentives for welfare recipients, the Self-Sufficiency Project. Out of a sample of 5,000 individuals, half were randomly assigned to the project. If they started full-time work within a year, they received a generous benefit. Card & Hyslop (2005) show that the effects of the experiment were large: After one year, the treatment group had a 14 percentage points higher employment rate than the control group. Chetty et al. (2013) calculate that this implies a participation elasticity of 0.38. After the experiment ended, there was no longer any difference in outcomes between the treatment and control groups.

Cross-national studies

Another method borrowed from the intensive margin literature (Blundell et al., 1998) that is used by a number of papers is group instrumental variables (IV). The idea is to divide the sample into groups by, e.g., age, education and gender, and use group membership as an instrument for tax rates or net wages. This is equivalent to simply running a regression on group averages. The method is similar to difference-in-differences.

Jäntti, Pirttilä & Selin (2015) apply this method to a cross-national dataset of 13 countries (the United States, Canada, Australia, Israel and nine European countries). They create 1,200 groups based on country, age, education and gender. When running a regression across all countries and years without any controls, they estimate an elasticity of 0.2. However, this estimate could be biased due to changes over time or differences between countries that are unrelated to taxation. When they add group and year fixed effects, the elasticity is reduced to only 0.01.

In a related paper, Bartels & Shupe (2018) perform a group IV regression on 12 EU countries over the period 2008–2014. Defining groups as Jäntti, Pirttilä & Selin (2015), they find an average elasticity of 0.14 for women and 0.08 for men.

Britain and the Netherlands

Blundell, Bozio & Laroque (2011) use a group approach to estimate participation elasticities in the United Kingdom. The groups are defined by gender and education level, and differential changes in after-tax wages between these groups over time are used to identify the elasticity. They find an elasticity of 0.25 for prime-age men and 0.34 for prime-age women.

In Britain, the Working Families Tax Credit (WFTC) was introduced in 1999 with the purpose of raising the employment rate of lone parents and reducing child poverty. This reform is analyzed by Gregg & Harkness (2003), Blundell et al. (2005), Francesconi & van der Klaauw (2007) and Leigh (2007). The papers examine the labour supply response of single mothers, using single women without children as a control group. While none of the papers report participation elasticities, they are calculated by Bettendorf et al. (2014) to be 0.61, 0.45, 0.6 and 0.07, respectively.[12] The outlier is Leigh (2007), who uses a considerably shorter follow-up period than in the other papers.

Meghir & Phillips (2010) analyze the labour supply behaviour of British men. In identifying the elasticity they make use of the fact that housing benefit is tied to the level of rent, which has varied over time across regions of the UK. Using this as an instrument for net income when working, they estimate an elasticity of 0.27 for single men and 0.53 for married men, when restricting the sample to men with low education. For men with high education, the estimates are not significantly different from zero.

Bettendorf, Folmer & Jongen (2014) study an extension of eligibility of the EITC in the Netherlands using a difference-in-differences approach as well as a regression discontinuity method. Before 2002, only those single parents who had a child aged 13 or less were eligible for the EITC. This cut-off was increased by four years in 2002. In their DD analysis, they compare the labour supply of single mothers with children aged 12 to 16 years with single mothers who had older or younger children. In the regression discontinuity analysis, the effect is estimated by analyzing mothers to children just above and below the cut-off point of 16 years of age. The cut-off creates a discontinuity that can be used for identification. None of the methods find any evidence of participation responses.

Sweden and Finland

In many countries, spouses are taxed jointly, which combined with a progressive tax schedule raises the participation tax rate for the secondary earner in the household, which often affects the labour supply of married women. In 1971, Sweden transitioned from taxing married couples jointly to taxing them separately. Selin (2014) analyzes this reform, noting that it increased work incentives for secondary earners. The incomes of husbands affect to what extent the policy change creates an incentive for their wives to enter the labour market. By comparing women married to high- and low-income earners, he estimates the elasticity to be 1, with a higher elasticity for women with children.

In a similar reform, the Swedish housing benefit was altered in 1997 to be based on individual rather than household income. In practice, this resulted in lower housing benefit for one-earner couples and unchanged benefit levels for two-earner couples. The participation tax rate for secondary earners (usually women) thus fell. Bastani, Moberg & Selin (2020) examine how this affected labour supply. Comparing low-income mothers, who are eligible for the benefit, with low-income women without children, who are ineligible, they show that the labour force participation of the former group increased in the years following the reform, corresponding to an elasticity of 0.13.

Sweden introduced an EITC in 2007, but because this tax credit is payable to all workers, no natural control group exists and the reform has not been possible to evaluate using quasi-experimental methods. (Edmark et al., 2016) However, workers over 65 are eligible for a larger EITC, as well as lower payroll taxes – a reform which was also implemented in 2007. Laun (2017) uses those born during the previous calendar year, and thus ineligible for the two tax breaks, as a control group and finds that the reform raised employment in the treatment group. The effect implies a participation elasticity of 0.22.

Kosonen (2014) studies the Finnish Child Homecare Allowance (HCA), a benefit system offered to mothers who stay home to care for their children. He exploits variation over time in the municipality-specific component of the HCA. Using a difference-in-differences methodology, the participation elasticity for mothers is estimated at 0.86. Kosonen further

[12] *Francesconi & van der Klaauw (2007) calculate an elasticity with respect to net income of 1.1.*

concludes that the participation elasticity is highest for mothers with low and high education while being lower for individuals with a medium education level.

Iceland and Switzerland

In three recent papers, economists have used so-called tax holidays to identify a participation elasticity. Such a tax holiday occurred in Iceland in 1987. Up until 1987, Icelanders paid taxes on last year's income, so the 1987 tax liability was calculated on 1986 earnings. However, in 1988 the tax collection system was changed so that taxes were paid on the current year's income. Thus the 1988 tax liability was based on 1988 incomes – and 1987 incomes were never taxed. Sigurdsson (2019) and Stefánsson (2019) both analyze this tax holiday.

Stefánsson (2019) shows that the participation rate – defined as the proportion having positive labour income – did not deviate from the trend in 1987, implying a participation elasticity of zero.[13]

Sigurdsson (2019) uses a difference-in-differences design, comparing individuals in different tax brackets before the reform. Theory predicts that individuals paying higher tax rates should increase their labour force participation more, as they receive the largest tax cut when the tax rate falls to zero. However, Sigurdsson finds no such differences between tax brackets, also implying an elasticity close to zero (in fact slightly negative).

The difference-in-differences method can only observe labour market exits, as entrants had no income in the year before the reform. In order to capture labour market entries, Sigurdsson constructs a life-cycle model, using adjacent cohorts as control groups. Hence he arrives at an extensive margin elasticity of 0.1.[14]

In the late 1990s and early 2000s, Switzerland also transitioned to taxing current incomes, although the year of transition, and therefore of the tax holiday, differed by canton. Martinez, Saez & Siegenthaler (2018) analyze this reform and find no participation responses.

The tax holiday papers use a very convincing identification strategy, but the external validity may be questioned. In the standard model, people respond more to temporary than to permanent policy changes – the Frisch elasticity is greater than the Marshallian elasticity. However, it is easy to think of optimization frictions that make it difficult to participate in the labour market for just one year (and for employers to temporarily increase their labour force). Therefore, the low participation elasticities estimated may not be surprising.

4. Meta-analysis

There is a great deal of variation in the cited estimates. The literature is a long way from consensus. Nonetheless, a few conclusions can be drawn. There is evidence that people respond to incentives when deciding whether to work. 27 of the 35 studies find an elasticity larger than 0.1, at least for women. Women respond more strongly than men. All studies that report elasticities disaggregated by gender find a larger elasticity for women.

The estimates are summarized in figure 1. The 35 papers report 45 elasticities in total. The mean is 0.36 and the median is 0.27. Elasticities seem to have declined over time, consistent with the findings of Heim (2007), Blau & Kahn (2007) and Kumar & Liang (2016). Figure 2 shows a downward trend of about 0.11 per decade, which is substantial. A likely

[13] *Stefánsson (2019) and Sigurdsson (2019), as well as an earlier paper by Bianchi, Gudmundsson & Zoega (2001), find that working individuals increased the number of weeks worked during 1987. However, this is properly classified as an intensive margin response and therefore we do not include it in the table.*

[14] *This can be calculated from table 6 in Sigurdsson (2019) by dividing the estimated semi-elasticity (0.07) by the employment rate (0.67).*

explanation is increased female labour force participation, whereby the available pool of nonworkers has shrunk over time.

The same decreasing trend can be seen when analyzing estimates by publication year. However, we would be careful in drawing conclusions from this, as the studies were published during a relatively short time period, and the methods and study populations have varied over time.

