Labor Supply Responses and Adjustment Frictions: A Tax-Free Year in Iceland*

Jósef Sigurdsson[†]

September 21, 2019

Abstract

Labor income earned in Iceland in 1987 was not taxed. I use this to study labor supply responses to temporary wage changes. I construct a new population-wide dataset of earnings and working time from digitized payslips and use two identification strategies to estimate intensive and extensive margin Frisch elasticities of 0.37 and 0.07, respectively. These average responses are driven by those with the ability to adjust in response: extensive margin responses are driven by young and close-to-retirement cohorts and intensive margin responses are driven by workers in temporally flexible and hourly paid jobs, though take-up of secondary jobs contributes to one-tenth of the overall response. Finally, I find that married women with children respond more than husbands, who themselves respond negatively to wives' tax cuts, consistent with substitutability in non-market time. Overall, the results suggest that adjustment frictions and household interdependencies reduce aggregate labor supply responses to tax cuts.

JEL Classification: E65, H24, H31, J21, J22

*This paper is a revised version of the first chapter of my Ph.D. thesis at Stockholm University. I am indebted to Emi Nakamura, Arash Nekoei, Torsten Persson and Jón Steinsson for invaluable advice, guidance and encouragement. For their helpful comments, suggestions and discussions, I thank Ingvild Almås, Josh Angrist, Richard Blundell, Konrad Burchardi, Mitch Downey, Peter Fredriksson, Karl Harmenberg, Xavier Jaravel, Per Krusell, Hannes Malmberg, Peter Nilsson, Robert Östling, Jon de Quidt, Benjamin Schoefer, Fabian Sinn, David Strömberg, Danny Yagan, participants of the 2019 NBER Summer Institute, joint labor and public economics session, and numerous seminar and conference participants. I thank Statistics Iceland for the collaboration and hospitality over multiple visits between the fall of 2016 and the spring of 2018, during the process of data construction. In particular, I thank Hrafnhildur Arnkelsdóttir and Margrét K. Indriðadóttir for their help in making this project possible. Special thanks go to Ómar Harðarsson, Stefán Jansen, Hildur Sigurðardóttir, Eyjólfur Sigurðsson and other staff members at Statistics Iceland for valuable discussions.

[†]Bocconi University and Central Bank of Iceland. Email: josef.sigurdsson@unibocconi.it

1 Introduction

One of the longest-standing questions in economics asks how people adjust their labor supply in response to temporary changes in pay. This response is summarized by the intertemporal elasticity of substitution of labor supply, or the *Frisch elasticity*. Knowing this elasticity is crucial to understanding how aggregate shocks and economic policies affect output and welfare in the economy. For example, the size of the Frisch elasticity is important for evaluating the consequences of pension reforms (e.g. Imrohoroğlu and Kitao, 2012), for determining optimal taxes on capital and labor income (e.g. Conesa, Kitao, and Krueger, 2009; Heathcote, Storesletten, and Violante, 2017; Stantcheva, 2017), and it is a crucial parameter in the research on cyclical movements in employment and wages (Lucas and Rapping, 1969).

Macroeconomic models with labor market clearing require a large Frisch elasticity to rationalize large business-cycle movements in employment with only modest movements in wages. But existing microeconometric estimates are often small and statistically insignificant.[1] A wide range of views arise from the fact that estimating the Frisch elasticity is notoriously difficult. Causal estimation requires an exogenous temporary change in the after-tax wage with limited income effects. Exogenous wage changes are hard to find, let alone those that are transitory.

In part, disagreement on the Frisch elasticity reflects differential views on adjustment frictions. Frictions are less likely to attenuate short-run responses to large changes (Chetty, 2012) and inattentive individuals are more likely to respond to changes that are salient and simple (Chetty, Looney, and Kroft, 2009). These features cloud how much we can learn about labor supply by studying tax reforms, unless they are sufficiently large and salient.

To tackle these difficulties, I exploit a tax-reform in Iceland resulting in a year free of labor income taxes. In 1986, the Icelandic government announced a tax reform, replacing a system where this year's taxes were based on last year's income by a pay-as-you-earn withholding-based system. To ensure that during the transition year of 1987, workers would not have to pay taxes simultaneously on their 1986 and 1987 earnings, no taxes were collected on 1987 labor incomes. As illustrated in Figure 1, the income earned in 1987 was tax-free. This tax-free year created a strong incentive for intertemporal substitution of work, but a minimal income effect. First, there was no windfall gain for taxpayers, as those earning the same in 1987 as in 1986 did not see a change in their cash-flow. Second, the reform did only imply a small change in life-time income. Therefore, the tax-free year offers a rare natural experiment suitable to estimate the Frisch elasticity.

To exploit this experiment, I built a new population-wide dataset. I constructed new data on the universe of workers and firms from payslips found at Statistics Iceland, which I converted into a machine-readable data set. Information on all pay and all working time in all jobs makes this an ideal data set to study labor supply. Combining this with individual data from tax returns, I obtained a new employer-employee panel data set for the entire workforce from 1981 until today. These rich data enable me to uncover the details of labor-supply adjustment along multiple dimensions.

I use two complementary quasi-experimental research designs to identify labor supply elasticities

[1] For a recent survey of the microeconometric literature, see, e.g., Keane (2011). Keane and Rogerson (2015) provide a recent discussion of the micro-macro controversy.

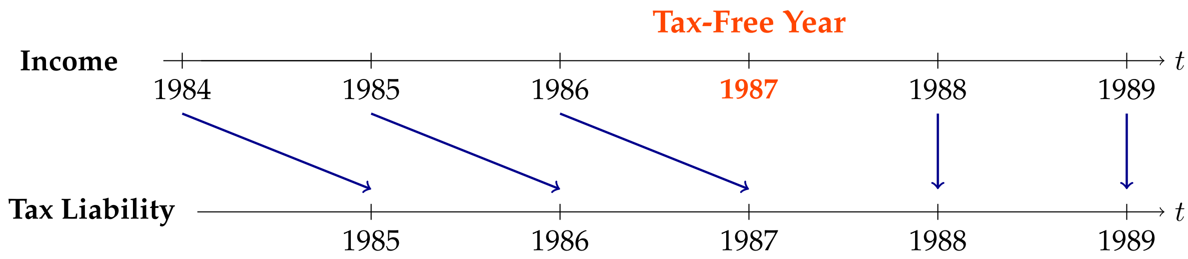


Figure 1: Income tax system before and after the tax reform

along the intensive margin (i.e. working hours among those working) and the extensive margin (i.e. labor force participation and employment arrangement). Figure 2 schematically describes my empirical approach.

First, the *tax-bracket difference-in-differences (DD)* design exploits cross-sectional variation in the size of tax-cuts arising from progressivity of the tax schedule, building on the seminal idea in Feldstein (1995). More precisely, while all workers were given a tax-free year in 1987, workers in a high tax bracket prior to the reform, receiving a large increase in after-tax wages, are expected to respond more strongly than those in a lower tax-bracket. I use this difference across groups in a difference-in-differences design. Relating the dose-responses to differences in intensities of marginal tax-rate changes enables me to identify labor supply elasticities. A key advantage of this design is to difference out aggregate trends in employment and macro shocks around this time. My detailed population-wide data and methodological approach distinguish my study from the previous work of Bianchi, Gudmundsson, and Zoega (2001) who study a small random sample of workers and compare their outcomes in the tax-free year to those in the year before and after.

Large and salient changes in after-tax wages, such as those resulting from the tax-free year, are likely to deliver responses closer to those in a frictionless labor market. However, frictions are still likely to shape the margins of labor adjustment and heterogeneity in responses. In particular, adjustment costs and indivisibilities may lead to large extensive margin elasticities. As the tax-bracket DD design exploits the variation in tax rates across groups of workers that are employed prior to the reform, it cannot, by construction, identify entry responses. This is an important limitation as the tax-free year generates a strong incentive for labor-market entry. Furthermore, analyzing heterogeneity in responses is also restricted by this design. If we are to paint a complete picture of intertemporal labor supply, studying responses along the relevant margins and dimensions of heterogeneity is necessary.[2]

To overcome these issues, I develop a new research design: *life-cycle difference-in-differences (DD)*. This design builds on two features of my setting. As the tax reform was unanticipated, the timing of the tax-free year is plausibly exogenous from an individual's life-cycle perspective. Moreover, as the labor supply of similar individuals is likely to evolve similarly over their life-cycle, labor supply

[2]A candidate explanation for the micro-macro divergence is that macro elasticities incorporate responses along all margins (Rogerson and Wallenius, 2009) while micro elasticities typically reflect restricted samples and specific sources of identifying variation. As I document in Section 6.2, almost all previous work is limited to the analysis of prime-aged men or particular occupations, such as bicycle messengers or taxi drivers, for identification and data-limitation reasons.

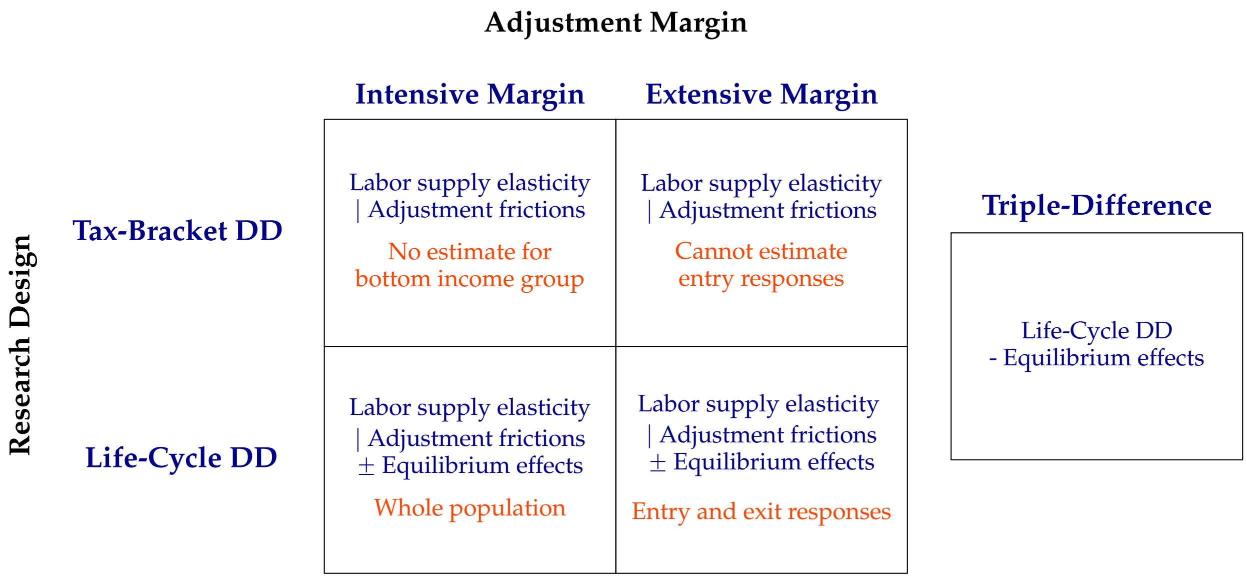


Figure 2: Overview of Empirical Methodology

elasticities can be identified by comparing two similar individuals at the same point in their life-cycle, one of whom receives an unexpected wage shock but the other does not. As an example, the labor supply elasticity for an individual who was 40 years old in 1987 can be estimated by matching him to an observationally similar individual with the same characteristics but who was aged 40 in 1986. The advantage of this method is to be able to identify labor supply responses for the whole population, including labor market entry.

Using the tax-bracket DD design, I estimate strong responses in both labor earnings and working time, with an intensive-margin elasticity at 0.37. Decomposing this estimate into different margins, I find that 30% of the overall response stem from more weeks of full-time work. This includes transitions from part-time to full-time employment, the exchange of vacation for working time and weeks worked on secondary-jobs. 70% are accounted for by more earnings within full-time weeks, such as over-time hours, more shifts and more effort induced on the job. I also demonstrate that the higher earnings reflect labor supply rather than reporting responses. Using the life-cycle DD design, I find that extensive margin responses are modest on average, with an employment semi-elasticity of 0.07. However, this modest average response masks an important heterogeneity. Most of the responses originate among those close to retirement and, in particular, cohorts younger than 25, who are still in school or out of the labor force for other reasons. Thus, the size of estimated elasticities will depend on the density of individuals at the tails of the work life-cycle distribution.

A reform as extensive as the tax-free year is likely to generate equilibrium effects. However, it is not clear ex ante in which direction these effects go. On the one hand, an increase in labor supply could lead to a reduction in wages if labor demand is not perfectly elastic. This may dampen the labor supply response to the tax-free year and lower the estimated elasticity. On the other hand, in partial equilibrium, workers receiving large tax cuts may spend less time on leisure but also on home production, such as home cleaning, cooking and child care. In all likelihood, this generates a demand for labor input in the sectors providing those goods and services, thus making it possible for individuals to work more during a tax holiday if they so desire, increasing the estimated elasticity. The life-cycle DD incorporates all such equilibrium effects, as well as other aggregate effects in the tax-free year, which are 'differenced out' in the tax-bracket DD. Comparing elasticity estimates using the two methods thus allows me to gauge the magnitude of these aggregate effects. I can do this by direct comparison or using a method that combines the two designs in a triple-difference. Both methods give the same difference, namely 0.1 – i.e., about one-fifth of the intensive-margin response.

My estimates of the average Frisch elasticities in the economy are most closely related to two studies.[3] In an important contribution, Bianchi, Gudmundsson, and Zoega (2001) highlight the rare opportunity offered by the Icelandic tax-free year to study labor supply. Among a small random sample of workers, they find that people work much more in the tax-free year than in the preceding and following year. Indeed, in the meta-analysis of Chetty et al. (2013), the Icelandic tax-free year is one of few informative data points on intertemporal labor supply. What distinguishes my study from theirs is that I use detailed population-wide data and an empirical approach comprised of two complementary identification strategies. This enables me to provide a more robust analysis of how Icelandic workers responded to a temporary incentive to work, as well as leading to some new insights into the anatomy of labor supply responses. In a study contemporaneous to mine, Martinez, Saez, and Siegenthaler (2018) study labor supply responses to a tax reform in Switzerland, leading to a two-year tax holiday. My results contrast theirs, which imply a modest response on average to this strong work incentive. I discuss this further in Section 6.2 and I argue that differences in labor-market flexibility constitute the most plausible explanation for the differences in labor supply elasticities of the average Icelandic and Swiss worker. The Icelandic labor market is more flexible and less regulated than labor markets on mainland Europe, closer to what is found in the US. I find that within Iceland, there are substantial differences in labor supply responses depending on workers' flexibility in work arrangements. Similarly, Martinez, Saez, and Siegenthaler (2018) report a relatively large elasticity for the self-employed subpopulation, a group which likely has more flexibility in labor adjustment than the average worker. Similar differences are likely to be found across countries as well.