Table 2 shows unconditional averages for different groups of papers. We see that elasticities are larger in North America. Women's elasticities are greater than men's. Married women (in this category we also include estimates pertaining to all women) seem to respond more than single women, possibly because they are typically the secondary earner in a couple, whose participation decision is more responsive to incentives.

Studies using difference-in-differences (DD) methodology find larger elasticities. There are several possible explanations for this, for example, that DD papers concentrate on the reforms where responses are likely to be the highest (e.g., EITC reforms targeted at single mothers).

Surprisingly, elasticities that are expressed with respect to net wages are on average lower than elasticities with respect to the financial gain from work, although, as shown in the theory section, the former definition always yields larger elasticities for a given magnitude of the participation response. The explanation could be that the use of this elasticity definition is correlated with unobserved study characteristics that cause a high estimate. It could also be due to chance.

Figure 1. Histogram of elasticity estimates (45 estimates from 35 papers)

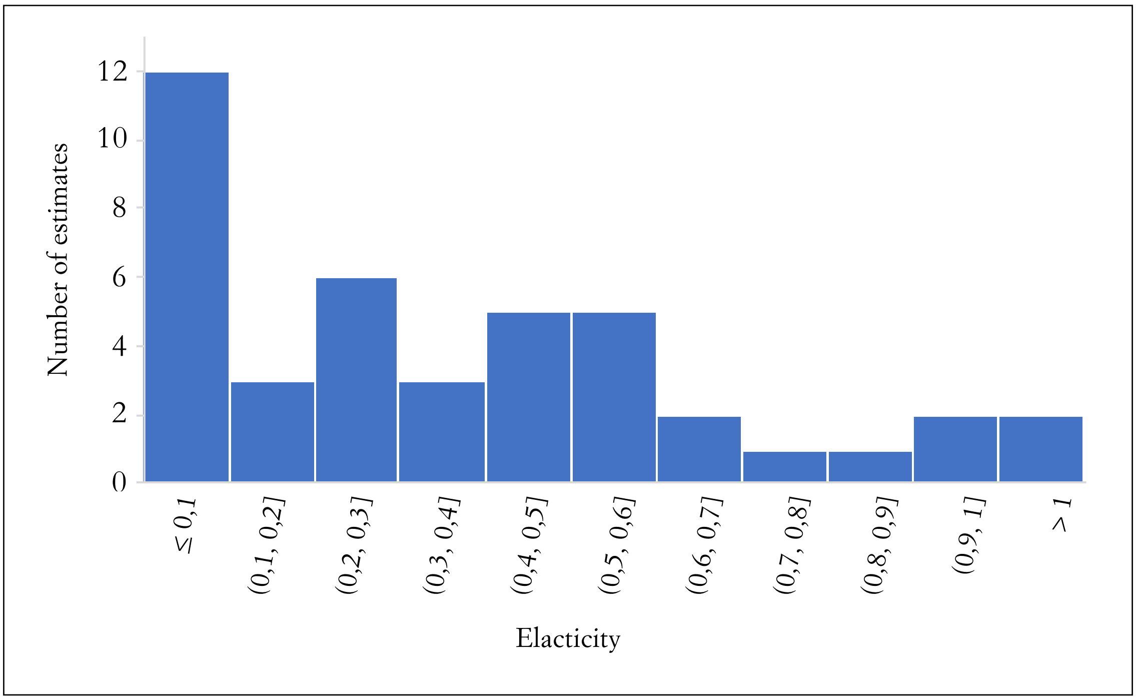


Figure 2. Elasticity estimates by year of the reform evaluated (or midpoint for studies spanning several years)

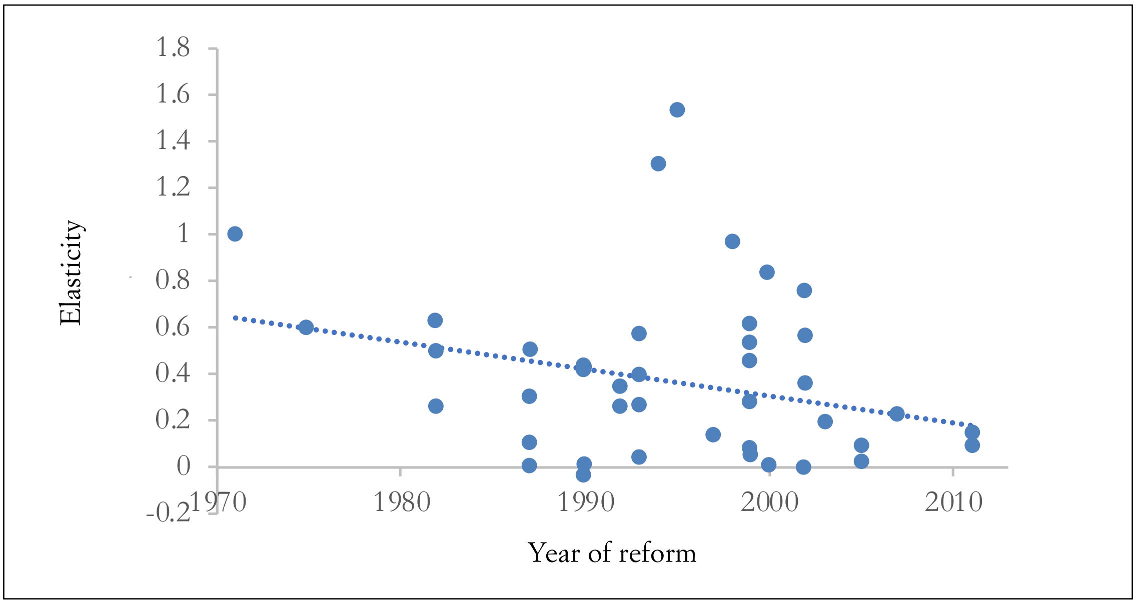


Table 2. Unconditional elasticity averages

Continent	Europe	North America	
	0.28 [20]	0.42 [25]	
Gender	Single women	(Married) women	Men/both genders
	0.4 [12]	0.44 [14]	0.28 [19]
Age	Older workers	Working age	
	0.38 [6]	0.36 [39]	
Elasticity denominator	Net wage	Financial gain from work	
	0.35 [23]	0.37 [22]	
Methodology	DD	Other methodologies	
	0.45 [20]	0.29 [25]	

Number of elasticity estimates in square brackets.

For policy purposes, it is important to have an estimate of the relevant elasticity for reforms affecting the full population. This matters when, e.g., parameterizing optimal income taxation models and evaluating general tax cuts or tax credits, such as the Swedish EITC. Most studies in our survey focus on specific reforms or subgroups, but nine estimate an elasticity for the general working-age population (typically ages 25–54) using techniques such as group IV or individual fixed effects regressions. The average elasticity in these papers is 0.1.[15] However, for general tax cuts, the responses of indi-

[15] *The nine papers are Gelber & Mitchell (2012), Lin & Tong (2017), McClelland, Mok & Pierce (2014), Bartels & Shupe (2018), Blundell, Bozio & Laroque (2011), Jäntti, Pirttilä & Selin (2015), Sigurdsson (2019), Stefansson (2019) and Martinez, Saez & Siegenthaler (2018).*

viduals at the margin of retirement are also important. As estimated elasticities for this subgroup are considerably higher than 0.1, we assess that the average full-population elasticity lies in the range 0.1–0.2.[16]

Overall, a clear pattern is that elasticities are greater for groups with a low employment rate, such as low-skilled single mothers. This is perhaps to be expected, as the elasticity by definition is decreasing in the employment rate for a given employment effect. In addition, a greater number of people out of work likely means a larger number of people at the margin of entering employment, which is what matters for the magnitude of the elasticity.[17]

5. Conclusion

Since the mid-1990s, a growing number of studies have used quasi-experimental methods to identify labour force participation responses to tax and benefit reforms. We have identified 35 such papers. We find that participation elasticities are larger for women and have declined over time. Americans seem to be more responsive than Europeans. The average elasticity across all studies is 0.36, although there is a large range from 0 to more than 1. Many papers focus on groups with a larger potential labour reserve, such as single mothers, where – as expected – estimated participation responses also are higher. We believe that the policy-relevant elasticity for the full population lies in the range 0.1–0.2.

We offer some advice for future research in this area. Researchers should use the elasticity definition that is now standard in the literature, i.e., the participation elasticity with respect to the financial gain from work. This allows for comparison between countries and reforms, and makes it easier to predict the effects of future reforms. Many papers do not report an elasticity at all. Although in some cases an elasticity can be calculated from information reported in the paper (which, e.g., Chetty et al., 2013, do), this also increases the risk of error.[18] Preferably, the elasticity should be calculated by researchers who have access to the underlying data.

It is also worth noting that a relatively small number of countries is covered by our survey. The majority of papers are from the United States. Two are from Canada and the rest are from Western Europe. Some examples of major high-income countries that are completely absent are Japan, Australia and Germany. This suggests that there is much room for continued research. □

References

Alpert, A. & Powell, D. (2014), "Estimating Intensive and Extensive Tax Responsiveness: Do Older Workers Respond to Income Taxes?", RAND Working Paper WR-987-1.

Angrist, J. & Pischke, J.-S. (2009), *Mostly Harmless Econometrics*, Princeton, N. J.: Princeton University Press.

Autor, D. H., Duggan, M., Greenberg, K. & Lyle, D.S. (2016), "The impact of disability benefits on labor supply: Evidence from the VA's disability compensation program", *American Economic Journal: Applied Economics*, 8 (3), pp.31–68.

Bargain, O. & Doorley, K. (2011), "Caught in the trap? Welfare's disincentive and the labor supply of single men", *Journal of Public Economics*, 95 (9–10), 1096–1110.

Bargain, O. & Peichl, A. (2016), "Own-wage labor supply elasticities: variation across time and estimation methods", *IZA Journal of Labor Economics*, 5 (1).

[16] The papers that focus on older workers are Alpert & Powell (2014), Brown (2013), Gelber et al. (2017) and Laun (2017). As shown in table 2, the average elasticity is 0.38.

[17] See the discussion in the theory section and in Bastani, Moberg & Selin (2020).

[18] Cf. footnote 12.

Bartels, C. & Shupe, C. (2018), "Drivers of participation elasticities across Europe: gender or earner role within the household?", IZA Discussion Paper No. 11,359.

Bastani, S., Moberg, Y. & Selin, H. (2020), "The anatomy of the extensive margin labor supply response", *Scandinavian Journal of Economics*, forthcoming.

Bastian, Jacob (2020), "The Rise of Working Mothers and the 1975 Earned Income Tax Credit", *American Economic Journal: Economic Policy*, forthcoming.

Bettendorf, L. J., Folmer, K. & Jongen, E. L. (2014), "The dog that did not bark: The EITC for single mothers in the Netherlands", *Journal of Public Economics*, 119, pp. 49–60.

Bianchi, M., Gudmundsson, B. R. & Zoega, G. (2001), "Iceland's Natural Experiment in Supply-Side Economics", *American Economic Review*, 91 (5), pp. 1564–79.

Blau, F. D. & Kahn, L. M. (2007), "Changes in the labor supply behavior of married women: 1980–2000", *Journal of Labor Economics*, 25 (3), pp. 393–438.

Blundell, R., Duncan, A. & Meghir, C. (1998), "Estimating labor supply responses using tax reforms", *Econometrica*, 66 (4), pp. 827–861.

Blundell, R., Brewer, M. & Shephard, A. (2005), "Evaluating the labour market impact of Working Families' Tax Credit using difference-in-differences", HM Revenue and Customs, Working Papers 4.

Blundell, R., Bozio, A. & Laroque, G. (2011), "Labor supply and the extensive margin", *American Economic Review*, 101 (3), pp. 482–86.

Brown, K. M. (2013), "The link between pensions and retirement timing: Lessons from California teachers", *Journal of Public Economics*, 98, pp. 1–14.

Carbonnier, C. (2008), "Spouse labor supply: Fiscal incentive and income effect, evidence from French fully joint income tax system", THEMA Working Paper No. 2008-20.

Card, D. & Hyslop, D. (2005), "Estimating the effects of a time-limited earnings subsidy for welfare-leavers", *Econometrica*, 73 (6), pp. 1723–1770.

Carrington, W. J. (1996), "The Alaskan Labor Market during the Pipeline Era", *Journal of Political Economy*, 104 (1), pp. 186–218.

Chetty, R., Friedman, J. N. & Saez, E. (2013), "Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings", *American Economic Review*, 103 (7), pp. 2683–2721.

Chetty, R., Guren, A., Manoli, D. & Weber, A. (2013), "Does indivisible labor explain the difference between micro and macro elasticities? A meta-analysis of extensive margin elasticities", In *NBER Macroeconomics Annual 2012, Volume 27*.

Cogan, J. F. (1981), "Fixed costs and labor supply", *Econometrica*, 49 (4).

Devereux, P. J. (2004), "Changes in Relative Wages and Family Labor Supply", *Journal of Human Resources*, 39 (3), pp. 698–722.

Edmark, K., Liang, C.-Y., Mörk, E. & Selin, H. (2016), "The Swedish Earned Income Tax Credit: Did It Increase Employment?", *FinanzArchiv: Public Finance Analysis*, 72 (4), pp. 475–503.

Eissa, N. (1995), "Taxation and labor supply of married women: the Tax Reform Act of 1986 as a natural experiment", NBER Working Paper No. 5,023.

Eissa, N. (1996), "Labor supply and the economic recovery Tax Act of 1981", In Feldstein, M. & Poterba, J.M. (eds.), *Empirical Foundations of Household Taxation*, Chicago: University of Chicago Press.

Eissa, N. & Hoynes, H. W. (2004), "Taxes and the labor market participation of married couples: the earned income tax credit", *Journal of Public Economics*, 88 (9–10), pp. 1931–1958.

Eissa, N. & Liebman, J. (1996), "Labor supply response to the earned income tax credit", *Quarterly Journal of Economics*, 111, pp. 605–637.

Fadlon, I. & Nielsen, T. H. (2015), "Family Labor Supply Responses to Severe Health Shocks", NBER Working Paper No. 21,352.

Francesconi, M. & van der Klaauw, W. (2007), "The socioeconomic consequences of 'in-work' benefit reform for British lone mothers", *Journal of Human Resources*, 42, pp. 1–31.

French, E. & Song, J. (2014), "The effect of disability insurance receipt on labor supply", *American Economic Journal: Economic Policy*, 6 (2), pp. 291–337.

Gelber, A. M., Jones, D., Sacks, D. W. & Song, J. (2017), "Using Non-Linear Budget Sets to Estimate Extensive Margin Responses: Method and Evidence from the Social Security Earnings Test", NBER Working Paper No. 23, 362.

Gelber, A. M. & Mitchell, J. W. (2012), "Taxes and time allocation: Evidence from single women and men", *Review of Economic Studies*, 79 (3), pp. 863–897.

Graversen, E. K. (1998), *Labor Supply and Work Incentives*, PhD dissertation, University of Aarhus.

Gregg, P. & Harkness, S. (2003), "Welfare reform and lone parents employment in the UK", Centre for Market and Public Organisation, Working Paper 03/072.

Gruber, J. (2000), "Disability insurance benefits and labor supply", *Journal of Political Economy, 108* (6), pp. 1162–1183.

Gruber, J. & Saez, E. (2002), "The elasticity of taxable income: Evidence and implications", *Journal of Public Economics*, 84 (1), pp. 1–32.

Gruber, J. & Wise, D. A. (1999), "Introduction and Summary", In *Social Security and Retirement Around the World*, Chicago: University of Chicago Press.

Heim, B. T. (2007), "The incredible shrinking elasticities married female labor supply, 1978–2002", *Journal of Human Resources, 42* (4), pp. 881–918.

Hotz, V. J. & Scholz, J. K. (2003), "The earned income tax credit", In Moffitt, Robert A. (ed.), *Means-Tested Transfer Programs in the United States*, Chicago: University of Chicago Press.

Hotz, V. J., Mullin, C.H. & Scholz, J.K. (2002), "The Earned Income Tax Credit and labor market participation of families on welfare", Mimeo, University of California–Los Angeles, Vanderbilt and University of Wisconsin–Madison.

Jonassen, A. B. (2013), *Regression Discontinuity Analysis of the Disincentive Effects of Increasing Social Assistance*, PhD thesis, Department of Economics and Business, Aarhus University.

Juhn, C., Murphy, K. M. & Topel, R. H. (1991), "Why Has the Natural Rate of Unemployment Increased over Time?", *Brookings Papers of Economic Activity* (2), pp. 75–142.

Jäntti, M., Pirttilä, J. & Selin, H. (2015), "Estimating labour supply elasticities based on cross-country micro data: A bridge between micro and macro estimates?", *Journal of Public Economics*, 127, pp. 87–99.

Keane, M. P. (2011), "Labor supply and taxes: A survey", *Journal of Economic Literature*, 49 (4).

Kleven, H. (2016), "Bunching", *Annual Review of Economics*, 8, pp. 435–464.

Kleven, H. (2019), "The EITC and the Extensive Margin: A Reappraisal", NBER Working Paper 26, 405.

Koning, P. & van Sonsbeek, J. M. (2017), "Making disability work? The effects of financial incentives on partially disabled workers", *Labour Economics*, 47, pp. 202–215.

Kosonen, T. (2014), "To work or not to work? The effect of childcare subsidies on the labour supply of parents", *The BE Journal of Economic Analysis & Policy, 14* (3), pp. 817–848.