The combination of a rare setting offered by the tax-free year and rich data enables me to go beyond estimating average labor-supply elasticities and trace out the mechanisms that drive the strong labor supply responses I find. My analysis is guided by a combination of machine-learning methods and causal estimation. More precisely, after obtaining estimates of labor supply elasticities at the individual level using the life-cycle DD method, I use the random forest algorithm (Breiman, 2001) to highlight the most important features shaping differences in labor supply responses. Using those features as a guide, my analysis yields three main results that provide information to ongoing debates in labor economics.

First, I find that individuals with low labor-market attachments – including the young cohorts and workers close to retirement – that work less than full time prior to the reform, are very responsive.

[3]In Section 6.2 I provide a more detailed summary and a meta-analysis of previous work.

In particular, the youngest cohorts, of which many were part-time employed, respond very strongly with elasticities as high as 2.

Second, I find that flexibility of jobs is key in shaping the labor-supply responses. My results demonstrate that the responses are strongest among workers in jobs with a more temporal flexibility – i.e. those with an a priori high ability to adjust their hours – and those with labor market contracts that build in compensation for marginal hours worked. These workers are less likely to be bound by hours constraints and have a greater ability to get compensated for additional hours in their primary jobs. The largest responses are therefore concentrated among exactly those groups that would be predicted by theory. However, prior evidence on how frictions influence intertemporal labor-supply responses is very limited.

While hours may be rigid within jobs, they may become flexible by workers holding multiple jobs. I find that workers indeed overcome hours constraints in their primary jobs by working at a secondary job. In addition, I find a reduced probability of primary-job change in the tax-free year, consistent with workers being less willing to engage in a time-consuming job search. When I decompose the overall responses, I find that one third of the increase in weeks worked and one tenth of the total increase in labor earnings are created by work on secondary jobs, while the remainder results from more work on continuing primary jobs. While taking a secondary job may be an important avenue for the labor supply adjustment of many workers, existing evidence on multiple-job holding and the macroeconomic importance of this margin remains very scarce. These results are important when viewed in the light of resent changes in the structure of labor markets. As a growing proportion of the workforce holds jobs in the "gig economy" – working through contracting and temporary arrangements – often alongside their primary jobs (Katz and Krueger, 2016; Hall and Krueger, 2018), labor supply elasticities accounting only for primary employment will become inaccurate descriptions of employment responses.[4]

Third, I estimate how family structure shapes labor supply responses. Married women have higher elasticities than married men, in part reflecting part-time and non-employment of mothers. However, single men and single women have similar elasticities. An extensive literature has studied gender differences in labor supply elasticities (see e.g. McClelland and Mok, 2012; Blundell and MaCurdy, 1999, for review). This result highlights that the gender differences in labor supply found in previous work may not reflect inherent differences, but rather reflect family ties and specialization in the household.[5] I estimate cross-elasticities for married couples and find negative cross-elasticities for husbands but no significant cross-responses for wives. This result is inconsistent with models of unitary household labor supply, which have the strong prediction that cross-elasticities of spouses should be symmetric (Chiappori and Mazzocco, 2017).[6] Accounting for both own and cross-

[4]Recent studies have found that workers take up secondary-jobs, such as ride-hailing for Uber, because of the flexibility they provide (Hall and Krueger, 2018) and in order to mitigate frictions and volatility in income on primary jobs (Farrell and Greig, 2016; Koustas, 2018). In addition, Angrist, Caldwell, and Hall (2017) estimate a large labor supply elasticity among Uber drivers, indicating that labor supply may be very elastic in secondary-jobs.

[5]Blau and Kahn (2007) and Heim (2007) document that the labor supply elasticities of married women have shown a decreasing trend in recent decades with rising participation rates.

[6]This finding complements a large empirical literature testing the restrictions imposed by unitary household models. For reviews see, e.g., Donni and Chiappori (2011) and Chiappori and Mazzocco (2017).

responses, my results imply that total household responses are about 20% smaller than if spouses would have been treated in isolation. Taken together, these results indicate coordinated labor supply responses within the household and substitutability in the non-market time of husbands and wives.

The paper proceeds as follows. Section 2 describes the empirical setting and the reform that gave rise to the tax-free year and Section 3 describes the data set I have constructed. Sections 4 and 5 explain the tax-bracket and life-cycle DD designs, respectively, and present estimates of labor supply elasticities obtained from these two designs. Section 6 discusses my estimates of average Frisch elasticities and puts them in the context of the findings in previous work. Section 7 studies the mechanisms behind the strong labor supply responses, highlighting the importance of heterogeneous adjustment frictions. Section 8 concludes the paper. Some additional background material and auxiliary analyses are relegated to an Online Appendix.

2 The Tax-Free Year and Background

2.1 Income Tax System and Tax Reform

On January 1, 1988, Iceland took up a withholding-based pay-as-you-earn income tax system, similar to what is now in place in most advanced economies. Prior to the reform, income taxes were collected with a one-year lag. The tax liability and tax payments due every month in year t were computed based on year $t-1$ income. This system resembled those in place in most developed countries prior to adopting a modern pay-as-you-earn tax system. When announcing the tax reform, the authorities also announced that labor income earned in 1987 would not be taxed. As Figure 1 depicts, this implies that while people were paying taxes every year, including in 1987 when they paid taxes based on their income earned in 1986, they would take home tax-free whatever they earned more in 1987 than in 1986.

The key features of the reform for the purpose of my analysis are that it generated a large, salient and unanticipated increase in wages that lasted only a single year. On December 6, 1986, the Finance Minister announced the tax reform. The Ministry of Finance began preparing the reform in the early fall of 1986 and later in the fall, the decision was made for it to take place in January 1988. The reform was therefore unanticipated by the taxpayers. Figure 4 plots the monthly count of the number of newspapers mentioning a withholding-based or pay-as-you-earn tax system between January 1980 and December 1988. As the figure documents, there was no discussion of a reform of this kind in the years before its announcement, whereas 30-40% of the newspapers printed in the weeks following the announcement had coverage of the reform.

The reform was very salient. Newspapers printed headlines such as "*A Tax-Free Year*" and "*Pay-as-you-earn tax system in 1988 – all income in 1987 tax-free*". In addition, the tax authorities sent out advertisements and explanatory flyers, as exampled in Figures A.11 and A.12 in the Online Appendix. These also advertised that a prerequisite for tax freedom was that taxes were filed for 1987 as usual. This was important as other taxes, such as those on capital income and wealth, and benefits were unchanged in 1987. From the perspective of my study, the quality of administrative data in 1987, such as tax returns, was not influenced by the reform.

The tax-free year generated a strong incentive for intertemporal substitution. The average tax rate fell to zero from about 10 percent, increasing the incentive for employment (extensive-margin). On the intensive margin, the changes in incentives were even stronger, as the after-tax wage increased by about 20 percent on average. While the whole population received an increase in wages, some workers received a larger cut than others due to the progressivity of the tax system. It is by harnessing these differences that I identify the intensive-margin labor supply elasticity. Furthermore, the tax-free year did not create an income effect for individuals that are myopic in their decision making. There was no windfall gain for taxpayers, as those earning the same in 1987 as they earned in 1986 did not see a change in their cash-flow.[7] In addition, taxes were only cut temporally for a single year, allowing me to study labor supply responses during that year. A one year change in incentives is the relevant frequency for a business cycle analysis of employment fluctuations. As a result of all these features, the tax-free year comes close to being the ideal natural experiment to study intertemporal substitution of labor supply and estimate Frisch elasticities.

The only change to the tax system made in 1987 was that income taxes were temporarily set to zero. However, the reform was accompanied by a simplification of the tax system that was put in place after the tax-free year. These changes were being worked out during the first months of 1987 as part of adapting the old tax system to tax withholding. The simplifications consisted of two main changes. First, the reform abolished a large share of deductions that could be made to taxable income before arriving at the tax base. Second, the progressive tax schedule was replaced with a flat tax. To summarize, the reform changed both the tax base and the tax rate, where the aim was to simplify the tax system but leave the average tax burden unchanged.

I argue that these changes are unlikely to influence people's responses to the tax-free year and my estimates of Frisch elasticities. The effects on later taxes were not as obvious and clear-cut as the tax free year. Understanding the effect on tax payments would involve understanding the interaction of tax deductions, tax allowances and tax rates which influenced the tax burden in opposing directions. Relatedly, these changes were much less salient than the tax-free year. Figure 4 shows that a change to a flat tax received limited media attention. Moreover, flyers and explanatory material from the tax authorities emphasized the fact that income in 1987 was tax-free and the changes in the structure of tax collection in 1988, but contained no information about changes in the tax schedule after 1987. As discussed in Section 4 and detailed in Online Appendix E, I perform a series of tests to evaluate this claim, finding my results to be robust to these concerns.

Details of the tax system pre and post the tax-free year are found in Online Appendix A and a discussion of the reform and the time-line of events in Online Appendix B.

2.2 Icelandic Labor Market in International Context

The Icelandic labor market is quite flexible, characterized by low unemployment, flexible hours, and variable participation and wages (OECD, 1991, 2007).[8] In this sense, its characteristics are more sim-

[7]Similarly, the reform did not influence the government's budget, as the tax revenue flows were uninterrupted.

[8]For an overview of the Icelandic economy, including characteristics of the labor market, see e.g. various previous issues of *Economy of Iceland*.

ilar to the US than to mainland Europe. The flexibility of the labor market has long played a key role in the rapid adjustment to macroeconomic shocks.[9]

Labor force participation in Iceland is high, exceeding 80 percent of the working-age population. The overall participation grew steadily until the mid 1980s, primarily due to the increased participation of women, who by the beginning of the 1990s accounted for close to half of the labor force, although a smaller share of total hours. Relative to the OECD countries, female participation is among the highest, as well as the participation rates among the young and the elderly.

In comparison with the other OECD countries, Icelandic firms have a considerable flexibility to lay off workers. Firms can easily adjust their level of labor input over the business cycle, either by hiring and firing of workers or by adjusting the number of hours of current employees. Changes in hours per worker account for about half of the variation in employment over the business cycle.

The labor market is highly unionized. Collective bargaining between the umbrella unions on both sides of the market negotiates general employee rights and minimum wages. However, this sets the base for wage bargaining at lower levels, such as in sectors and firms, where the flexibility to account for local conditions is greater. Therefore, in spite of this centralization, real wages are very flexible in Iceland as compared to other OECD countries (Central Bank of Iceland, 2018).

3 Data

For the purpose of this project, I construct a new administrative data set on the universe of the Icelandic working age population back to 1981. The data set has two main sources: an employer-employee data set constructed from newly digitized payslips, and individual tax records. In addition to these main sources, I draw on additional data, including Statistics Iceland's Education Register and the Population Register from the National Registry. Below I describe the two main data sets in detail.

3.1 Payslips: Employer-Employee Data

At the end of each year, all employers are obligated to compile a payslip for each employee in their establishment, or for every job if the employee holds more than one job at the establishment. This applies to all firms and establishments, including self-employed workers. Employers send copies of payslips both to the respective employee and to the Directorate of Internal Revenue. Information from payslips is then used as inputs for many purposes, such as for individuals' income taxation, the computation of accident insurance, and the computation of firms' payroll taxes.

Since the early 1990s, an increasing number of employers compile and send payslips to the Directorate of Internal Revenue in a machine-readable format, and currently almost all payslips are received electronically. Before that time, in the 1980s and in the early 1990s, all payslips were compiled in paper format. The records were then stored in various forms, including on magnetic tape cartridges and mainframe tapes. In collaboration with Statistics Iceland, I have converted all payslips

[9] As an example of this emphasis, the Director of the European Department of the International Monetary Fund (IMF), Thomsen (2018) notes in a recent speech that "*While I was not familiar with Iceland's economic history before coming here, we soon realized that Iceland had a history of quickly adjusting to shocks, not least because of labor market flexibility.*"

back to 1981 into data in a machine-readable form. The resulting product is a panel data set covering the universe of jobs in Iceland, connecting all employers and their employees, for each year from 1981 to 2015.

Payslips contain information on all labor earnings and related compensation. This includes wage payments, contractor payments, piecework pay in fishing, pension payments, bonuses and commission, remuneration to a company's board members and accountants, travel allowances and other allowances (car, clothes, food, etc). Each of these components is reported separately on each payslip for a given job. In addition, and importantly for the current project on labor supply, the payslips also contain information about working time in each job. Time is measured in weeks worked, with the reference week amounting to 40 working hours. Employers are obligated to report the number of weeks an employee worked on a given job based on his actual working time during the year and employment arrangement, such as part-time employment. The same is true for self-employed workers that must report working time in the same way for themselves as well as for their spouses and children that may work for them. A worker can at most be recorded working 52 weeks on a given job during the year. However, workers can have more than one job and therefore be registered as working more than 52 weeks. For example, a full-time employee holding a single job and working at least 40 hours per week is recorded as working 52 weeks. Another worker that holds two part-time jobs on which he works a parallel 20 hours per week is recorded as working 26 weeks on each job (reported separately) and 52 weeks in total.

The reason why employers (and self-employed workers) were required to report the working time of their employees was twofold. First, the calculation of worker's accident insurance fees was based on the number of weeks an employee worked during the year. This insurance covered accidents, leading to a worker's injury or death, that occurred on the job or on the way to or from work. The insurance fee, which was updated every year on January 1 and paid by the employer, varied from job to job and differed by occupation and the risk of injury and accidents in a given job.[10] Second, the payroll tax levied on firms to fund the public unemployment insurance system was based on the total number of weeks worked by all workers in the given firm in each year. In contrast to the insurance fee, this tax was independent of occupation and sector.[11] Therefore, the number of weeks registered for each worker on his payslip is to reflect the number of weeks worked during the year rather than the number of weeks employed. In addition, these are the only universal data on employment and labor input by sectors and occupations based on which official statistics are constructed, which put pressure on correct filing.

Each payslip includes a unique personal identifier of the worker and a unique identifier of firms. In addition to the detailed information on payments and working time, payslips include demographic information about workers, including occupation according to a two-digit classification based on the International Standard Classification of Occupations (ISCO), and the firm, including the firm's sector according the International Standard Industrial Classification of All Economic Activities (ISIC).