Kostol, A. R. & Mogstad, M. (2014), "How financial incentives induce disability insurance recipients to return to work", *American Economic Review, 104* (2), pp. 624–55.

Kumar, A. (2016), "Lifecycle-consistent female labor supply with nonlinear taxes: evidence from unobserved effects panel data models with censoring, selection and endogeneity", *Review of Economics of the Household, 14* (1), pp. 207–229.

Kumar, A. & Liang, C. Y. (2016), "Declining female labor supply elasticities in the United States and implications for tax policy: Evidence from panel data", *National Tax Journal, 69* (3), pp. 481–516.

Laun, L. (2017), "The effect of age-targeted tax credits on labor force participation of older workers", *Journal of Public Economics, 152*, pp. 102–118.

Leigh, A. (2007), "Earned income tax credits and labor supply: evidence from a British natural experiment", *National Tax Journa*l, 60 (2), pp. 205–224.

Liebman, J. & Saez, E. (2006), "Earnings responses to increases in payroll taxes", NBER Retirement Research Center Paper No. NB 04–06.

Lin, E. Y. & Tong, P. K. (2017), "Married couple work participation and earnings elasticities: evidence from tax data", *International Tax and Public Finance*, 24 (6), pp. 997–1025.

Lundberg, J. (2017), "Analyzing tax reforms using the Swedish Labour Income Microsimulation Model", Working Paper 2017:12, Department of Economics, Uppsala University.

Manoli, D. & Weber, A. (2011), "Nonparametric Evidence on the Effects of Retirement Benefits on Labor Force Participation Decisions", NBER Working Paper No. 17, 320.

Manoli, D. & Weber, A. (2016), "Nonparametric Evidence on the Effects of Financial Incentives on Retirement Decisions", *American Economic Journal: Economic Policy*, 8 (4), pp. 160–182.

Martinez, I. Z., Saez, E. & Siegenthaler, M. (2018), "Intertemporal Labor Supply Substitution? Evidence from the Swiss Income Tax Holidays", NBER Working Paper No. 24, 634.

McClelland, R. & Mok, S. (2012), "A review of recent research on labor supply elasticities", Congressional Budget Office Working Paper 2012-12.

McClelland, R., Mok, S. & Pierce, K. (2014), "Labor force participation elasticities of women and secondary earners within married couples", Congressional Budget Office Working Paper 2014-06.

Meghir, C. & Phillips, D. (2010), "Labour supply and taxes", In Adam, Stuart et al. (eds.), *Dimensions of Tax Design: The Mirrlees Review*, Oxford: Oxford University Press.

Meyer, B. & Rosenbaum, D. (2001), "Welfare, the earned income tax credit, and the labor supply of single mothers", *Quarterly Journal of Economics*, 116, pp. 1063–1114.

Milligan, K. & Stabile, M. (2007), "The integration of child tax credits and welfare: Evidence from the Canadian National Child Benefit program", *Journal of Public Economics*, 91 (1–2), pp. 305–326.

Neisser, C. (2017), "The elasticity of taxable income: A meta-regression analysis", IEB Working Paper 2017/10. IZA DP No. 11958

Saez, E. (2002), "Optimal income transfer programs: Intensive versus extensive labor supply responses", *Quarterly Journal of Economics*, 117 (3).

Saez, E., Slemrod, J. & Giertz, S. H. (2012), "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review", *Journal of Economic Literature*, 50 (1).

Selin, H. (2014), "The rise in female employment and the role of tax incentives. An empirical analysis of the Swedish individual tax reform of 1971", *International Tax and Public Finance*, 21 (5), pp. 894–922.

Sigurdsson, J. (2019), "Labor Supply Responses and Adjustment Frictions: A Tax-Free Year in Iceland", Mimeo, Bocconi University.

Stefánsson, A. (2019), "Labor supply response to a tax holiday: The take-home from a large and salient shock", Essay 1 in *Essays in Public Finance and Behavioral Economics*. PhD thesis, Uppsala University.

Appendix: Comparison with Chetty et al. (2013)

The only previous general survey of the quasi-experimental participation elasticity literature that we know of is Chetty et al. (2013), which cites 15 papers. We include seven of those in our survey. Below are the papers left out, and our motivations for doing so.

Juhn, Murphy & Topel (*1991*): Regional wage trends (presumably before-tax) are used for identification – not policy variation.

Graversen (*1998*): The parametric and nonparametric DD estimates have different signs, indicating non-parallel trends that are difficult to control for.

Devereux (*2004*): Before-tax wages are used.

Liebman & Saez (*2006*): The authors report many different estimates, and state that it is unlikely that a suitable control group can be found.

Carrington (*1996*): The paper studies a labour demand shock. We are interested in the effects of policy.

Gruber & Wise (*1999*): Cross-country evidence only – not quasi-experimental.

Bianchi, Gudmunndsson & Zoega (*2001*): The elasticity reported concerns the number of weeks worked, which departs from the conventional definition of the extensive margin.

Manoli & Weber (*2011*): In the published version (Manoli & Weber, 2016), the authors report a semielasticity and state that it is difficult to translate into an elasticity (p. 172).

Switching Costs in the Finnish Retail Deposit Market*

Tuomas Takalo
Bank of Finland and VATT

Abstract

I calibrate the switching cost for the Finnish retail deposit market by using the approach developed by Oz Shy (2002). It turns out that switching costs faced by deposit customers of the main Finnish banks manifest large variation and are high, ranging from 200 euros to nearly 1,400 euros. Over a 20-year period, switching costs have increased by roughly 50% in real terms, but in relation to average account balance, switching costs have not essentially changed. Changes and differences in the banks' competitive strategies might explain the variation in switching costs across time and banks..

JEL Classification: G21,L13,L49.

**Contact e-mail: tuomas.takalo@bof.fi. I thank two anonymous referees, Rune Stenbacka and Sara Stenvik for detailed comments and useful discussions. A part of this paper contains data and calibration methods behind the results reported in Section 2 of Rune Stenbacka and Tuomas Takalo (2019) "Switching Costs and Financial Stability", Journal of Financial Stability,41,14–24. I also thank Juho Anttila for research assistance.*

Editor: Kari Heimonen

1. Introduction

Switching costs shape bank competition by conveying market power and affecting pricing (see, e.g., Gehrig and Stenbacka, 2007, Degryse and Ongena, 2008, ICB, 2011, and Ciet and Verdier, 2019). Bank switching costs are also recognized as a determinant of financial stability (Stenbacka and Takalo, 2019, and Brown et al., 2020). The influential Vickers report (ICB, 2011) elevates switching costs to a central role in banking regulation, and reviews many policy tools to affect switching costs. Direct evidence of the magnitude of bank switching costs is, however, scant. In this paper I measure the switching cost in the Finnish deposit market by using the approach developed by Oz Shy (2002). As Shy (2002) also applied the method to the Finnish deposit market, comparing the results of these two studies reveals how bank switching costs have changed over a 20-year period in Finland.

I find that switching costs faced by customers of the main Finnish banks manifest large variation and are high, ranging from 200 euros to nearly 1,400 euros. In relation to the average account balance of a customer, switching costs range from 2% to 15%. Comparing these numbers with those reported by Shy (2002) suggests that while switching costs have increased some 50% in real terms over 20 years, switching costs per average account balance have not changed.

The cross-sectional variation in switching costs might partially reflect differences in the banks' loyalty programs: For example, the Savings Banks Group does not run loyalty programs invariantly and imposes the lowest switching costs to its clients. In contrast, the clients of the OP Financial Group face the highest switching costs. The OP Group is a cooperative entity running a sophisticated loyalty program where loyalty bonuses accumulate from an owner-customer's use of the OP Group's banking and insurance services and can only be used for paying the OP Groups' banking service charges and insurance fees. Such loyalty bonuses cannot be transferred to another bank and they thus work much like frequent flier miles for airlines.

My results indicate that the OP Group's loyalty program might have been successful in locking in their owner-customers. The OP Group's loyalty program has also raised competition policy concerns: The Finnish Competition and Consumer Authority (FCCA) launched an investigation into the OP Group's loyalty program in December 2015 after a rival insurance provider, If P&C, filed a complaint, accusing the OP Group for abusing its dominant position by bundling the Group's banking and insurance products via its loyalty program. In its decision the FCCA, while considering the OP groups' loyalty program problematic from the competition policy point of view, finds no clear evidence that the loyalty program would significantly restrict competition in the non-life insurance market (FCCA, 2019). More generally, increasingly wide-spread loyalty programs and other changes in bank competition might explain relatively high switching costs found by this study.