[10]As an example, in 1987 the insurance payment for a blue-collar factory worker per week worked amounted to about 0.14% of his average monthly earnings which was more than threefold the weekly fee for office clerks.

[11]In 1987, this tax per week worked equalled 0.31% of the average weekly earnings of blue-collar workers.

3.2 Individual Tax Returns

The second primary data source I use in this paper is a panel of individual tax returns. Like the data set I have constructed from payslips, these data extend back to 1981. The data sets are easily linked via a unique personal identifier. Individual tax returns have information on all income, including labor income, financial income, pension, social security and transfer payments as well as other sources of income. These data also record all tax payments, both at the national and local level, as well as deduction and tax allowances. I use this detailed data to construct marginal tax rates.[12] Because a wealth tax was levied in Iceland during most of my sample period, and in periods when a wealth tax has not been levied, the structure of tax returns has not been altered and the data set includes detailed information on all assets and liabilities back to 1981. In addition, the tax records include a range of demographic variables, as well as family identifiers linking married or cohabiting couples. Further information about the data sets and summary statistics are provided in Online Appendix C.

4 Tax-Bracket Difference-in-Differences

In this section I estimate labor supply responses using a difference-in-differences research design which exploits the intensity at which workers' after-tax wages were influenced by the tax-free year and the dose-response in labor supply.

4.1 Research Design

In general, to identify the causal effect of the tax-free year on labor supply, a proper counterfactual is needed for what would have happened in its absence. Alternatively, if the population is treated with different 'doses' of tax cuts, causal effects can be identified from the differential treatment intensity, provided that they generate differential responses. In the current context, while the whole Icelandic population was given a tax-free year in 1987, non-linearities in the pre-reform tax schedule generated substantial differences in the changes of after-tax wages. Therefore, this setting naturally lends itself to a difference-in-differences research design. My strategy follows a strand of literature dating back to the seminal paper of Feldstein (1995), exploiting a cross-sectional variation generated by the 1986 tax reform in the US.[13]

The tax schedule prior to the reform was progressive with four brackets, consisting of three national level brackets and a local-level municipal tax. Taxable income, from both labor and capital, was taxed in the same way at the national and municipal levels. As detailed in Section 2, all taxes on labor income were set to zero in 1987.

Figure 5a plots the evolution of tax rates by tax brackets from 1981 to 1990. In 1986, the average worker in the bottom tax bracket faced a marginal tax rate of 10.2%, corresponding to the average mu-

[12]Since marginal tax rates are not directly observed in individuals' tax returns, I have built a tax calculator for the Icelandic tax system to construct marginal tax rates. This method predicts actual tax liabilities with great precision. See Online Appendix C.1 for details.

[13]For a summary and discussion of the empirical literature on labor supply and taxable income elasticities see, e.g., Saez, Slemrod, and Giertz (2012).

nicipal tax rate, while the average tax payer in the top bracket faced a marginal tax of roughly 48.7%.[14] As documented in the figure, while tax rates had been on a slightly decreasing trend throughout the 1980s, the difference across brackets had remained stable. Tax rates were frequently reviewed in relation to the government's budget and tax-bracket thresholds, which were set in nominal values, were generally reviewed and updated each year to account for changes in prices and wages. As a result, which I document in Figure 5b, the tax-bracket thresholds corresponded to roughly the same income percentile throughout the 1980s and therefore the income groups in each bracket were stable and similar over time.

Assigning treatment status. My empirical strategy to estimate elasticities is to relate the differential labor supply responses of workers in higher vs. lower tax brackets to their differential tax relief. As the tax rates faced each year are endogenous to labor income, which is my outcome of interest, I follow Feldstein (1995) and later work by assigning treatment status based on a lagged tax bracket. The lagged tax bracket is unrelated to current income. Since income and other factors influencing the tax-bracket position are persistent, the tax-bracket position is persistent as well, as documented in Figure A.13 in the Online Appendix. As a result, a lagged tax bracket serves as a valid and strong instrument for the current tax bracket. In my main analysis, the treatment group consists of workers that faced marginal taxes in the three top brackets, while workers in the bottom bracket constitute the main control group. In order to get a larger sample size for inference and detailed later analysis, I pool together the estimates for the three top tax-brackets, under the assumption that labor supply elasticity is the same across tax brackets, providing a weighted average elasticity. In addition, I will estimate disaggregated responses, by tax bracket, as well.

Sample and restrictions. With the aim of analyzing a sample of comparable workers facing different tax rates, I restrict my sample of the working-age population, age 16-70, in two ways. First, I use a balanced sample of individuals observed in all years 1981 to 1987. Since everyone aged 16 years and older is required to file taxes, independent of their labor market status, this excludes workers that die, emigrate from Iceland or young people that are not observed during the pre-treatment period and for whom trends in labor supply cannot be assessed. Second, for each of the pre-reform years, I restrict the sample to workers that were employed in the previous year, defined as having labor earnings greater or equal to a base-income threshold, roughly corresponding to minimum-wage earnings for a low-skilled worker.[15] Thus, I obtain a sample of workers in one of the four brackets that can be assigned a treatment status in the current year. Restricting the sample in this way corresponds to restricting the sample to those with earnings above the 20th percentile, including zeros. Since the unemployment rate was between 1% and 2% throughout the 1980s (Figure A.9 in Online Appendix), this restriction mainly serves as a means of excluding those entering and exiting the labor market due to life-cycle patterns, which may generate differential trends across tax brackets depending on

[14] In 1986, the municipal tax rate ranged from 5 to 11.5%.

[15] Similar restrictions are frequently imposed in studies of the core labor force, see e.g. Kindlund and Biterman (2002). The base income threshold equals $1.5\times$ guaranteed income (*tekjutrygging*), where guaranteed income is a reference amount used in calculations of various kinds of income support provided by the government and the municipalities, such as for the elderly and disabled. Using the guaranteed income as a reference point has the advantage compared to, e.g., minimum-wage earnings by sectors and occupations, that it is updated each year to account for inflation.

where workers enter and exit. In my analysis, I define employment in the same way when estimating extensive margin responses, i.e. having labor earnings exceeding this threshold. The research design developed in Section 5 does not rely on a tax-bracket comparison and can be used to study labor supply responses for the whole working-age population, both along the intensive and extensive margins.

Before proceeding, some unique features of my empirical setting relative to previous work estimating labor supply elasticities using tax reforms are worth highlighting. During the tax-free year, all tax-brackets were collapsed into a single bracket with a zero marginal (and average) tax rate. Importantly, taxes are zero in 1987 independent of the earnings that year. In most settings, this is not true, requiring the researcher to construct an instrument for the tax rate in the treatment period (Gruber and Saez, 2002). In addition, as most previous work studies long-term responses to permanent tax changes, income shocks may move workers between brackets, influencing the after-tax wage and therefore the empirical estimates. In my study, which is focused on short-term responses within the tax-free year, concerns about such switches are not relevant.

Estimating equation. The reduced-form labor supply responses to the tax-free year are estimated using the following differences-in-differences (DD) specification

$$y_{it} = bracket_{i,t-1} + \delta_t + \eta \cdot B_{i,t-1} \times \delta_{t=1987} + \boldsymbol{X}'_{it}\gamma + \mu_{it} \tag{1}$$

where y_{it} is the outcome of interest of individual i in year t, $bracket_{i,t-1}$ is an indicator function for tax brackets in year $t-1$ (treatment status), and δ_t are time fixed effects included to control for time effects affecting all individuals. The identification of the labor supply response to the tax-free year is brought by η, the coefficient on the interaction of $B_{i,t-1}$, which is an indicator function for being in one of the top three tax brackets, interacted with a dummy for the tax-free year of 1987. The regression controls for individual characteristics, collected in the vector $\boldsymbol{X}_{it}$, which includes a full set of dummies for individual characteristics such as age, marital status, number of children, education, living in the capital area, and, in some cases, occupation and sector of employment. Since these variables, apart from age, may themselves be influenced by the reform, they are defined in pre-reform levels. The error term is denoted by μ_{it} and captures other determinants of labor supply. The importance of accounting for serial correlation in outcomes in a DD setting has been emphasized by Bertrand, Duflo, and Mullainathan (2004). I cluster standard errors at the individual level to allow for an arbitrary correlation over time in the error term.

In order to obtain an elasticity estimate, I relate differential labor supply responses – i.e. the dose-response – to the differential increase in the after-tax wage generated by the tax-free year. Intuitively, in its simplest form, the elasticity estimate corresponds to the Wald estimator, which is the ratio of the reduced form and first stage, that can be obtained from estimating equation 1. Following this logic, I employ the following two-stage least squares (2SLS) difference-in-differences specification

$$y_{it} = bracket_{i,t-1} + \delta_t + \varepsilon \cdot \log(1-\tau_{it}) + \boldsymbol{X}'_{it}\gamma + \nu_{it} \tag{2}$$

where τ_{it} is individual i's marginal tax rate in year t. Instrumenting the log net-of-tax rate $\log(1-\tau_{it})$

with the reduced-form interaction $B_{i,t-1} \times \delta_{t=1987}$, the coefficient ε identifies the elasticity (e.g. of labor supply) to a change in the net-of-tax wage.

4.2 Results

Graphical evidence and validity of identifying assumptions. The key identifying assumption underlying the empirical design is that absent a tax-free year, the labor supply of workers in high and low tax brackets would have run parallel. As always, the parallel trends assumption cannot be directly tested, as the counterfactual scenario is never observed. However, the plausibility of this assumption can be evaluated by assessing if outcomes follow parallel trends prior to the reform.

Figures A.5, A.6 and A.7 in the Online Appendix illustrate the research design by visually implementing difference-in-differences by plotting the time series of labor earnings, weeks worked and marginal tax rates for the average individuals in the three top tax-brackets relative to those in the bottom bracket. In order to provide a comparison that corresponds to the regression analysis, where I control for individual characteristics that differ across tax brackets, I non-parametrically weight the group-by-year distributions of the control group to align with that in the treatment group, using the frequently applied reweighting method of DiNardo, Fortin, and Lemieux (1996). The figures provide compelling evidence of differential labor supply responses.

In order to provide a formal test of the parallel trends assumption, I estimate a version of the DD regression (1), where the treatment status is interacted with all time dummies 1982-1988. The results for both labor earnings and weeks worked are presented in Figure 6. The set of pre-reform coefficients tests for parallel trends, with each coefficient corresponding to a placebo test for the given year. The tests indicate no false positives.[16] While there is no significant difference in 1988 in terms of weeks worked, there is a difference for labor earnings. There can be two reasons for this difference. First, the tax-free year may generate an effect of labor supply that extends beyond 1987. Second, labor supply in 1988 and onwards is possibly influenced by changes in the tax system taking place in 1988. As the focus of this paper is to study the short-term effect of a transitory tax-cut, I limit my sample to 1981-1987 with 1987 being the single treatment year. While I comment briefly on responses extending beyond 1987, the analysis of permanent effects is reserved for further research (in ongoing work).

Regression results. Table 1 presents estimates of the elasticity of earned income, defined as the sum of labor earnings in all jobs including self-employment. Each column-by-row entry in the table corresponds to one regression estimate. In Column (1), the regressions control flexibly for individual characteristics, including dummies for gender, age, education, marital status, number of children and location. In the top row, I present estimates of the elasticity of labor income, estimated using equation (2). The elasticity estimate is 0.374 and it is highly statistically significant at the 1% level. The estimate implies that a 10% increase in the after-tax wage causes labor earnings to increase by almost 4% on average. Conceptually, the elasticity estimate consists of two components. First, the

[16] Figure 6 demonstrates how this identification strategy is useful in dealing with possible effects of macroeconomic shocks. As documented in Figure A.9 in the Online Appendix, the Icelandic economy was hit by a negative macroeconomic shock in 1983, related to a resource shock in the fishing sector and a drop in the real exchange rate. As Figure 6 illustrates, this macroeconomic shock did not differentially affect outcomes across the tax-brackets in a statistically significant way.

reduced form, presented in the middle row, which is a DD estimate of equation (1) on log labor income, which is estimated at 0.077. Second, presented in the middle row, the first stage which is a similar DD estimate where the outcome variable is the log net-of-tax rate, estimated at 0.207. The elasticity is essentially the ratio of the reduced form to the first stage, but here it is estimated using 2SLS. In Column (2), I evaluate the sensitivity of this estimate to the inclusion of occupation and sector fixed effects, which results in a slightly lower estimate. This indicates that there are cross-sectional differences in elasticities across types of jobs, a heterogeneity which I will explore in detail in Section 7.

Previous research has highlighted that a DD design can be effectively combined with matching methods to produce a more robust inference (Heckman et al., 1997; Blundell and Dias, 2009). Matching will generate more comparable treatment and control groups and DD will 'difference out' unobserved differences. In order to leverage these benefits, I augment my DD estimation with a non-parametric coarsened exact matching (Iacus, King, and Porro, 2012). More precisely, I first match individuals coarsely on pre-reform characteristics (age, marital status, number of children and education) and then estimate DD on the matched sample, using the weights obtained from matching. Since the set of covariates used in the matching procedure is very general, I am able to match 99.96% of the sample in this way. The results, reported in Column (3) of Table 1, are very similar to the main specification, implying an elasticity of 0.401. The robustness of the main specification to this alternative implies that systematic differences in the characteristics of individuals across the different brackets have limited effects on the estimates.[17]

Table 2 presents estimates of the effect on weeks worked and is organized in the same way as Table 1. The variable collects total weeks worked across all jobs held by the individual. The regression estimates reflect strong responses in weeks worked. The reduced form estimate implies that workers in the top three tax brackets increased their working time by about 1 week more than those in the bottom bracket. A treatment effect of 5 additional weeks relative to a pre-reform average of 48.43 weeks implies an elasticity of about 0.10 (4.926/48.43).[18]

It is important to highlight what these results imply and what to expect. As discussed in Section 3, the working time recorded on the payslips is reported in weeks worked. This reflects time spent working, not duration of employment, with a standard week corresponding to 40 hours. The caveat is that weeks are capped at 52 per job. In total, workers can work less than 52 weeks per year, e.g. if not working all weeks in the year or if part-time and not working 40 hours per week. But they can work more than 52 weeks if they hold more than one job. Therefore, an additional week reflects the exchange of vacation for working time, more full-time employment and work on secondary jobs. However, this measure does not capture overtime and other changes in working time within the week, exceeding 40 hours.

The earnings elasticity of 0.374 incorporates all margins of labor supply leading to increased earnings, including more weeks worked. My estimates imply that 30% of the overall response are brought

[17]Figure A.8 in the Online Appendix investigates where these responses originate in the earnings growth distribution, documenting that responses reflects more and higher earnings increases but also less earnings decreases

[18]As reported in Table A.5 in the Online Appendix, this implied elasticity is similar to one obtained from a specification in logarithms of weeks worked.

about by more weeks worked (e.g. less vacation, more full-time employment, secondary-jobs) and 70% by more earnings within those weeks (e.g. over-time hours, more work effort).

Table 3 documents the estimated effect on employment which, as explained above, is defined as having an income equal to or exceeding an income threshold. I find no significant effect on employment. When interpreting this result, a few features of the research design are important to bear in mind. First, recall that labor supply responses are identified from differential responses of workers in different tax brackets. Hence, by construction, the research design is unable to uncover labor-market entry responses: the sample is restricted to workers that are employed prior to the reform. Second, while the design is able to uncover the potential effect on labor market exit, for reasons such as delayed retirement, the estimates imply that there are no *differential responses* along the extensive margin across tax brackets. However, it is still possible that some workers delayed retirement in response to the temporary incentive created by the tax-free year, relative to what they would not done in a normal year. I revisit this question in Section 5, where I develop a research design that is able to detect extensive margin responses both through entry and exit.

Real Labor Supply Responses, not a Reporting Phenomenon. A critical reader will ask the important question: can the estimated earnings elasticity be interpreted as labor supply elasticity? While it is clear that my finding of an effect on weeks worked stems from more working time, the earnings effect might reflect – in addition to more jobs, weeks, daytime and overtime hours – some form of reporting responses or tax avoidance. I conduct a further analysis along several dimensions to shed light on this question, demonstrating that my findings reflect, at least to a large extent, real labor supply responses.

First, I estimate responses separately for employed workers and for self-employed and business owners, defined as having at least one job as self-employed. Self-employed individuals are likely to have more flexibility in adjusting their labor supply and hence, we might expect to find larger responses for them. However, self-employed workers might also be able to increase their income in the tax-free year through tax avoidance, e.g. by misreporting capital income as labor income or shifting income from other years to the tax-free year. Such avoidance is less likely to be possible for employed workers, as their employers have no direct incentive to collude. Table A.6 in the Online Appendix reports estimates of the elasticity of labor earnings. For wage earners, the elasticity is almost exactly the same as for the whole sample, 0.373, while the elasticity is larger for the self-employed (0.484). However, as documented in Table A.7 in the Online Appendix, there are similar differences in the elasticity of weeks worked between the two groups. Although some of the differences between these groups may be due to reporting behavior, these findings rather imply that they reflect differences in hours flexibility. Such flexibility may be tempting for workers in less flexible jobs. In Table A.10 in the Online Appendix, I investigate whether there was an increased take up of self-employment in the tax-free year, finding a significant entry response.[19]

Second, I study whether my estimates are likely to be explained by income shifting. During the tax-free year, workers may have negotiated with their employer to adjust their compensation in some

[19] The estimated semi-elasticity of self-employment implies that a 10% increase in the after-tax wage increases self-employment by 1 percentage point, relative to an average of 14.9 percent.

way or to front-load some payments. While such behavior is likely to be more difficult and costly to achieve through wages and salaries, e.g. due to payroll taxes, other forms of payments may have been used. To investigate this possibility, I estimate equation (2) separately for each sub-payment on the payslip (in real $ values) and report the effect relative to the total. The results are reported in Table A.8 in the Online Appendix. Overall, the results do not show an unexpected pattern. Increases in wages and salaries make up 94% of the increase in payments and most of the remainder consists of payments such as fringe benefits, and travel allowances, which are likely linked to more work. Potential suspects, such as sales commission and bonuses, as well as gifts, make up only 0.8%. Taken together, this implies that my results are not explained by income shifting through discretionary payments.[20]

Third, I estimate the effect on capital income. While labor and capital income were taxed according to the same tax schedule both pre and post reform, labor income was tax-free in 1987, and capital income was taxed as before. Although it does not provide a pure placebo test, estimating the effect on capital income allows for investigating potential misreporting and tax avoidance. The reporting behavior would manifest itself in a negative effect on capital income, as taxpayers report more of their capital income as labor earnings in the tax-free year. A negative effect on capital income would therefore indicate that at least part of the estimated earnings elasticity is masking reporting behavior. However, we might not expect a zero effect on capital income. As a large part of capital income, such as business income and dividends, is an implicit function of labor supply in the economy, there may be equilibrium effects on capital income resulting from an increased labor supply. Table A.9 in the Online Appendix reports the estimates, documenting a small positive effect on capital income, which is only 2 percent of the treatment effect on labor income. This contradicts the hypothesis of misreporting.

Lastly, there is other, more circumstantial, evidence implying that the Icelandic population was working very hard during the tax-free year. When there is a strong temporary incentive to work, individuals have the incentive to avoid or postpone other activities that take time from working. While a natural example is leisure activity, workers might also be more reluctant to stay at home when they themselves or their family are ill. Figure A.15 documents that workers in Iceland took less sick leave in 1987. The average share of hours on sickness leave of total paid hours was 2.4% both in the years prior to and after 1987, but fell to 1.6% during 1987.[21]

Robustness. In addition to what is described above, I perform further analyses to assess the robustness of my results. First, I evaluate the robustness of my strategy of assigning treatment status based on last year's bracket. While individuals' tax-brackets positions tend to be persistent (Figure A.13), tax brackets correspond to same income quantiles over time (Figure 5b), and analysis of pre-reform years finds no evidence of false positives (Figure 6), a potential bias might arise due temporary mean

[20] These results are consistent with evidence from other Nordic countries, indicating a limited tax avoidance in labor earnings because of third-party reporting by firms (Kleven et al., 2011).

[21] In Figure A.15 I also document that fewer people were receiving sickness benefits in 1987 than in the years before. To the extent that this evidence indicates that workers were working themselves very hard in 1987, it rhymes with a recent study by Ólafsdóttir et al. (2016), which finds an increased likelihood of a heart attack among middle-aged and old men in 1987 and 1988, in particular in the group of self-employed.

reverting income shocks. For example, some people that are in a high tax bracket in the previous year are there because of an income shock that reverts to the mean in the current year, generating a downward bias in the earnings elasticity. I evaluate the validity of this concern in Online Appendix D, performing exercises where I use further lags of the tax-bracket position, as well as a richer set of information, to predict the current tax bracket. This ensures more stable tax-bracket positions over time. The results are very similar to my main specifications, indicating a limited bias due to mean reverting income shocks.

Second, I study the differences in responses across tax brackets and evaluate the robustness of the choice of control group. In the above, I pool together the estimates for the three top tax-brackets in a weighted average elasticity. In addition, the difference-in-differences estimates assume that elasticities are homogeneous across tax brackets. If that assumption is violated, elasticity estimates will be biased. For example, if the elasticity for workers in the top bracket is lower than that for those in a bottom bracket, the elasticity estimate will be biased downwards relative to the true underlying elasticity. Tables A.15 and A.16 in the Online Appendix present estimated effects on earnings and weeks worked separately by tax brackets. The results imply that labor supply elasticities are largest for the lower-middle bracket (0.484) but the smallest for the top bracket (0.236).[22] While the relative earnings response (the reduced form) increases with higher tax brackets, it does so less than the difference in the tax cuts (the first stage), resulting in smaller elasticities. One natural explanation for smaller elasticities in the higher brackets is more frictional adjustment of working time, although this may also reflect differences in preferences. To the extent that these results imply less elastic labor supply of workers in higher than lower tax brackets, this indicates that my main estimates may be downward biased.[23]

While the tax-free year generated a temporary incentive in 1987, its effect may have been more persistent, particularly in the light of the strong responses I document. In addition, the tax system saw several permanent changes in 1988, which themselves may have generated effects on labor supply. In Online Appendix E, I discuss whether and how the permanent reform in 1988 may affect my Frisch elasticity estimates and perform a range of tests to evaluate the robustness of my results, which all broadly support my main results reported above. I argue that the 1988 changes, which were announced in the spring of 1987, were neither as simple nor as salient as the tax-free year. While the progressive tax schedule was replaced with a flat tax rate, substantial changes in tax deductions influenced the tax base at the same time. Judging from media coverage (Figure 4) these changes, such as the introduction of the flat tax, seem to have been much less salient than a *tax-free year*, which caught much attention.

[22]In the Online Appendix I also present results from other robustness tests. In one set of exercises I estimate elasticities for the top- and upper-middle tax brackets, employing the lower-middle bracket as a control group. As documented in Table A.17, this results in elasticity estimates between 0.232 and 0.289, similar to the bracket average in Table A.15. Studying the elasticity of weeks worked, as reported in Table A.18, yields a similar conclusion. I have also explored the sensitivity of my main estimates of earnings responses that are in natural logarithms. Estimating earnings elasticities using an inverse hyperbolic sine transformation of earnings instead of logs, or $\log(\text{labor earnings} + 1)$, gives broadly similar results.

[23]Importantly, estimates using the identification strategy presented in Section 5 do not suffer from this potential bias as the control group does receive a treatment – i.e. experience a tax-free year – during the sample period.

5 Life-Cycle Difference-in-Differences

In this section, I develop a new research design to estimate labor supply responses along both the intensive and extensive margins. I compare people of a certain age to workers of the same age before the tax-free year, exploiting the fact that the tax-free year was an unanticipated event. This design complements the tax-bracket DD design and has the comparative advantage of being able to identify labor supply responses of the whole population along the extensive and intensive margins, including potential labor market entry.

5.1 Research Design

Intuition and Graphical Illustration. Before drawing up the details of the estimation and presenting results, I use a stylized graphical example to describe the intuition behind the *life-cycle difference-in-difference* approach. For this purpose, I borrow the model in MaCurdy (1981) and the associated graphical representation.

Following an example given in MaCurdy (1981), first consider comparing two individuals, A and B, for which Figure 3 draws their life-cycle wage profiles. As the two individuals are similar in all aspects, they have the same life-cycle wage profile at all ages, except at age T when B is treated with an unanticipated wage increase of Δ that lasts a single period. This causes B's labor supply at age T to be higher than that of A by $\varepsilon \times \Delta$, but lower by a small income effect, where ε is the Frisch elasticity.[24] In this experiment, ε can be estimated by relating the difference in working hours of A and B at age T to the wage difference Δ.

In my empirical setting, there exists no comparison such as between A and B. However, since individuals experience the tax-free year at different points over their life-time, my setting allows for an alternative comparison, enabling me to estimate the labor supply elasticities. To illustrate that comparison, Figure 3 plots a wage profile for individual C who is identical to B except that he experiences an unanticipated wage increase associated with a tax-free year when he is one year older, at age T_{+1}. As documented by the figure, at age T, individual C is the counterfactual for B, as they follow the same wage paths. Therefore, the intertemporal elasticity ε can be estimated by relating the wage increase Δ to the difference in labor supply of B and C at age T, when C has not received the wage increase. The life-cycle difference-in-difference strategy builds on this idea.

Matching Procedure. The research design leverages two features. First, from the individual's perspective, at which age she experiences a tax-free year was as good as random. Second, absent a tax-free year, the labor supply of similar individuals is likely to follow similar paths over their life-cycle. Therefore, for a given worker experiencing a tax-free year, workers in other birth cohorts with similar characteristics, when observed at the same age, are likely to constitute a good counterfactual.

A key challenge is to pair workers being treated with a tax-free year to an appropriate comparison

[24] Online Appendix G discusses the MaCurdy (1981) model and the examples therein. As I illustrate there, a one-period wage change as depicted in Figure 3 results in a change in B's life-time labor supply due to an income effect arising from the fact that the rise in wages increases the income in that year even if the labor supply is unchanged. This is different from the tax-free year, where the income remains unchanged at the same labor supply as the year before.

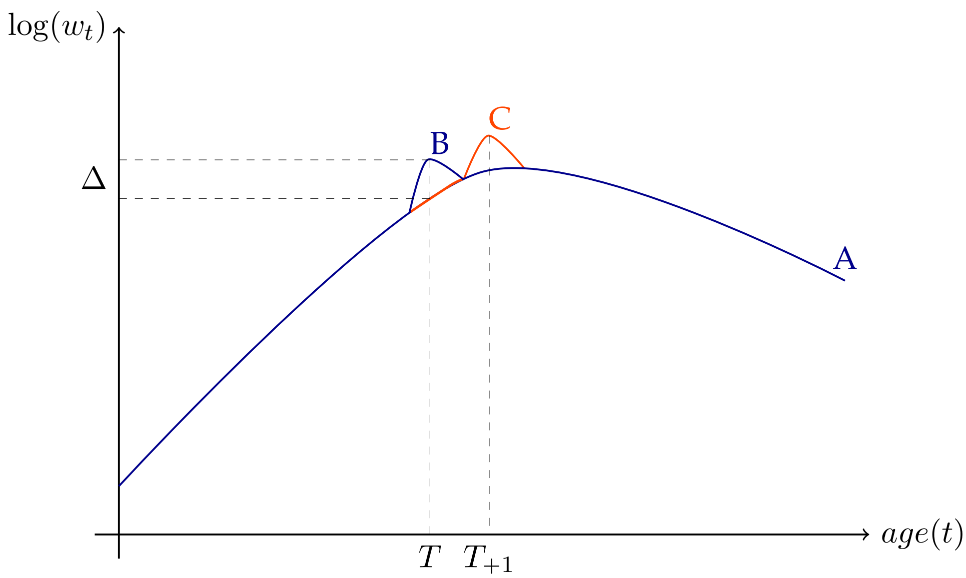


Figure 3: Life-cycle DD research design illustrated using a MaCurdy (1981) Graph

group with a parallel life-cycle labor supply trend, which can be pinned down using difference-in-differences. To this end, I construct a control group by implementing a "Coarsened Exact Matching" (CEM) procedure (Iacus, King, and Porro, 2012), where each birth cohort is paired with individuals of the same age and lagged characteristics in other birth cohorts. The general argument for applying matching in observational studies is to achieve a balance in covariate distributions across treatment and control group, with the aim of replicating a randomized experiment as closely as possible (Rosenbaum, 2002; Rubin, 2006). As opposed to methods relying on estimating a propensity score, CEM is a nonparametric procedure to achieve a sample balance ex ante. Therefore, with reference to the design of randomized experiments, the method enables me to construct "blocks" within which individuals may be expected to follow similar trends in labor supply, but receiving treatment at a given age is plausibly random.