While the existence of switching costs in deposit markets is well documented (see, e.g., Kiser, 2002, Carbo-Valverde et al., 2011, Hannan and Adams, 2011, and Brunetti et al., 2016), the evidence is, however, often indirect. A notable exception is Shy (2002) who develops a method of estimating switching costs in the banking industry directly and applies the method to the Finnish deposit markets. Shy's (2002) method only requires information about bank service charges and market shares. In contrast to Shy (2002), I can use banks' real names, and market shares are based on accurate numbers. However, determination of bank service fees is much more complicated today than it was in Shy (2002) due to more sophisticated product versioning and loyalty programs of banks. Prior to this study, Shy's method has been used to measure switching costs in the banking industry at least by Egarius and Weill (2016) but they do not analyze deposit market switching costs separately.[1]

Another method to estimate bank switching costs is developed by Kim et al. (2003). Their method uses bank accounting data, and is applied to deposit markets at least by Silva and Lucinda (2017). Silva and Lucinda (2017) report even higher estimates of switching costs relative to deposit account balance than in this study.

[1] *Shy's method has also been employed in thesis work – see, e.g., Carlström (2010) and Stenvik (2016).*

Customer loyalty programs have been extensively studied (see, e.g., Basso et al., 2009, and Kari et al., 2017, for a discussion of the issues). According to this literature, loyalty programs could be seen as a way for firms to increase switching costs, to lock in customers, and even to deter entry, since a customer will lose their loyalty benefits if they switch to a rival. Some loyalty programs could also be seen as a form of product versioning where a firm with a market power attempts to price discriminate its customers. Loyalty programs also provide firms with valuable information about their customers, allowing for more accurate customer tracking and database marketing. The competitive implications of customer loyalty programs are not clear; as in the case of switching costs more generally, they can make markets more or less competitive depending on the circumstances (see, e.g., Basso et al., 2009; Ruiz-Aliseda, 2016). In the banking context, the competitive implications of switching costs have been shown to be particularly complex since they may also depend on the banking regulation such as deposit insurance, and information-disclosure and bail-out policies (see, e.g., Gehrig and Stenbacka 2007, Ciet and Verdier, 2019, and Stenbacka and Takalo, 2019).

I next replicate the main parts of Shy's (2002) model. Then, in Section 3, I explain the institutional environment of the Finnish banking industry and collection of the data. I combine the data with the model in Section 4 so as to provide new evidence of the deposit market switching costs. Section 5 concludes and discussed policy implications.

2. The Model

I replicate here the key features of the model in Shy (2002), referring the reader to the original source for more details and proofs (see also Shy, 2001).

Consider a market with k banks, $\{k \in \mathbb{N} | k \geq 2\}$, indexed by $i = 1, ...k$. Each bank i has initially $N_i \in \mathbb{N}$ customers who face the choice of either remaining in the bank or switching to another one. A customer's utility is given by

$$(1) \qquad U_i = \begin{cases} -f_i, & \text{if the consumer stays,} \\ -f_j - \delta_i, \forall j \neq i, & \text{if the consumer switches,} \end{cases}$$

in which $f_i \in \mathbb{R}_+$ is the service fee charged by bank i, and $\delta_i \in \mathbb{R}_+$ is the switching cost in the case where bank i's customer decides to change their banking relation to the rival bank j.

The profits of bank i are then given by

$$(2) \qquad \pi_i(f_1, ..., f_k) = f_i q_i$$

in which $q_i \in \mathbb{N}$ is the number of customers who will choose to deposit in bank i.

The banks are indexed according to a decreasing market share order so that bank 1 has the largest market share and bank k has the smallest market share. It is further assumed that i) each bank i, $i \neq k$, fears being undercut by bank k, and sets its fee f_i in reference to f_k, and that ii) the smallest bank k fears being undercut by the largest bank 1, and therefore sets its fee f_k in reference to f_1.

Under these assumptions there exists a vector of fees $(f_1, ..., f_k)$ that satisfies the Undercut-proof Property (UPP). In price competition, firms have an incentive to undercut a rival's price in order to attract customers from their competitor. Intuitively the UPP is satisfied when no bank can increase its profits by undercutting a rival bank and no bank can increase its service fee without being undercut by a rival.

Formally, when the UPP is satisfied, each bank i, $i \neq k$, chooses its fee f_i to maximize $\pi_i(f_i, f_k)$ (as given by equation (2)) subject to the constraint

(3) $$f_k q_k \geq (f_i - \delta_i)(N_i + N_k), \quad i \neq k,$$

taking f_k as given. Bank k in turn chooses f_k to maximize $\pi_k(f_k, f_1)$ subject to

(4) $$f_1 q_1 \geq (f_k - \delta_k)(N_i + N_k), \quad i \neq k,$$

taking f_i as given.

Equations (2)–(4) imply that the banks choose the highest possible prices satisfying constraints (3) and (4). Therefore constraints (3) and (4) hold as equalities. Furthermore, in an UPP equilibrium it must hold that $q_i = N_i \; \forall i$. Substituting N_i for q_i in equations (3) and (4), and solving for δ_i yields the UPP switching costs as

(5) $$\begin{aligned} \delta_i &= f_i - \frac{N_k f_k}{N_i + N_k}, \quad i \neq k, \\ \delta_k &= f_k - \frac{N_1 f_1}{N_1 + N_k}. \end{aligned}$$

Equation (5) implies that estimating switching costs only requires information about banks' service fees and the relative number of retail customers in each bank.

3. Institutional Environment and Data

3.1 Finnish Retail Banking Industry

Since the Finnish banking crisis of the early 1990s, there has been a large number of mergers in the Finnish banking industry. As a result the Finnish retail banking market is concentrated. As shown by Table 1 the deposit market shares of the two and four largest banks are over 65% and 80%, respectively. In what follows, I will focus on the four largest banking groups, the OP Group, Nordea, Danske Bank and the Savings Banks Group.

The Finnish retail banking market is also characterized by the use of customer loyalty programs, which reward customers for concentrating all their banking services and assets on the same bank. Typically, a customer gets bonuses, discounts, or other benefits once they have a threshold amount of assets (e.g., deposits and loans) at their bank.

Table 1: Bank Deposit Market Shares in Finland in 2016

Bank	Deposits (M€)	Market share (%)
OP Group	55,198	37.5
Nordea	40,723	27.7
Danske Bank	18,411	12.5
Savings Bank Group	6,072	4.1
Others	26,694	18.1
Total	147,098	100

Notes: This table lists deposit account balances (excluding deposits from financial institutions) at the largest banks in Finland, and the corresponding deposit market shares at the end of year 2016. Deposit and market share figures are in million euros and percentages, respectively. Source: Finance Finland (2017).

Of the four main banks in Finland, three run a customer loyalty program. The market leader, the OP Group, is a cooperative, offering loyalty discounts to those customers who are also its owners. The amount of discounts awarded to a customer depends on the customer's average monthly assets and loans at the OP Group. The loyalty benefits at Danske Bank and Nordea, the two main commercial banks in Finland, depend on the amount of assets in the bank; the key details of their programs are listed in Table 2. Out of the four main banking groups, only the Savings Banks Group does not run customer loyalty programs invariantly.

Table 2: Loyalty Programs of Nordea and Danske Bank

Nordea			
Regular customer	Key customer		
Assets ≥6,000 €	Assets ≥30,000 €		
Products from ≥3 different categories	Products from ≥5 different categories		
Regular monthly income ≥500 €	Regular monthly income ≥500 €		
Danske Bank			
Level 1	Level 2	Level 3	Level 4
Assets 0–10,000 €	Assets 10,000–50,000 €	Assets 50,000–150,000 €	Assets ≥150,000 €

Notes: This table lists the requirements for each level of the customer loyalty programs of the two main commercial banks in Finland, Nordea and Danske Bank, in 2017. "Assets" includes both savings and loans.

3.2 Measuring Market Shares and Service Fees

While equation (5) suggests that estimating switching costs only requires information about service fees charged by each bank and the relative number of retail customers in each bank, I do not have these figures but need to approximate them from available data.

I approximate a bank's market share in terms of retail customers by a bank's market share in terms of retail deposits, as given in Table 1. Shy (2002) suggests of using the number of bank accounts as a proxy for the bank's customer base. The problem with this proxy is that many accounts are inactive. For example, according to the Bank of Finland sources, there were in total 16,211,877 bank accounts in Finland in 2016, which corresponds roughly 3.5 bank account per adult person.[2] The deposit market share proxy circumvents this problem but I cannot take into account the skewed distribution of deposits across customers in calculations.

Service fees are typically monthly or annual fees. Hence, when a customer contemplates switching a bank, relevant consideration is the discounted sum of fees that the customer expects to pay if they stay with their current bank or switch to another bank. I therefore calculate lifetime fees by discounting the infinite sum of monthly and annual fees with the same four percentage real interest rate that Shy (2002) also used. More specifically, the lifetime fee $f_{l,i}$ for bank i is calculated from the bank's monthly fee $f_{m,i}$ with the formula $f_{l,i} = 12 \cdot f_{m,i}/(1-d)$ whered $d = 1/(1+r)$ is the discount factor when the real interest rate is $r \in \mathbb{R}_+$. With $r = 0.04$, $f_{l,i} = 312 \cdot f_{m,i}$.