For each birth cohort, the control group is selected from the adjacent birth cohort, born one year earlier. That is, workers at a given age in 1987 are matched to workers at the same age in 1986. I make this restriction in order to achieve three goals. First, this limits the set of workers to be paired to those that are most likely to be comparable in their life-cycle patterns of labor supply as well as other aspects. Second, this allows me to restrict the sample period for both the treatment and control groups to 1987 and earlier, enabling me to exclude later years where labor supply may be influenced by the tax-free year itself or changes in the tax code taking place in 1988, thus escaping possible effects of the reform on the control group. Third, and importantly, the control group within each birth-cohort pair does not experience a treatment until after the end of the sample period.[25] Within adjacent cohort pairs, I further match on a set of characteristics other than age that may correlate with trends in labor supply. I limit the set of characteristics to gender, marital status, number of children,

[25]This setup allows me to circumvent the problems discussed in Borusyak and Jaravel (2018) related to event-study designs where the control group eventually becomes treated within the sample period.

location dummy for living in the capital area, education coarsened into 3 levels (compulsory, junior college and university), and lagged labor income coarsened into deciles. Given the general set of characteristics, I have a broad support and am able to match 99.98% of the sample.[26] Cases where no match is found are dropped and in cases of multiple matches, observations are weighted according to the size of the treatment group.[27]

The matching procedure provides a sample of treatment and control groups that are comparable in factors confounded with trends in labor supply behavior. However, my research design does not impose the assumption that outcomes measuring the labor supply of the comparison groups are at an equal level. Rather, it assumes that they follow parallel life-cycle trends in labor supply. Therefore, causal effects are obtained using difference-in-differences. In this way, I leverage the advantages of matching in generating comparable treatment and control groups and DD to difference out unobserved differences.

Identifying assumptions and graphical example. The primary identifying assumption is that, in the absence of a tax-free year, the labor supply of similar individuals in adjacent cohorts would have followed parallel life-cycle paths. In addition, the research design rests on the assumption that labor supply only deviates from these life-cycle trends in 1987 due to the tax-free year.

Figure 7 provides a graphical example illustrating the research design for a sample of three birth cohorts, born in 1940, 1939 and 1938. Panel (a) plots the marginal tax rates, illustrating the staggering of when the birth cohorts experience the tax-free year over their life-time. In panel (b), I plot the average weeks worked, documenting that these cohorts work on average about 51 weeks in normal years, but increase their working time to about 53 or 54 weeks in the tax-free year. Panel (c) plots the evolution of real labor earnings, normalizing the averages to 100 in 1986. The figure documents similar trends among the three cohorts in the years prior to 1987 but a clear temporary divergence from that trend in the tax-free year. I make that point clearer in panel (d), which plots the difference in earnings for each cohort relative to the cohort born one year earlier. This removes the common trend and illustrates the clear differential change in earnings. The figure lends support to the key identifying assumption underlying my research design, illustrating how adjacent cohorts follow similar life-cycle trends in labor supply.

A potential threat to identification would be if there were shocks contemporaneous to the tax-free year that influence the outcome of the treatment group relative to the control group. An example of such threats would be shocks to labor demand leading to an increased labor input in equilibrium and a reverse causality. A potential scenario would be that some sectors or occupations were hit by shocks in the tax-free year which would influence their labor market outcomes and be captured by my estimates. In general, such a shock would be a cause for concern as the results might be driven by particular subgroups receiving additional treatment. I evaluate the robustness of my result to these concerns below. Moreover, and importantly, no other reforms coincided with the tax-free year, such as changes to social security or taxes on firms. The only change to individuals' taxes and benefits

[26] Due to the "curse of dimensionality", the nonparametric matching procedure delivers fewer matches the larger the set of characteristics matched on. As a robustness check, I have also performed matching with more characteristics, including occupation and sector, arriving at broadly similar results.

[27] As a robustness check, I have performed one-to-one matching, which delivers similar results.

in 1987 was that income taxes were zero. Lastly, I emphasize that while macroeconomic shocks unrelated to the tax-reform would invalidate the life-cycle DD design, it identifies the market-wide impact of a change in wages, including equilibrium effects arising from the reform.

Empirical Framework and Estimating Equation. The sample consists of individuals i belonging to birth cohorts c, where c denotes year of birth. Age is defined as $a = t - c$, where t is "calendar time". I denote the age at which a birth cohort experiences the tax-free year treatment by $A_c = 1987 - c$. As emphasized and illustrated in the above examples, the relevant concept of time in this empirical framework is life-time, i.e. age. In that context, it is useful to refer to *age cohorts* as the group of individuals who are being observed at the same points in their life-time.

As described in detail above, workers at age a from cohort c are matched to workers of the same age a from the adjacent birth cohort $c-1$. Matched cohort-pairs $\{c, c-1\}$, i.e. age cohorts, are denoted by g. Within each age cohort g, I define "event time" as $k = a - A_c$, or age relative to age at the event of treatment. Then, I define the treatment indicator as $D_{gk} = 1$ if $a = A_c$, but zero otherwise. All age cohorts are observed during and prior to the treatment event. Importantly, this implies that the treatment indicator D_{gk} uniquely defines the treatment group (c) and the treatment period within each age cohort, as the control group ($c - 1$) does not experience the treatment until after the end of the study period.

Using this notation, the difference-in-differences are estimated by:

$$y_{ik} = \alpha_{ig} + \delta_k + \eta \cdot D_{gk} + \boldsymbol{X}'_{ik}\gamma + \mu_{ik} \tag{3}$$

where y_{ik} measures the outcome of interest for individual i at event time k, α_{ig} are match-group fixed effects, i.e. fixed effects for each cell (or block) within which individuals are matched, which absorbs the average differences between the treatment and the control group, and δ_k are event-time fixed effects. The vector $\boldsymbol{X}_{ik}$ collects characteristics that we may want to control for, but that are not used in the matching process, such as occupation and sector fixed effects. The error term, μ_{ik}, captures other determinants of labor supply. To address potential concerns regarding serial correlation in outcomes within groups across periods (Bertrand, Duflo, and Mullainathan, 2004), I cluster standard errors μ_{ik} at the match-group level. Estimating equation (3) for the sample of all age cohorts, the difference-in-differences coefficient η gives an estimate of the average treatment effect on labor supply across the population. In order to obtain an estimate of the (semi-) elasticity of labor supply, I estimate the following equation:

$$y_{ik} = \alpha_{ig} + \delta_k + \varepsilon \cdot \log(1 - \tau_{ik}) + \boldsymbol{X}'_{ik}\gamma + \nu_{ik} \tag{4}$$

where the logarithm of the net-of-tax rate $\log(1 - \tau_{ik})$ is instrumented by the treatment indicator D_{gk}. An estimate of the labor supply elasticity is then created by the coefficient ε which, when equation (3) is estimated for the sample of all age cohorts, is the average across the population.

5.2 Regression Results

In Table 4, I report estimates of the labor earnings elasticity. The table is organized in the same way as those in Section 4, where the bottom row reports the first-stage estimates, the middle row reports the reduced form and the top row reports elasticity estimates. The top row of Column (1) reports an elasticity estimate of 0.654, which is highly significant at the 1% level. In Column (2) I include both occupation and sector fixed effects with the aim of absorbing e.g. demand-side shocks contemporaneous to the tax-free year that might affect the estimates. Now estimates are identified from the variation within sectors and occupations. As documented by the table, the estimates are almost identical to those in Column (1) and the estimates are therefore insensitive to these controls.[28] In Column (3) I include individual fixed effects, which also produce a similar estimated elasticity as in the two prior specifications.

When interpreting these estimates and comparing them those presented earlier, it is important to bear in mind that they do not only incorporate intensive margin responses, as those reported in Section 4, but also extensive margin responses. Different from the tax-bracket DD design which, by construction, excluded workers that were out of the labor force or part-time employed, the current method is able to identify responses for the entire working-age population, measuring the average aggregate elasticity.

Table 5 reports the effect on weeks worked. The semi-elasticity of weeks worked, documented in the top row of Column (1), is 3.014 additional weeks. Based on pre-reform averages, this translates into an elasticity of about 0.078 (3/38.37). Controlling for occupation and sector fixed-effects in Column (2) and individual fixed effects in Column (3) results in similarly sized estimates.

As emphasized above, an important motivation underlying the development of the life-cycle DD design is its suitability for studying labor supply responses of workers that are marginally-attached to the labor market and extensive margin responses. Recall that the estimates reported in Section 4 implied a zero differential employment response across tax brackets. Table 6 reports employment semi-elasticity estimates, where the dependent variable is, as before, an indicator for labor earnings exceeding the base income level. Since an individual's decision whether to enter or exit the labor market is likely based on the total financial incentives of working which, in turn, are influenced by the dis-incentives generated by the tax burden he expects to bear if employed, the employment semi-elasticity relates the employment probability to the *avergage* tax rate that individuals face if working rather than the tax paid on the marginal dollar earned.[29] As documented in the top row of Column (1), this produces an employment semi-elasticity of 0.068 which is highly statistically significant. That is, the estimate implies that a 10% decrease in the take-home pay increases the employment rate by about 0.7 percentage points. In Column (2) I include individual fixed effects, finding a similar effect. This result highlights that the tax-free year generated a modest but both economically and statistically significant effect on employment.

[28]Occupation and sector fixed effects are defined based on the previous year's employment and job status. Workers that do not have an occupation and sector, i.e. who are not employed, are allocated to a residual group. While not reported, the results are also robust to including occupation×period and sector×period fixed effects.

[29]The average tax rate is the ratio of the income-tax payments and income-tax base. Employment semi-elasticity estimates relate the employment rate to the net-of-average-tax rate.

Robustness. I have conducted further analyses along several dimensions in order to evaluate the robustness of the results reported above. I document the main findings below.

Iceland being a small open economy, macroeconomic volatility has traditionally been driven by external shocks, such as in exports, or shocks in its natural resources, e.g. biological shocks in fish supply. At the time of the tax reform, the Icelandic economy had been in an upswing where a key driver of the growing economy was a booming fishing sector. Marine exports had been growing strongly following a positive terms-of-trade shock, mainly due to higher fish prices in North-America and other close markets. While on a downward trend throughout the 20th century, fishing and fish processing constituted about 15% of GDP in the 1980s and this sector employed about the same share of workers. Therefore, there might be a concern that my results are influenced by some form of export or fishing-sector shock. To evaluate this claim, I perform my analysis on a sample excluding all workers and firms in the fishing and fish processing sector. Reported in Table A.28 in the Online Appendix, I find my results to be very robust to this exclusion. I have also carried out my analysis separately for workers employed in the tradable and non-tradable sectors separately, estimating similar but slightly larger estimates for the latter.

Another concern that the reader might have is whether my estimates are picking up some differential trends or shocks in labor supply (or demand) around the timing of the tax reform, either in the economy as a whole or for particular cohorts. Naturally, this concern cannot be ruled out or tested directly. However, as a way of evaluating its plausibility, I conduct placebo tests for the years leading up to the reform as 'placebo tax-free years'. More precisely, I first drop 1987 from my sample and then follow the same procedure as described in Section 5.1, estimating equation (3) for each cohort. Reassuringly, as documented in Figure A.16 in the Online Appendix, placebo-year coefficients are scattered around zero and are rarely statistically significant, indicating no systematic patterns or false positives, while the corresponding estimates for the tax-free year are always orders of magnitudes larger.

Triple-differences. Although the life-cycle DD design allows for identifying labor supply elasticities from differences between individuals likely to be on parallel life-cycle trends, I cannot rule out the possibility of aggregate shocks, other than the tax-free year, affecting my estimates. To further evaluate the robustness of the results, I develop a research design that marries the life-cycle DD and the tax-bracket DD in a triple-difference (DDD) design.[30] In this empirical strategy, differences across adjacent birth cohorts are augmented with within-birth-cohort differences across tax brackets. The benefit of this design is that, in addition to comparing similar individuals expected to be on similar life-cycle labor supply paths, it differences out all possible common time effects, reducing the identifying variation to cross-sectional variation only.

Table 7 reports the results for the three main outcomes. The elasticity of earned income is 0.431.[31]

[30] For earlier examples of studies employing triple difference designs, see, e.g., Gruber (1994) and Chetty, Looney, and Kroft (2009).

[31] The estimating equation, which combines equations (3) and (1), is:

$$y_{ik} = \alpha_{ig} + \delta_k + bracket_{i,k-1} + \alpha_{ig} \times bracket_{i,k-1} + \beta_D D_{gk} + \beta_B B_{i,k-1} + \eta \cdot D_{gk} \times B_{i,k-1} + \boldsymbol{X}'_{ik}\gamma + \nu_{ik} \quad (5)$$

The coefficient of interest is η, which identifies the triple-difference. This captures the variation in labor supply specific to

As expected, this is similar to estimates using the tax-bracket DD but lower than estimates using life-cycle DD, which also incorporates equilibrium effects. The semi-elasticity on weeks worked is about 2.5 weeks. The employment semi-elasticity is close to zero and not statistically significant.

6 Discussion of Frisch Elasticity Estimates

6.1 Summary of Elasticity Estimates

Figure 8 summarizes my estimates of Frisch elasticities across the two adjustment margins using the two research designs. Employing the tax-bracket DD design, I estimate an elasticity of 0.374. Based on the structure of the estimation method, and the fact that it identified no effect on employment, I interpret this estimate as an intensive margin elasticity. Using the life-cycle DD design, I estimated an average elasticity of 0.654, which captures both intensive and extensive margin responses. The employment elasticity, measuring extensive margin responses, is 0.068.

Two main factors separate my estimates from the two designs. First, as the tax-bracket DD identifies elasticities from a cross-sectional variation in tax rates, the estimates are restricted to the employed population whereas estimates using the life-cycle DD design also incorporate extensive margin responses. Second, as the life-cycle DD exploits both the time-series and cross-sectional variation, it will incorporate all macroeconomic effects in the tax-free year, including equilibrium effects. Ex ante, it is not clear whether equilibrium effects contribute positively or negatively to the elasticity estimate. On the one hand, if labor demand is not perfectly elastic, strong labor supply responses may lead wages to fall. This would dampen the labor supply response and attenuate the estimated elasticity using the life-cycle DD compared to the tax-bracket DD if these effects are common across tax brackets. On the other hand, workers receiving large tax cuts may spend less time on leisure but also on home production, such as home cleaning, cooking and child care. This may generate demand for labor input in the sectors providing those goods and services, thus facilitating more work for those that desire to work longer hours during a tax holiday, amplifying the overall labor supply response. Moreover, if these workers disproportionately fall into lower brackets, elasticities estimated using the tax-bracket DD will be biased downwards relative to those estimated using the life-cycle DD.

As my administrative data does not include information on hourly wage rates, I cannot directly estimate the effects of the tax-free year on wages. However, using data on wage rates drawn from a survey on paid hourly wages collected by the Icelandic Wage Research Committee I can evaluate whether there appear to be reductions in wage rates in 1987. Figure A.10 in the Online Appendix plots hourly wage rates by occupations through the 1980s and provides little evidence for a reduction in wages during the tax-free year.[32]

the treated birth-cohorts (relative to the control birth-cohorts), for the workers in high-tax brackets (relative to those in low tax brackets), during the tax-free year (relative to the years before). As before, the elasticity of labor supply is identified by estimating a version of the equation above that includes the logarithm of the net-of-tax rate $\log(1-\tau_{ik})$, which is then instrumented with the triple-difference interaction term $D_{gk} \times B_{i,k-1}$. To enhance the comparison between the two research designs, Table 7 reports estimates based on the same sample as in Section 4. However, estimates for the full sample are broadly similar.

[32]The wage increase that is visible in end of 1986 and beginning of 1987 likely results mostly from national-level collective wage negotiations that took place during 1986 and new agreements were signed that year.

In order to arrive at comparable estimates based on the two methods and to get an estimate of the size of these equilibrium and aggregate effects, I first apply the life-cycle DD method to the same sample that was used in the tax-bracket DD estimation, yielding an elasticity estimate of 0.529. Next, I apply the triple-difference method, which gives an elasticity estimate of 0.431, which is very similar to my baseline tax-bracket DD estimate. The difference between the two estimates is 0.098, or about one-fifth of the size of the intensive-margin elasticity. Together with the lack of evidence for reduction in wages, this implies that the tax-free year may have generated positive equilibrium effects through increased activity and demand for various services. However, since this estimate may also include macro shocks not related to the tax reform, I am cautious in this interpretation.

6.2 Comparison to Existing Evidence

Is the size of the Frisch elasticity I estimate reasonable? Due to its centrality in economic theory and for policy evaluation, most work on labor supply focuses on obtaining elasticity estimates. Therefore, results can be compared across studies. However, as I discuss below, reliable estimates of Frisch elasticities are few and the evidence is mixed. An alternative approach to evaluating the size of my estimates is via comparison to estimates of other parameters – such as of Marshallian and Hiscksian labor supply elasticities – for which empirical evidence is more abundant. I pursue this approach in Online Appendix F.

To obtain a point of reference, I conduct a meta-analysis of previous estimates in Figure 9.[33] Figure 9a summarizes estimates of intensive-margin Frisch elasticities. The figure is organized in three sections by the samples studied, from left to right: the population (either as a whole or a representative sample), prime-aged men and specific occupational groups. For reference, I also plot (circled in orange) estimates for the corresponding sample in my study.

Close studies. My paper stands closest to two earlier studies. Bianchi, Gudmundsson, and Zoega (2001) highlight the rare opportunity that the tax-free year offers to study labor supply. Using a random sample of 9,300 individuals, they compare outcomes in the tax-free year to the year before and the year after and report strong responses. However, two reasons make it challenging to compare their estimates to those of others, including mine. First, as the authors document, employment and output had been on an upward trend in the years prior to the reform. This makes it difficult to separate the responses to the tax-free year from pre-trends and the business cycle. In the current paper, I take this concern seriously by using difference-in-difference and matching methods to generate comparable treatment and control groups on common trends prior to the reform and then difference out common trends and unobserved differences. Second, due to limited available data at the time, the estimates in Bianchi, Gudmundsson, and Zoega (2001) are based on *average* tax rates in 1986. In order to enhance the comparability, I compute an elasticity of 0.77 using average changes in earnings reported in their study and my computation of average marginal tax rates in the population.[34] This

[33] I do not attempt to provide an exhaustive survey, but rather to provide an informative comparison. Extensive surveys of the literature on labor supply elasticities include Killingsworth (1983), Pencavel (1986), Blundell and MaCurdy (1999), Meghir and Phillips (2010) and Keane (2011).

[34] More specifically, based on information in Table 6 in Bianchi, Gudmundsson, and Zoega (2001), I compute the elasticity

difference between their elasticity estimate and mine, which is about half its size, is that it does not separate labor supply responses during the tax-free year from employment and business cycle trends, equilibrium effects, as well as the effects of changes to the tax system taking place in 1988.

Martinez, Saez, and Siegenthaler (2018) estimate a Frisch elasticity using a tax reform in Switzerland. Switzerland changed the base for income taxation from the previous *two* years' income to pay-as-you-earn. As a result, the reform led to a two-year tax holiday. Using staggering of the reform across cantons, the authors estimate a small elasticity of 0.05 with a small standard error. In contrast, my intensive margin estimate of 0.37 is almost an order of magnitude larger. This difference, in particular the small response of Swiss workers, is surprising. Given that the tax holidays in Iceland and Switzerland both created clear and strong incentives for workers to temporarily increase their labor supply, why do they generate such different responses? There are at least three candidate explanations. First, it may reflect differences in the degree of adjustment and organizational frictions between the Icelandic and Swiss labor markets. As I highlighted in Section 2.2, the Icelandic labor market is more flexible and less regulated than labor markets on mainland Europe, closer to what is found in the US labor market. As I will document in the coming sections, labor supply responses vary substantially across workers depending on the flexibility of their work arrangements. It is not unlikely that the same is true across countries. Indeed, Martinez, Saez, and Siegenthaler (2018) report a relatively large elasticity of 0.3-0.4 for the subpopulation of self-employed workers, a group which is likely to have more flexibility in labor adjustment than the average worker. Second, the methodology in Martinez, Saez, and Siegenthaler (2018) compares average outcomes in areas experiencing a tax holiday to others undergoing a reform in different years, identifying reduced-form aggregate responses. This includes the direct effect, where workers faced with larger tax cuts have stronger incentives to shift labor supply to years without taxes, as well as all indirect equilibrium effects, negative and positive. While negative local-labor market equilibrium effects cannot be ruled out as one possible explanation, strong negative effects would be inconsistent with my results for Iceland, which indicate that the tax-free year generated positive equilibrium effects. Third, unlike the reform in Iceland, the reform in Switzerland may have been, at least to some extent, anticipated. The federal tax law was passed in 1990, following which the cantons were free to adopt the new system when they wanted. The earliest switching cantons had a tax holiday in 1997 and 1998. If the anticipation effects are strong, workers may choose to work less in the years before the tax holidays, and thus intertemporally substitute labor supply towards the tax-free years. Unfortunately, as reported in Martinez, Saez, and Siegenthaler (2018), the data are missing for the early switching cantons when suitable control groups for such an analysis are available, rendering it infeasible to answer this question empirically. However, one might also expect that the anticipation effects would lead to stronger responses during the tax holiday. Therefore, of these three, differences in labor-market flexibility seem to be the most plausible explanation for the differences in labor supply responses.

as the weighted average percentage change in earnings for men and women in 1987 relative to the average in 1986 and 1988, divided by the change in net-of-tax rates for the same years. The standard error is computed from the standard errors reported for the changes in earnings using the Delta method. I interpret this as an intensive-margin elasticity as the calculations are based on individuals working in 1986. This is the same procedure as that used in Chetty (2012), whereas their calculations are based on averages across the tax-bracket schedule.

In recent work, Stefánsson (2019) revisits and extends the analysis in Bianchi, Gudmundsson, and Zoega (2001). Using difference-in-differences across tax-brackets, he provides earnings elasticity estimates implying a lower Frisch elasticity than found in Bianchi, Gudmundsson, and Zoega (2001). These estimates are complementary to the analysis I present in Section 4.[35] Relying on the details of my new data set, such as information about working time, characteristics of jobs and information about work on different jobs that people hold, the focus of the current paper is on studying the mechanisms behind peoples' responses and their heterogeneity, revealing the anatomy of labor supply responses. Stefánsson (2019) focuses on adjusting elasticity estimates to potential biases due to income effects from the reform and heterogeneity in elasticities. Notably, his results imply a very limited bias due to income effects. This is comforting and in line with the fact that the tax-free year did not create a cash-flow effect as income taxes are due every year. However, his analysis indicates a bias due to heterogeneity across tax brackets. Importantly, this bias can be limited by using alternative research designs, such as the life-cycle DD method that I introduce in Section 5.

Other earlier work. Most of the existing evidence on Frisch elasticity is based on regressions of working hours on wages of prime-age men, predominantly in the Panel Study of Income Dynamics (PSID) in the US. This includes the seminal studies of MaCurdy (1981) and Altonji (1986). As Figure 9a illustrates, the estimates reported in this literature tend to be small and imprecisely estimated and often statistically insignificant from zero. There are several reasons why my estimates might differ from this literature. First, the instrumental variable approach used in much of this literature is based on individual characteristics, traditionally age and education, as predictors of changes in wages.[36] While this literature brought the insight that these factors can be good predictors of the level of wages, later work has found them to perform poorly in predicting changes, leading to weak instruments (Keane, 2011). In addition, these characteristics may also predict individuals' taste for leisure, something that would violate the exclusion restriction. Second, there has been much emphasis on issues of measurement in survey data on wages and hours, the PSID data set in particular, which may lead to either a positive or a negative bias (see, e.g., Heckman, 1993; French, 2004a; Barrett and Hamermesh, 2017).[37]

Small and insignificant earlier estimates may reflect both the empirical challenge of estimating Frisch elasticity and adjustment frictions and inattention to small changes in pay. This has motivated several studies that study particular occupations, such as bicycle messengers and taxi drivers, for whom finding exogenous changes in wages is plausible and who are more free in choosing their daily labor supply. Summarized in Figure 9a, a finding that clearly emerges from this literature is that these groups are particularly elastic. For reference, the figure also presents my elasticity estimate for the subsample of taxi and transportation drivers, which is larger than my population estimate.

[35] In most cases, the set of tax-bracket comparison groups differs from my main analysis, which pools together the estimates for the three top tax-brackets in a weighted average elasticity. However, when comparable, as in Table A.17 in the Online Appendix, the earnings elasticity estimates presented in the two papers are broadly similar.

[36] Pistaferri (2003) differs from others in this literature. In a novel approach, he estimates a life-cycle model using data on people's subjective beliefs about earnings to isolate unexpected variation in wages from expected variation.

[37] In addition to the quasi-experimental literature surveyed here, an extensive literature estimates Frisch elasticity using structural methods. In Figure A.14 in the Online Appendix, I survey estimates from prominent papers in this literature. A general pattern that emerges is that my estimates are closer to the parameter values reported in this literature than to the estimates I have surveyed above.

While these studies provide clear causal estimates pertaining to particular subgroups in an environment with minimum frictions, it is not clear how informative they are for learning about business cycle variation in employment for the average worker and the economy as a whole. For example, while Fehr and Goette (2007) find a very elastic behavior of Swiss bike messengers, the evidence in Martinez, Saez, and Siegenthaler (2018) indicates that their findings are not representative of the Swiss population.

Figure 9b reports estimates of extensive margin Frisch elasticity.[38] Apart from Martinez, Saez, and Siegenthaler (2018), who estimate a Frisch elasticity of zero during the Swiss tax holidays, and Carrington (1996), studying employment in Alaska during an oil pipeline boom in the 1970s and documenting a very large employment response, the existing evidence is concentrated within the population of workers close to retirement. Consistent with this literature, I document extensive-margin responses of older workers. However, the strongest employment responses in my sample are among the youngest cohorts, for which no previous estimates exist.

7 Anatomy of Labor Supply Responses

I have uncovered strong labor supply responses to the temporary work incentive generated by the tax-free year. If there is significant heterogeneity in behavior at the micro level, estimates of the average labor supply elasticity do not identify a structural parameter. Aggregation issues and nonlinearities imply that the aggregate elasticity will depend on the demographic structure of the economy (Attanasio et al., 2018). This implies that understanding business cycle fluctuations in employment and improving public policies requires knowing how labor supply responses are influenced by individuals' characteristics and constraints, and how those factors shape the margins of response. However, due to the lack of large-scale natural experiments and detailed microdata, previous work has been unable to study the macroeconomic relevance of heterogeneity.

One may approach this analysis from two vantage points. Based on theory of labor supply, where the Frisch elasticity measures the intertemporal exchange of time on and away from market work, one approach is to study how increased working time comes at the expense of less leisure, less home production, and less human-capital production. This would be a *functional approach*. An alternative is to study how institutions, such as the family and the workplace, shape the ways in which individuals arrange and adjust their working time. This approach, which I follow, can be described as an *institutional approach*.

7.1 What Features Shape Labor Supply Responses?

In studying heterogeneous responses, how can we systematically direct our attention in the most productive directions? There are multiple margins along which heterogeneity may arise and a vast literature provides a range of standing theories. My approach is to use machine-learning methods as a way of drawing up a roadmap, uncovering signposts that give the most important directions, and

[38]For a recent meta-analysis of extensive margin elasticities, see Chetty et al. (2013).

then arrange my analysis around those landmarks.

My methodology involves four steps. First, I estimate labor supply elasticity at the individual level using life-cycle DD, matching each individual to a counterfactual constructed from a group of individuals with the exact same characteristics. Next, I use the random forest algorithm, developed by Breiman (2001), to predict labor supply elasticity using a broad set of characteristics.[39] Third, I exploit the comparative advantage of the random forest algorithm relative to other machine-learning methods, allowing me to rank characteristics by their importance. Fourth, I proceed by arranging my analysis around the set of most important predictors.