I collect information about banks' service fees from the VertaaEnsin.fi on-line platform in January 2018. VertaaEnsin.fi is a part of the CompareEurope-Group, a leading provider of online comparison platforms for financial services in Europe. VertaaEnsin.fi contains up-to date information about various retail banking packages, customer loyalty pro-

[2] *There were 5.503 million people in Finland in 2016, of which 84% were at least 15 years, see Statistics Finland, http://www.stat.fi/tup/suoluk/suoluk_vaesto.html, last accessed on 30 October, 2017.*

grams, and the associated account and payment card fees in Finland. To facilitate a customer's comparison of banks and their service fees, the platform also selects the most relevant service packages for each bank. I include all these packages in the service fee calculations, and double-check the accuracy of the information for these packages from the banks' own websites.

VertaaEnsin.fi, however, contains no information about the Savings Banks Group. It provides service fee information for Oma Säästöpankki, the largest savings bank in Finland, but Oma left the Savings Banks Group in 2015. I therefore use Nooa Säästöpankki as the representative of the Savings Banks Group. Nooa is owned by the other group member bank and is a large savings bank operating in the Helsinki metropolitan area. I obtain Nooa's fee information from the bank's website. Using the fee information for Oma from VertaaEnsin.fi as a representative of savings banks fees instead of Nooa's fees gives essentially the same results (see Section 4.3).

Using the collected service fee information, I calculate the average monthly and lifetime fees for the banks. Table 3 displays the results. The first service package featured in Table 3 for each bank is a mandatory banking service package: A customer residing in Finland has a statutory right to basic banking services that include a current account, a payment card, and internet banking services (Amendment to Act on Credit Institutions §1054/2016). The other packages I consider typically include a more advanced payment card and some other services. The packages in the table are labeled according to the most advanced payment card included in a package. (In some premium packages, a customer can have access to another payment card and bank account for the same fee.)

Table 3 shows that the lifetime fees for the mandatory service package and for a package with a standard combined debit-credit card are roughly 1,000–2,000 euros. Customers having access to the highest loyalty benefit package in Nordea face the lowest fees. To reach such loyalty benefit levels, a customer needs to hold some non-negligible amount of assets in the bank (see Table 2). Therefore it is likely that such a customer pays other fees to Nordea, such as mortgage interest rates and repayment fees, or fund management fees, which are not captured by the service fee calculations here. Customers willing to purchase a premium service package at the lowest loyalty benefit level in Danske Bank face the highest fees, but such customers are probably rare.

Table 3 also reveals that the banks' average fees across all customer categories of a bank, except in the case of Danske Bank, are close to each other, approximately five euros per month or roughly 1,500–1,600 euros over the lifetime. However, Danske Bank's larger average fee is driven by the high price of the premium (Platinum) service package for the lowest loyalty benefit levels. If the Platinum package is excluded from two or three lowest benefit levels, Danske Bank's average fee becomes similar to the rivals' average service fee.

Table 3: Service Fees of the Largest Finnish Banks

Bank and Package	Monthly (€)	Lifetime (€)
Savings Banks Group		
Debit/Credit	4	1,248
Gold Debit/Credit	6.25	1,950
All customers, average	5.13	1,599
OP Group		
Non-owner customers		
Electron	5.45	1,700
Owner-customers		
Debit/Credit	2.95	920
Gold Debit/Credit	6.50	2,028
All customers, average	4.95	1,550
Nordea		
Basic customers		
Electron	7.5	2,340
Regular customers		
Debit/Credit	5.25	1,638
Gold Debit/Credit	6.7	2,090
Key customers		
Gold Debit/Credit	0	0
All customers, average	4.86	1,517
Danske Bank		
Benefit level 1		
Debit	6.8	2,122
Gold Debit/Credit	6.9	2,153
Platinum Debit/Credit	18	5,616
Benefit level 2		
Debit	4.8	1,498
Gold Debit/Credit	5.9	1,841
Platinum Debit/Credit	12	3,744
Benefit level 3		
Debit	1.6	499
Gold Debit/Credit	3.8	1,187
Platinum Debit/Credit	9	2,808
Benefit level 4		
Debit	1.5	468
Gold Debit/Credit	3.7	1,154
Platinum Debit/Credit	8	2,496
All customers		
Average	6.83	2,132
Average excl. Platinum for bl. 1–3	4.83	1,505

Notes: The first column explains service packages at each bank and the second column their corresponding monthly service fees. The lifetime fees in the third column are calculated by using four percentage real interest rate, as in Shy (2002). All service packages include at least the statutory banking services (a bank account, internet banking, and a payment card). The service packages in the first column are labeled according to the most advanced payment card included in the package. "Electron" means that a package only includes the Visa Electron debit card, "Debit/Credit" means that a package includes a standard combination card that has both debit and credit payment features, and "Gold" and "Platinum" mean that a package includes a premium combination debit-credit card (Visa Gold, Mastercard Gold, or Mastercard Platinum). Visa is the main provider of cards for the Savings Bank and OP Groups, and Mastercard for Nordea and Danske Bank. The service packages and fees are collected in January 2018 from the VertaaEnsin.fi online comparison platform and banks' websites. The Savings Bank Group is represented by Nooa Säästöpankki. "Average" is an average service fee across all customer categories of a bank, and "Average excl. Platinum for bl. 1–3" is an average service fee of Danske Bank when the Platinum package is excluded from the benefit levels 1–3 but included in the benefit level 4.

4. Results

4.1 Calibration Procedure

Using the model of Section 2, and the deposit market shares and service fees calculated in Section 3, I can attempt to measure the switching costs. A challenge in this exercise is that I do not know the distribution of customers across various levels of the banks' customer loyalty programs. Thus, while Table 3 suggests that three main banks with the largest market shares engage in product differentiation, there is no point to extend the single fee model of Section 2 to capture this phenomenon. I thus proceed as if the all banks would set a single fee as in the model of Section 2.

In equation (5), I first let $k = 4$, and then use the third column of Table 1 to set $N_1 := N_{OP} = 0.375$, $N_2 := N_N = 0.277, N_3 := N_{DB} = 0.125$ and $N_4 := N_{SB} = 0.041$ in which subscripts OP, N, DB, and SB refer to the OP Group, Nordea, Danske Bank and the Savings Banks Group, respectively. Of the four banks considered the OP Group has the largest market share and the Savings Banks Group the lowest. Therefore the model is based on the assumption that the Savings Banks Group sets its fee by using the fee of the OP Group as the reference point, and the other three banks set their fees in reference to the fee of the Savings Bank Group.

As an example of switching cost calculation, let us consider the OP Group. I approximate the OP Group's service fee by its average fee across its customer categories. The assumption is heroic. It is plausible to think that a majority of the OP Group's customers are also its owners and use a standard combined Visa Debit/Credit card. Thus, using the average service fee approximates the service fees upwards. Yet, the service fee calculations only take into account the basic internet banking account fees and fixed annual fees from a payment card. Since most customers use some other banking services (e.g., withdraw cash, use credit features of a payment card, exchange currency, and so on), the service fees in my calculations are approximated downwards. Furthermore, since market shares are based on account balances and since it is plausible to think that customers using a Visa Gold card have larger account balances, the OP Group's service fee and, by implication, the switching costs of its customers, relative to account balances are more accurately captured than the average service fee and the switching cost in terms of euro amounts. In any event, by using Table 3, it is easy to calculate alternative switching costs by using alternative weightings of customer segments.

Under these assumptions, Table 3 reveals that the average lifetime discounted sum of service fees charged by the OP Group (f_{OP}) is approximately 1,550 euros. Similarly, the average life-time fee charged by the Savings Bank Group (f_{SB}) is approximately 1,599 euros. Then, equation (5) suggests that the switching costs facing the OP Group's customers are given by

$$\delta_{OP} = f_{OP} - \frac{N_{SB} f_{SB}}{N_{OP} + N_{SB}} = 1,550 - \frac{0.041 \cdot 1,599}{0.375 + 0.041} \approx 1,392. \tag{6}$$

Proceeding in the way outlined by equation (6) gives the switching costs for three remaining banks. In the case of Danske Bank, I use the average service fee that excludes the Platinum package from the benefit levels 1–3 but include it in the benefit level 4. To measure the switching costs per average account balance, I calculate the average account balance by dividing the total account balance in the Finnish banking industry from Table 1 by the total number of bank accounts in Finland in 2016.