Figure 10 plots the relative importance of characteristics in predicting labor supply elasticity, measured with the gain achieved by splitting along the dimensions of a given characteristic. The characteristics used in the random forest prediction can be broadly categorized into two groups: characteristics of the individual, such as gender and age, and characteristics of his job and employment arrangement, such as occupation and working time. The figure presents results from three models. First, presented in the first bar, a model only based on individual characteristics highlights age to be an important feature, followed by whether and how many children individuals have. In the second bar, I present a model based only on employment and job characteristics, which are all defined in pre-reform values. It highlights the importance of *weeks*, which bundles the importance of weeks worked in three pre-reform years, as well as labor earnings and net wealth. The third bar plots results from a full model, incorporating both individual and employment characteristics, as well as characteristics of spouses of married individuals. It documents that working time in the years prior to the reform is the single most important feature.[40] This is followed by earnings, age, wealth, measures of spousal labor-market activity, and then characteristics of jobs, such as sector, occupation and firm size. Following the pattern it reveals, my analysis will evolve around and illustrate the importance of three themes. First, *labor-market attachment*, which is highlighted by the importance of working time, age, earnings, and wealth. Second, the importance of *flexibility of employment arrangement* is highlighted by the weight of weeks worked and characteristics of jobs. Third, *family ties and coordination* is highlighted by the importance of spouses' labor market activity and children. I now study these themes in turn.

7.2 Labor Market Attachment

Figure 11a plots the elasticity of earned income by age. As explained in Section 5, the life-cycle DD design builds on pairwise cohort-by-cohort differences and therefore naturally produces separate elasticity estimates by cohort. Figure 11a plots separate estimates for each cohort by age in 1987.

The figure displays a very interesting pattern. For the prime-age population, the elasticity is

[39] Athey and Imbens (2016) and Wager and Athey (2017) develop a methodology that uses random forests to estimate heterogeneous treatment effects. This methodology relies on random assignment and can therefore be readily applied to RCTs. Differently, my research design builds on difference-in-differences, implying that their method cannot be readily applied. Therefore, I use my research design to first obtain causal effects at the individual level and then use the random forest algorithm to characterize the heterogeneity in the effects. I then proceed to a more thorough analysis guided by the patterns revealed.

[40] Figure A.18 in the Online Appendix plots a decision tree with the most important splits. It documents that the single most important split is whether working more or less than 25 weeks in 1985.

stable and between 0.4 and 0.5. Older cohorts display slightly stronger responses, in particular those at or around the statutory retirement age of 67.[41] The young cohorts – between the age of 18 and 30 – display the largest elasticities, as high as 2 among the youngest cohorts. Although the elasticity is largest only for the few youngest cohorts, this has an important implication for the aggregate elasticity. The population aged 18-30 corresponds to about 22% of the population, which pulls up the average elasticity depicted with a solid horizontal line. In a similar fashion, Figure 11b explores the heterogeneity in extensive-margin responses, plotting the employment semi-elasticity by age. This figure highlights that the modest aggregate employment responses reported in Section 5 mask an important heterogeneity. All employment responses are driven by workers younger than 25 and older than 60, with the former group displaying a very elastic employment behavior. For the prime-age population, the employment elasticity is zero.

Strong labor supply responses of the youngest and oldest cohorts highlight the importance of labor-market attachment. In particular, the young cohorts, most of whom are out of the labor force, part-time employed or at early stages in their careers, are likely to have more of their time endowment available to be exchanged for more working time at the expense of leisure and other activities. Indeed, the evidence presented in Figure 10 demonstrates the interplay between age and working time, highlighting that age is less important on its own when interacted with weeks worked pre reform.[42] In addition, the fact that elasticities are larger for cohorts close to retirement than the prime-age population rhymes with anecdotal evidence of some pensioners having postponed retirement in 1987 to earn tax-free income.[43]

7.3 Flexibility of Employment Arrangement

The canonical model of labor supply assumes that workers hold a single job on which they can flexibly choose their hours of work.[44] As a result, workers choose to work the number of hours that maximizes their utility at the given wage. Since hours can be varied freely, workers are always on their labor supply curve and preferences determine the hours response to wage changes. A growing literature casts doubt on this assumption, proposing that workers face various adjustment frictions, such as inflexibility in work schedules and inability to get remunerated for additional hours. As a result, estimates of short-run labor supply elasticities will be muted as actual hours cannot be easily adjusted to a new desired level in the event of a wage change. In what follows, I study how such adjustment frictions influence the heterogeneity in labor supply responses.

Temporal Flexibility. Jobs appear to vary greatly in the temporal flexibility they offer. Recently,

[41] While workers receive pension and are eligible for old-age benefits from age 67, it is common to retire later and some choose to retire earlier, e.g. at the time when their spouse reaches the statutory retirement age.

[42] As documentd in Figure A.18, plotting a 'tree' with the most important splits in my random forest prediction, being younger or older than 29 is the second most important split, following the split between working more or less than 25 weeks in 1985. This age cut-off is well supported by Figure 11a.

[43] At the statutory retirement age of 67, individuals become eligible to receive old-age benefits, conditional on having earnings below a certain threshold. According to records of the Social Insurance Administration, benefits were reduced for a significant share of individuals due to receive benefits in 1988, as their 1987 income exceeded the income threshold.

[44] An equivalent interpretation is that workers freely choose between employers offering different hours and wage packages.

Goldin and Katz (2016) exemplify pharmacists as an occupation with high temporal flexibility. In particular, a characteristic feature of this occupation are simple transitions between part-time and full-time employment due to high substitutability between workers.[45] As a result, there is a substantial dispersion in working time within the occupation. Another example are occupations where taking on additional shifts is relatively easy, such as for Uber drivers (Hall and Krueger, 2018), leading to a large dispersion in working time. Building on the idea that underlies these examples, I construct a measure of temporal flexibility based on the dispersion in working time within occupations. More precisely, I measure temporal flexibility with the coefficient of variation (CV) in working time within occupations:

$$CV(W_{ot}) = \frac{\sigma_{ot}}{\mu_{ot}}, \quad \sigma_{ot} = \left[\frac{1}{N_{ot}-1}\sum_{i=1}^{N_{ot}}(W_{iot}-\mu_{ot})^2\right]^{\frac{1}{2}}, \quad \mu_{ot} = \frac{1}{N_{ot}}\sum_{i=1}^{N_{ot}} W_{iot} \tag{6}$$

where W_{iot} is the number of weeks worked by individual i in occupation o in year t, N_{ot} is the number of jobs in occupation o in year t, and μ_{ot}, σ_{ot} are, respectively, the average and standard deviation of weeks worked in occupation o in year t. I calculate $CV(W_{ot})$ for three years prior to the tax-free year and use the average in my analysis.[46].

How should this metric be interpreted? If there is much dispersion in working time within an occupation, e.g. many workers have part-time jobs while others work full-time, it is characterized by high temporal flexibility. However, if the dispersion is low, e.g. if the occupation only allows for full-time employment, the occupation has a low temporal flexibility. In other words, occupations with a higher temporal flexibility are those that offer a broader menu in terms of employment arrangement. According to this measure, occupations with the most temporal flexibility are elementary workers in the service sector (e.g. restaurant workers), workers in cleaning and related activities, and elementary workers in agriculture. The least flexible occupations are managers in construction and manufacturing, and blue-collar workers employed by the U.S. Navy.[47]

Figure 12a plots the earnings elasticity by occupation against the temporal flexibility of the occupation. Employing the life-cycle DD design, I estimate the occupation-level elasticity after conducting the matching procedure described in Section 5.1 within the set of workers employed in each occupation pre-reform. This enables me to compare elasticities across occupations without the difference being driven by compositional differences in characteristics such as age, gender and family characteristics, which I establish to be an importance source of heterogeneity. The size of the dots on the graph is proportional to the number of workers employed in the occupation. The figure documents an upward slope: workers in the most flexible occupations have substantially higher elasticities than those in the most rigid occupations. This difference is further tested in Table 8, reporting that while the earnings elasticity is larger among workers in more temporally flexible jobs, workers in more rigid jobs still display sizable labor supply responses.

[45] Recent research suggests that workers in greater need of flexibility, such as women with young children, put more value on flexible jobs and choose them more actively (Mas and Pallais, 2017).

[46] Figure A.17 in the Online Appendix plots the distribution of $CV(W_{ot})$

[47] These are workers employed at the U.S. Naval Air Station Keflavik (NASKEF). The army base was built during World War II by the United States Army and closed in 2006.

Hours Constraints. Although the notion of hours constraints may be clear from a theoretical standpoint, empirical measures, or even indicators, of the prevalence of such frictions are hard to come by.[48] Ideally, one would like to observe individuals' employment contracts and compare responses of workers who have latitude in deciding how many hours they work to workers locked into a wage-hours bundle and working more hours on the given job is not feasible. As such information is not available, I proxy for hours constraints by whether a certain worker holds a job with a fixed monthly salary or is paid by the hour.

An institutional feature of the Icelandic labor market is that some workers have a fixed-salary contract, receiving the same salary irrespective of the number of hours (including overtime) they work. In most cases, these contracts specify the number of hours the employee is expected to work each month at the minimum. Although fixed-salary contracts exist throughout the labor market, they are more prevalent in some sectors and occupations than others.[49] To identify groups where such contracts are more prevalent, I employ an employer-employee data set with comprehensive information on wages and working hours.[50] As these data only extend back to 1998, they cannot be directly merged into my main data set at the level of individuals or firms. Therefore, I measure the average share of fixed-salary workers by occupation over a ten-year period and assign these shares to workers in my main data based on their occupation. Since these data exclude the public sector and some industries in the private sector, this measure cannot be computed for all occupations. I measure "flexibility of remuneration structure", or "hours flexibility" for short, with 1 minus the fixed-salary share. Occupations with the least hours flexibility according to this measure are professionals (e.g. engineers) in the construction sector (42%) and managers in the construction sector (37%), while those with most flexibility are elementary workers in construction (0.05%) and manufacturing (0.2%).

Figure 12b plots the earnings elasticity by occupation against hours flexibility of occupation. As in 12a, elasticities are computed at the occupation level and the size of the dots is relative to the number of workers in each occupation. The figure depicts a positive and statistically significant relation, implying that workers in occupations with more hours flexibility have larger elasticities. Interestingly, however, the figure also documents a sizable elasticity for workers in many of the occupations for whom hours constraints are likely to be binding. This difference is further tested in Table 8, which contrasts workers in high vs. low hours-flexibility occupations. The table documents that while the earnings elasticity is higher in the flexible occupations, the elasticity of weeks worked is higher in those with less flexible hours. This contrast highlights a difference in the nature of responses across these groups. Workers in occupations that are paid by the hour and have more latitude in adjusting their hours may respond along margins which are to a lesser extent captured by my measure of working time (full-time weeks), such as overtime, whereas those with fixed hours-wage contracts are

[48] Previous studies have indirectly inferred that hours constraints seem to be important from the fact that hours are found to be more flexible for job changers than stayers (Altonji and Paxson, 1988, 1992; Martinez-Granado, 2005). Studies using survey measures on the ability to adjust hours (Biddle, 1988; Ball, 1990) and measures of desired hours of work (Kahn and Lang, 1991; Dickens and Lundberg, 1993; Stewart and Swaffield, 1997; Euwals, 2001) have also found supporting evidence for the existence of hours constraints.

[49] See e.g. various issues of Statistics Iceland's *Statistical Series*, "Earnings in the private sector by occupational group".

[50] In the data, I observe wage rates, regular working hours, overtime etc. Working hours for salaried workers are contractual hours but actual hours worked for the hourly paid. For details, see Sigurdsson and Sigurdardottir (2016).

more likely to convert vacation into working time and take a secondary job, a topic to which I will return.

As this measure can only be computed for a subset of my sample, I construct another measure based on actual pre-reform working time. I define workers to be hours constrained in their primary job if they are recorded to be working exactly 52 weeks in that job in the prior year. This measure is likely to capture similar features as the measure defined above. Indeed, the cross-sectional correlation between the two measures is high (0.7). Documented in Table 8, workers that are constrained according to this measure are less responsive than those with more room for adjustment, in line with the results above.[51]

7.4 Overcoming Adjustment Frictions: Multiple Jobs and Job Changes

The previous section documented important heterogeneities in labor supply responses by temporal flexibility and flexibility in the remuneration structure. Interestingly, however, I find significant responses even for those workers most likely to be constrained. How are they able to overcome frictions?

Secondary jobs and primary-job changes. While hours may be rigid within jobs, they may be flexible across jobs. As a result, constrained workers may choose to change jobs to adjust their labor supply to a new desired level. Although job changes may be an operating margin for long-term adjustment, it is likely to be too costly a margin for temporary adjustment. Alternatively, therefore, workers may choose to take up secondary jobs – i.e. to *moonlight* – as a way of overcoming hours constraints.[52]

I exploit an unusual detail of my data, where I separately observe all jobs that workers hold, to study multiple-job holding (*moonlighting*) and primary-job changes as possible margins of adjustment. In Figure 13, I report estimates of the effect of the tax-free year on secondary-job holding and primary-job change. Primary jobs are defined as the job where the worker earns the highest income. Since my data include unique identifiers of firms, I can track each job over time and define primary-

[51] In addition to the analysis reported above, I have studied the heterogeneity in responses across firms, sectors and occupations by other factors that may influence workers' ability to adjust hours of work. Three results are worth mentioning. First, I find substantially larger elasticities among workers in smaller firms than large firms. Ex ante, it is unclear how firm size influences workers' ability to adjust their hours. If jobs require cooperation between workers, coordination of hours becomes a key constraint. If workers are complements, more work may not be feasible unless complemented by more input from coworkers. Alternatively, if workers are substitutes, all workers in the firm cannot adjust their hours at the same time. Within-firm coordination in hours may also attenuate labor supply responses (Labanca and Pozzoli, 2018; Battisti et al., 2016). My results lend support to the hypothesis of reorganization frictions being less severe in smaller than larger firms. Second, I estimate responses by sectors with different levels of capital intensity. One might expect labor supply adjustment to be more difficult for workers employed in capital intensive sectors and sectors where capital-labor complementarity is high. I find evidence supporting this hypothesis, although the differences are not very pronounced. Third, I study differences in responses across occupation by their *"Routine Task Intensity"* (RTI), a measure developed by Autor et al. (2003) and Autor et al. (2006), which I merge to my data using the mapping to International Standard Occupation Classification (ISCO) in Goos et al. (2014). It is not clear ex ante how job routineness correlates with labor supply responses. Workers in less routine jobs might have more flexibility choosing their working time, but accumulating more time on work, such as by working more shifts, might be easier in routine jobs. I find that workers in high RTI jobs have larger elasticities than those in low RTI jobs, although differences are much smaller than by my measures of flexibility.