4.2 Results

The main results are summarized in Table 4. The two bottom rows display the calibrated switching costs. The mean lifetime switching cost is 1,004 euros, and 11% in relation to the average account balance. The Savings Bank Group's customers can switch a bank more cheaply than the customers of the other banks. Shy (2002) also finds that the customers of the smallest bank face much lower switching costs than the customers of its rivals.

Table 4: Switching Costs in the Finnish Banking Industry in 2017

	OP Group	Nordea	Danske Bank	Savings Banks
Market share (%)	37.5	27.7	12.5	4.1
Average monthly fees (€)	4.95	4.86	4.83	5.13
Lifetime fees (€)	1,550	1,517	1,505	1,599
Switching costs (€)	1,392	1,311	1110	202
SC/avg. bal. (%)	15	14	12	2

Notes: The last row expresses switching costs per average account balance (9074 euros). The average account balance is calculated by dividing aggregate balance (147,098 M€), obtained from from Table 1, by the total number of bank accounts (16,211,877) in 2016, obtained from the Bank of Finland. Market shares are from Table 1, and monthly and lifetime fees are from Table 3. The fees reflect the situation at the beginning of year 2018 and other variables at the end of year 2016.

It is useful to compare the results of Table 4 to the results in Table 2 in Shy (2002). Since Shy (2002) uses the Finnish deposit market data from 1997 and here the data comes approximately around the year 2017, the comparison reveals that switching costs have increased by roughly 50% in real terms during the 20-year period from 1997 to 2017. However, the comparison shows no essential changes in switching costs per average account balance over the 20-year period in the Finnish banking industry. This pattern of higher switching costs and deposit account balance could reflect an increase in wealth and opportunity cost of time of the Finnish depositors since 1997.

Silva and Lucinda (2017) use a different method – the one developed by Kim et al. (2003) – to estimate switching costs in the Brazilian deposit markets. In the most comparable set up to mine, Silva and Lucinda (2017) report that switching costs faced by the depositors of the largest Brazilian banks range from 26% to 30% relative to average deposit account balance, which is even a higher figure than here.

An explanation for the Savings Bank Group's lower switching costs might be that savings banks are stakeholder banks where managers might have lower incentives to lock in their clients. This explanation is put forward by Egarius and Weill (2016) who find that across all banking activities and in lending markets (they do not consider deposit markets separately), the customers of cooperative banks tend to have lower switching costs than the customers of other bank types. In my data, however, the customers of the cooperative bank (the OP Group) face the highest switching costs. Thus the differences in profit-maximization objectives provide no obvious explanations for the findings here.

An alternative explanation could arise from the fact that as a cooperative, the OP Group attracts members based on common bonds. Such bank customers face higher switching costs. The importance of common bonds as a rationale for the cooperative bank membership has, however, diminished over time in Finland (Jones et al., 2016). Rather, I interpret the findings as to suggest that the OP Group's loyalty program has been successful to lock in their owner-customers, and the absence of the loyalty program in the Savings Banks Group might be a major reason for its lower switching costs.

Increasingly wide-spread adoption of loyalty programs might also contribute to the documented increase in the switching costs since Shy's (2002) study. For example, the OP Group introduced its loyalty program in 1999 after Shy collected his data. Moreover, after 2011 the only way to use the OP Group's loyalty bonuses has been to pay for the OP Group's service fees (FCCA, 2019).

These loyalty bonuses cannot be transferred to another bank and they thus work much like frequent flier miles for airlines, generating switching costs. The extent to which the loyalty programs of Danske Bank and Nordea outlined in Table 2 create switching costs is less clear, although the FCCA (Saarinen, 2014) appears to regard the loyalty programs of all major Finnish banks as switching barriers. The loyalty programs of Danske Bank and Nordea at least obscure the comparison of banks' service charges which – according to the Vickers report (2011) – is linked to high switching costs.

High switching costs documented in this study might indicate weak competition in the Finnish banking industry. As also documented in this study, the market shares of the largest banks are high. According to some measures, the Finnish banking industry is the most concentrated in Europe (see, e.g., Saarinen, 2014 and Savolainen, 2016). The Finnish banking industry has repeatedly attracted attention from the competition policy authorities during this millennium. Even prior to the case related to the OP Group's loyalty program, the lack of competition and high switching costs due to loyalty programs have been a concern to the FCCA (see, e.g., Saarinen, 2014). The FCCA has also raised the concern that the Finnish banks use Finance Finland – the industry association of the Finnish financial sector firms – as a collusive device so as to raise service fees (see, FCCA, 2016).[3] The banks' public announcements about the future mortgage margins (see, e.g., Rintakoski, 2015) and their cooperation in the automatic teller machine market have also concerned the FCCA (see, e.g., Kopsakangas-Savolainen and Takalo, 2014).

However, the relationship between switching costs and competition in the banking industry is not straightforward (Gehrig and Stenbacka, 2007, Carbo-Valverde et al., 2011, Ciet and Verdier, 2019, and Stenbacka and Takalo, 2019). As suggested by Carbo-Valverde et al. (2011) and Stenbacka and Takalo (2019), an increase in switching costs in the deposit markets should weaken competition with inherited customers but intensify competition for new customers. Since the Finnish banks often compete for new customers via mortgage interest rates, this competition should become more intense with the increasing level of switching costs. In line with this prediction, a rough calculation suggests that the average mortgage margin in Finland decreased by 46% between December 1997 and December 2017.[4]

4.3 Robustness

My measurement exercise involves a number of strong assumptions. I have therefore conducted several robustness checks by using alternative shortcuts. For example, using a bank's mortgage market share instead of its deposit market share as a proxy for the bank's customer base would yield similar results but the fourth largest bank in terms of granted mortgages would be Aktia, just ahead of the Savings Banks Group.

I report here in more detail results from the robustness check where I use the fee information for Oma Säästöpankki from the VertaaEnsin.fi online comparison platform as a representative of savings banks fees instead of the hand-collected information for the fees of Nooa Säästöpankki (see Section 3.2).

---

[3] *The OP Group's loyalty program case has also a link to the banks' cooperation within Finance Finland: As mentioned in the introduction, the case began after If P&C filed a complaint about the OP Group's loyalty program to the FCCA. The complaint prompted the OP Group to withdraw from Finance Finland but just prior to the FCCA's decision, it returned back to the association. According to the OP Financial Group's Chief Executive Timo Ritakallio, the main reason for the return was "the desire to increase the cooperation within the industry [translation from Finnish by the author]", https://twitter.com/ritakti/status/1009045585574400000, last accessed on June 10, 2020.*

[4] *I calculated this reduction from the Bank of Finland statistics, https://www.suomenpankki.fi/fi/Tilastot/rahalaitosten-tase-lainat-ja-talletukset-ja-korot/ (last accessed on June 25, 2020). The Bank of Finland readily calculates the average margin on new mortgages in Finland and it was 0.94 percentage points at the end of 2017. However, the Bank of Finland's average mortgage margin time series only begin from 2010. Following the Bank of Finland's method to calculate the average margin on new mortgages, I find that the average mortgage margin was between 1.69 and 1.76 percentage points in Finland at the end of 1997. Assuming that the average mortgage margin was 1.73 percentage points, the margin thus decreased by 0.79 percentage points or by 46% between December 1997 and December 2017.*

VertaaEnsin.fi gives only one the monthly fee for Oma Säästöpankki: five euros per month for a standard banking service package with a combined debit-credit card. Using the four percentage real interest rate, the corresponding life time fee is 1,560 euros. Using this euro amount instead of 1,599 for the savings banks life time fee implies that the switching costs for the customers of the OP, Nordea, Danske Bank and Savings Banks are in euros 1,397, 1,390, 1,120, and 160, respectively. In words, the switching costs of the Savings Banks Group's clients are slightly lower and those of the other banks' clients correspondingly slightly higher. Compared to the average account balance, there are no essential changes.

My estimates of switching costs and especially their increase from Shy (2002) may seem high. The estimated costs were lower (higher) if I assumed that customers would be disproportionately distributed on the lowest (highest) fee categories for each bank. However, since I do not know the distribution of customers across different service packages, the "average fee" assumption used in this study is a natural starting point. Furthermore, switching costs per account balance reported in this study are in line with the ones in Shy (2002) and lower than in Silva and Lucinda (2017), supporting the meaningfulness of the estimated costs. Also, it is comforting that the increase in switching costs appears to be matched by an equal reduction in the average mortgage margin.

5. Conclusion

I measure switching cost for the Finnish retail deposit market by using the approach developed by Shy (2002). In Section 5 of his article, Shy (2002) also uses the Finnish deposit market as an example of switching cost measurement. As the data in Shy (2002) comes from 1997 and here around the year 2017, the results also show how bank switching costs have changed over 20 years in Finland. In contrast to Shy (2002), I can use banks' real names, and market shares are based on accurate numbers.

I find that switching costs faced by customers of the largest banks exhibit large variation, ranging from 200 euros to nearly 1,400 euros. While the costs are calculated from the discounted lifetime banking fees assuming an once-in-a-lifetime switch, these estimated switching costs can be seen as high, especially at the top end. In relation to the average account balance of a customer, switching costs range from 2% to 15%. Comparing these numbers with those reported by Shy (2002) suggests that while switching costs have increased some 50%in real terms over 20 years, switching costs per average account balance have not essentially changed. The OP Group's customers appear to face the highest switching costs whereas the Savings Banks Group's customers the lowest – this finding together with the results reported in Egarus and Weil (2016) suggests that switching costs of stakeholder banks appear to be different than those of shareholder banks. Clearly, more research on the role of organizational structure for competition, performance and stability in banking would be warranted. Ferri et al. (2014a,b) provide advances towards this direction.

I conjecture that the differences in switching costs among the Finnish banks might be explained by differences in their loyalty programs. The spread of these loyalty programs could also explain the increase in the real switching costs over 20 years documented in this study. High switching costs could also indicate weak competition in the Finnish retail deposit market, although the relationship between switching costs and competition in the banking industry is complex (Gehrig and Stenbacka, 2007, Carbo-Valverde et al., 2011, Ciet and Verdier, 2011, and Stenbacka and Takalo, 2019). The theoretical results in Stenbacka and Takalo (2019) and the evidence documented in this study suggest a hypothesis according to which an increase in switching costs in the Finnish deposit markets is linked with more intense competition for mortgagors.

A future work should extend Shy's (2002) method to product versioning to accommodate different banking service packages and loyalty programs. The method should also be extended to allow for multiple switches by a customer over her lifetime.

The implications of customer loyalty programs and switching costs for competition and stability in the banking industry should also be evaluated carefully. On one hand, the Basel III liquidity regulations and the findings in Brown et al. (2020)

indicate that tighter depositor relationships consistent with higher switching costs are likely to make depositors less likely to run on a bank in a crisis. On the other hand, high switching costs may result in a fierce competition for new customers and increase bank failure rates (Stenbacka and Takalo, 2019). To make a comprehensive evaluation of the stability implications of switching costs, it would be valuable to build a bank run model with multiple banks in which depositors can switch deposits from one bank to another (see, e.g., Chen and Hasan, 2006, for a contribution to this direction).

High switching costs in the banking industry documented by this and earlier studies support the call by the Vickers report (ICB, 2011) for more regulatory attention to bank switching costs. Optimal regulation of these costs in practice is, however, challenging since bank switching costs can be affected by many regulatory policies and authorities such as those concerning competition, consumer protection, and financial stability. □

References

Basso, L. J., T.M. Clements, and T. J. Ross (2009). "Moral Hazard and Customer Loyalty Programs." *American Economic Journal: Microeconomics* 1, 101–23.

Brown, M., B. Guin, and S. Morkötter (2020). "Deposit Withdrawals for Distressed Commercial Banks: Client Relationships Matter." *Journal of Financial Stability* 46.

Brunetti, M., R. Ciciretti, and Lj. Djordjevic, (2016). "The Determinants of Household's Bank Switching." *Journal of Financial Stability* 26, 175–189.

Carbo-Valverde, S., T. H. Hannan, and F. Rodriguez-Fernandez (2011). "Exploiting Old Consumers and Attracting New Ones: The Case of Bank Deposit Pricing." *European Economic Review* 5, 903–915.

Carlström, M. (2010). Switching Costs in Local Finnish Retail Bank Lending. Hanken School of Economics, Department of Economics, Economics Master's Thesis, January 20, 2010.

Chen, Y., and I. Hasan (2006). "Transparency of the Banking System and the Efficiency of Information-based Bank Runs." *Journal of Financial Intermediation* 15, 307–331.

Ciet, N., and M. Verdier (2019). Competition and Welfare Effects of Bailout Policies. Cred Working Paper No. 2019–21, Paris.

Egarius, D., and L. Weill (2016). "Switching Costs and Market Power in the Banking Industry: The Case of Cooperative Banks." *Journal of International Financial Markets, Institutions and Money* 42, 155–165.

Ferri, G., P. Kalmi, and E. Kerola (2014a). "Does Bank Ownership Affect Lending Behavior? Evidence from the Euro Area." *Journal of Banking & Finance* 48, 194–209.

Ferri, G., P. Kalmi, and E. Kerola (2014b). "Organizational Structure and Exposure to Crisis Among European Banks: Evidence from Rating Changes." *Journal of Entrepreneurial and Organizational Diversity* 3, 35–55.

Finance Finland (2017). Pankkien markkinaosuudet 2016. https:// www.finanssiala.fi/materiaalit/FK-Pankkien-markkinaosuudet-2016.pdf, last accessed on August 27, 2018.

Finnish Competition and Consumer Authority (2016). FCCA's Decision Narrows Banks' Possibilities to Increase Charges and Fees for Consumer Credit. https://www.kkv.fi/en/current-issues/press-releases/2016/28.11.2016-fccas-decision-narrows-banks-possibilities-to-increase-charges-and-fees-for-consumer-credit/, last accessed on July 1, 2020.

Finnish Competition and Consumer Authority (2019). The Finnish Competition and Consumer Authority Terminates Investigation into the OP Financial Group Bonus System. https://www.kkv.fi/en/current-issues/press-releases/2019/the-finnish-competition-and-consumer-authority-terminates-investigation-into-the-op-financial-group-bonus-system/, last accessed on June 26, 2020.

Gehrig, T., and R. Stenbacka (2007). "Information Sharing and Lending Market Competition with Switching Costs and Poaching." *European Economic Review* 51, 77–99.

Hannan, T. H., and R. M. Adams (2011). "Consumer Switching Costs and Firm Pricing: Evidence from Bank Pricing of Deposit Accounts." *Journal of Industrial Economics* 59, 296–320.

Independent Commission on Banking (2011). Final Report: Recommendations. http://bankingcommission.independent.gov.uk/, last accessed on June 24, 2020.

Jones, D. C., I. Jussila, and P. Kalmi (2016). "The Determinants of Membership in Cooperative Banks: Common Bond versus Private Gain." *Annals of Public and Cooperative Economics* 87, 411–432.

Kari, M., I. Kiema, P. Kuoppamäki, and E. Lehto (2017). Customer Loyalty Programs and Consumers. Labour Institute for Economic Research Reports No. 36, Helsinki.

Kim, M., D. Kliger, and B. Vale (2003). "Estimating Switching Costs: The Case of Banking." *Journal of Financial Intermediation* 12, 25–56.

Kiser, E. K. (2002). "Predicting Household Switching Behavior and Switching Costs at Depository Institutions." *Review of Industrial Organization* 20, 349–365.

Kopsakangas-Savolainen, M., and T. Takalo (2014). "Competition Before the Sunset: The Case of the Finnish ATM Market." *Review of Network Economics* 13, 1–33.

Rintakoski, K. (2015). "Kilpailuvirasto OP:lle: Marginaaleista ei saa puhua." *Taloussanomat*, 28 October. https://www.is.fi/taloussanomat/art-2000001892948.html, last accessed on June 26, 2020.

Ruiz-Aliseda, F. (2016). "When Do Switching Costs Make Markets More or Less Competitive?" *International Journal of Industrial Organization* 47, 121–151.

Saarinen, M. (2014). "Pankit nostavat urakalla palvelumaksuja – Viranomainen huolissaan." *Talouselämä*, 30 August. https://www.talouselama.fi/uutiset/pankit-nostavat-urakalla-palvelumaksujaan-viranomainen-huolissaan/b8cedcc0-5595-3407-9db8-6123b6b774b1, last accessed on June 30, 2020.

Savolainen, E. (2016). "Suomen luottolaitosten keskittyneisyyden mittareita." *Euro&Talous*, 30 November. https://www.eurojatalous.fi/fi/2016/artikkelit/suomen-luottolaitossektorin-keskittyneisyyden-mittareita/, last accessed on July 2, 2020.

Shy, O. (2001). *The Economics of Network Industries*. Cambridge, UK: Cambridge University Press.

Shy, O. (2002). "A Quick-and-Easy Method for Estimating Switching Costs." *International Journal of Industrial Organization* 20, 71–87.

Silva, M. O., and C. R. Lucinda (2017). "Switching Costs and the Extent of Potential Competition in Brazilian Banking." *EconomiA* 18, 117–128.

Stenbacka, R. and T. Takalo (2019). "Switching Costs and Financial Stability." *Journal of Financial Stability* 41, 14–24.

Stenvik, S. (2016). Estimating Switching Costs in the Finnish Retail Banking Industry. Hanken School of Economics, Department of Economics, Economics Master's Thesis, August 31, 2016.