[52] For early literature on moonlighting, see, e.g. Perlman (1966); Shishko and Rostker (1976); O'Connell (1979); Krishnan (1990) and Paxson and Sicherman (1996). Interestingly, and related to my results in Section 7.5, Krishnan (1990) finds that husbands' decisions to moonlight are a substitute to wives' earnings.

job change as an event where either the worker leaves his primary job to take up another job or if a previous secondary-job becomes his primary job. Secondary-job holding is defined as working at least one week in a job other than the primary job.

Panel 13a reports a semi-elasticity of secondary-job holding of 0.052. This result, when compared to the average propensity of 0.297, implies strong responses along this margin. In Panel 13b, I present the effect on primary-job change, reporting a negative semi-elasticity of 0.048. This result is both interesting and intuitive. As the tax-free year only generated a temporary incentive, most workers are unlikely to make costly decisions such as changing primary jobs. Moreover, if searching for and taking up new jobs is costly in terms of forgone working time, workers are likely to temporarily postpone otherwise planned job changes.

A simple model with hours constraints in a primary job would imply that multiple-job holding and job changes are more operative margins for constrained workers.[53] To test that prediction, I separate the estimates by whether workers are constrained in their primary job, which as before is measured as working 52 weeks in the primary job in the previous year. As documented by the figure, I find that the effect on secondary-job holding is entirely driven by constrained workers. Similarly, the figure reports a decreased propensity of primary job change among constrained workers, whereas, if anything, it increased among those that were unconstrained.[54]

Decomposition. In order to evaluate the aggregate implications of these margins, I evaluate how much weight secondary jobs and job changes carry in explaining the overall labor supply response. To answer this question, I decompose the total labor supply effect into the contributions from continuing primary jobs, new primary jobs and secondary jobs. Total labor supply, E_T, measured either at the level of real labor earnings or weeks worked, can be written in terms of its subcomponents as

$$\begin{aligned} E_T &= E_p + E_s \\ E_T &= E_p^{\text{Cont}} + \gamma \cdot (E_p^{\text{New}} - E_p^{\text{Cont}}) + E_s \end{aligned} \tag{7}$$

where E_p^{Cont} is a continuing primary job, γ is the propensity of primary-job change and E_s are secondary jobs. The total effect of the tax reform ($d\tau$) can then be decomposed as follows

$$dE_T = \underbrace{dE_p^{Cont}}_{\text{Continuing primary job}} + \underbrace{\gamma \cdot (dE_p^{\text{New}} - dE_p^{\text{Cont}}) + d\gamma \cdot (E_p^{\text{New}} - E_p^{\text{Cont}})}_{\text{Primary job change}} + \underbrace{dE_s}_{\text{Secondary jobs}} \tag{8}$$

The components of equation (8) can be estimated using the DD framework in equation (2). Figure 13c reports the results from the decomposition. I find that 93% of the total earnings effect stem from increased earnings on continuing primary jobs and earnings on secondary jobs amount to 6.8% of the total, arising from both new secondary-jobs and more work on existing jobs. Of the additional weeks worked, 34% of the responses are created by more time on secondary jobs while the remainder arises

[53]In a recent study, Tazhitdinova (2017) exploits a German tax reform to show that the take-up of secondary jobs is likely to be driven by hours constraints.

[54]I have also explored the heterogeneous responses of workers along these margins by flexibility of jobs, finding similar results.

from increased working time in continuing primary jobs. Primary-job changes account for only 0.2% of the effect on labor earnings and contribute negatively to the change in weeks worked, which is consistent with a search cost in terms of foregone working time. As highlighted by equation (8), the contribution from job changes is a result of two opposing forces. First, as documented above, I find a decreased propensity of job change during the tax-free year. Second, those workers that do change jobs, however, increase their labor supply, possibly being able to overcome hours constraints in the previous job. As the decomposition highlights, these two effects almost exactly cancel each other.

7.5 Family Ties and Coordination

Changes in the take home pay, whether experienced by one or more members of a family, are likely to result in coordinated family responses. Interdependencies in spousal labor supply may run through at least three channels. First, as we expect couples to enjoy spending time together, they will coordinate their working time. That is, there is a complementarity in their leisure-time allocation, implying that a change in working time of one spouse will induce a same-sign response of the other. Second, in the spirit of Becker (1965), husbands and wives may engage in a shared effort of home production. As a result, if spouses are substitutes in home production, an increase in the take-home pay of one spouse, with a consequent increase in labor supply, will reduce the hours worked by the other.[55] Third, in addition to these indirect effects, there may be a direct income effect if the spouse's earnings are allocated to consumption which is public in the household.

I present a stylized model of collective labor supply in Online Appendix H, arriving at two predictions. First, within couples engaging in home production, individuals' own-elasticity of labor supply is stronger the more specialized they are in home production and the more important their labor input is in the process. This is because time allocated to home production is a closer substitute to time in market work than leisure. Second, cross-elasticities of labor supply are stronger the more time the individual spends on home production but falling in his input elasticity. In other words, in households where both spouses take part in home production but wives play the leading role, their own-wage elasticity will be larger than that of their husbands. As chores are likely to be influenced by the presence and number of children, with child care being a primary example of home production, mothers of more children are likely to have larger elasticities than women with fewer or no children. However, the cross-elasticity may be stronger (more negative) for married men than for married women if relatively more time-input is needed from them to substitute for their wives' time. I study how these mechanisms affected labor supply responses to the tax-free year.

Own-elasticities by family status. Figure 14 plots the elasticity of earned income by marital status and number of children. These are obtained using the life-cycle DD, which is important as patterns are likely to be shaped by responses of individuals working part-time or out of the labor force, such as married women with children. In the left-hand panel, Figure 14 documents that single men and

[55] Several studies have argued that home production influences the labor supply over the life-cycle (Rupert, Rogerson, and Wright, 1995, 2000) and over the business cycle (Benhabib, Rogerson, and Wright, 1991), implying that it may be an important factor in explaining the macro-micro discrepancy on the size of the Frisch elasticity. However, empirical evidence on spousal interdependencies in intertemporal labor supply remains very scarce.

women are very intertemporally elastic. Moreover, and interestingly, I find no statistical gender difference between single men and women, neither parents nor childless singles.

In the right-hand panel, Figure 14 plots elasticities for married and cohabiting individuals. By contrast, this figure shows a clear pattern. First, married men have a consistently lower elasticity than both their single counterparts and married women. Second, the elasticity among married women is steeply rising in the number of children, with mothers of four children or more having the largest elasticity estimate. Both the fact that married women have larger elasticities than their husbands and that elasticities are larger for mothers of more children are consistent with the predictions of my stylized model. If wives contribute a larger share of their time to home production, which is more time consuming in the presence of (more) children, we expect a larger elasticity for them than for their husbands.

Cross-elasticities of Married Couples How do spouses coordinate their labor supply responses? In order to answer this question, I estimate cross-elasticities for married men and women. That is, I estimate how individuals respond to changes in their spouses' marginal tax rate and, by extension, their labor supply. As income taxes in Iceland are collected at the individual level, an individual's marginal tax rate depends on his own earned income, but not that of his spouse.[56] This implies that the tax-free year generated different changes in the tax rates of husbands and wives. Since these differences vary across households, cross-elasticities can be identified. I estimate cross-elasticities using the following modification of equation (2):

$$y_{it} = bracket_{i,t-1} + \delta_t + \varepsilon^{own} \cdot \log(1 - \tau_{it}) + bracket^{spouse}_{i,t-1} + \varepsilon^{cross} \cdot \log(1 - \tau^{spouse}_{it}) + \boldsymbol{X}'_{it}\gamma + \nu_{it} \quad (9)$$

where the two endogenous variables, the individual's and the spouse's net-of-tax rate, are instrumented with an interaction between indicators of the tax-free year and the treatment status for the individual and his spouse separately. The coefficient ε^{cross} identifies the cross-elasticity.

Figure 14 plots cross-elasticities for married men and women by the number of children. For men, the cross-elasticity is negative and on average larger (in absolute value) for fathers than for childless husbands. In clear contrast, women's cross-elasticities are close to, and not statistically different from, zero. I study this pattern further in Table 9. In Column (1), the upper panel reports a cross-elasticity of -0.172 for husbands, while the lower panel reports a small, positive but insignificant cross-elasticity for wives. The cross-elasticities are identified under the exclusion restriction that the spouse's tax rate only affects individuals' labor supply via their spouses' labor supply. The estimates may, however, be influenced by income effects resulting from increased household income due to the spouse's response. I assess this in Column (2) by including spouse's income as an additional regressor, finding a negative coefficient for men, indicating a small income effect from spousal labor supply, but a positive for women.[57]

In Columns (3)-(8), I evaluate how the coordination revealed by the cross-elasticities depends on

[56]This is different from the tax system in place in many countries, including the US, where married couples are taxed on their joint income and couples therefore face the same marginal tax rate.

[57]I use the inverse hyperbolic sine function of the spouse's income, instead of the logarithm, in order to account for the possibility of non-earning spouses.

indicators of home-production activity and labor-market frictions. First, I find the cross-elasticity for fathers with young children (0-6 year olds) to be almost twice as large as the estimate for other men. Second, the cross elasticity is large and negative for men younger than 60, while not significant for older men.[58] Third, the cross-elasticity for husbands constrained in their primary job is more than twice as large as for those unconstrained, for whom it is only marginally significant.[59]

What do these results imply about overall household responses? To gauge this, I compare total household responses including and excluding cross-responses and spousal income effects. More precisely, I first estimate equation (9) in levels of income for both spouses separately and estimate the increase in total household income accounting for both own responses and effects from their spouse's responses. Comparing this increase in total household response to those assuming no cross-responses or income effects implies that actual household responses are 23% smaller than if spouses would have been treated in isolation.

The unitary model of household labor supply, which models spouses as a single decision-making unit, makes strong predictions about cross-elasticities (Becker, 1973, 1976). More precisely, it predicts that the Slutsky matrix should be symmetric: the cross-elasticities for husbands and wives should be equal (Chiappori and Mazzocco, 2017). This prediction in rejected in my setting. Furthermore, my findings indicate that coordinated responses arise from substitution in tasks and chores within the household. Married women with children respond strongly to a temporary tax-cut. As non-working time is, at least partly, spent on home production, increased market work must be met either by increased market-produced consumption or through an increased input from the spouse. My results show that while men work less in responses to their wives' incentive to work more, the reverse is not true.[60]

An extensive literature has studied gender differences in labor supply, and frequently finds larger elasticities for women than for men (McClelland and Mok, 2012; Blundell and MaCurdy, 1999). The evidence presented above sheds an interesting light on the gender differences in Frisch elasticities and the underlying mechanisms. The results indicate that these differences are not inherent to gender differences per se, as displayed by equal elasticities for single men and women, but rather to the presence of children and specialization within the household. Since the time allocated to home production is a closer substitute to market work than pure leisure, spouses spending relatively more time on home duties, who traditionally are women, will respond relatively more strongly to changes in the take-home pay.

[58]These results confirm those of Aguiar et al. (2013) for the US, finding evidence of a strong substitutability between market work and home production over the business cycle, and more strongly for married workers than singles.

[59]Large responses of mothers may result from adjustment frictions that make it harder for primary earners to adjust relative to secondary earners, who may work less market-hours but more hours at home. I study this in Figure A.19 in the Online Appendix, where I estimate both own and cross-elasticities separately for men and women depending on a combination of their own and their spouses' full-time vs. part-time employment status.

[60]In general, whether spouses' labor supply are complements or substitutes remains an open question. Studies on the 'added worker' effect have found evidence of substitutability in spousal labor supply in response to job loss (Lundberg, 1985; Cullen and Gruber, 2000; Stephens, 2002) and non-recipiency of disability benefits (Autor et al., 2017). Other studies have found evidence of complementarity in retirement decisions (Blau, 1998; Gustman and Steinmeier, 2000) and in responses to permanent tax reforms (Gelber, 2014; Goux et al., 2014). Recent structural work finds presence of children to be important in shaping cross-responses of spouses (Blundell et al., 2018).

8 Conclusion

Understanding how labor supply responds to changes in incentives has been a long-standing research program in micro- and macroeconomics. Exploiting a tax-free year in Iceland as a natural experiment, I find that people respond strongly to a temporary but large and salient change in pay. Using detailed microdata, I study the key mechanisms behind these responses, finding that both labor-market structure and family structure are important determinants of aggregate employment responses to temporary shocks. My results strongly indicate that labor supply responses cannot be boiled down to a single number and average elasticities cannot be interpreted as estimates of a deep structural parameter. But my results also indicate that voluntary changes in work constitute a key mechanism in the transmission of aggregate shocks. Hence, understanding business cycles and improving policies requires us to know which individuals are most responsive to changes in pay and how they respond.

Two important questions have been omitted from the paper. First, large temporary shocks may have permanent effects. For example, decisions to increase the working time or enter the labor market in response to a temporary incentive may be 'sticky' and generate permanent effects. Second, strong labor supply responses resulted in a predictable and large increase in income. Studying how increased labor supply affected consumption and savings may provide valuable information for understanding business cycles and the transmission of policies. I aim at studying both these questions in the near future.

Recent and ongoing changes in the US and European labor markets put my results into perspective. Employment arrangements are changing rapidly through more flexible scheduling, working from home, and part-time work (Katz and Krueger, 2016), as well as through the fragmentation of workplaces and the rise of secondary-jobs held in the "gig" economy. This means that conventional models and estimates of elasticities within primary jobs become less and less accurate descriptions of labor supply responses to shocks. Another related and pressing issue, where the labor supply forces play an important role, is the labor-displacing effects of automation. In analyzing these effects and evaluating how advances in robotics technology may reduce employment and wages, the size of the Frisch elasticity is a key measure (Acemoglu and Restrepo, 2017). These changes highlight the importance of studying many new aspects of labor supply.

References

AARBU, K. O. AND T. O. THORESEN (2001): "Income responses to tax changes—evidence from the Norwegian tax reform," *National Tax Journal*, 319–335.

ACEMOGLU, D. AND P. RESTREPO (2017): "Robots and jobs: Evidence from US labor markets," NBER Working Paper No. 23285, working Paper.

AGUIAR, M., E. HURST, AND L. KARABARBOUNIS (2013): "Time use during the great recession," *American Economic Review*, 103, 1664–96.

ALTONJI, J. G. (1986): "Intertemporal substitution in labor supply: Evidence from micro data," *Journal of Political Economy*, 94, S176–S215.

ALTONJI, J. G. AND C. H. PAXSON (1988): "Labor supply preferences, hours constraints, and hours-wage trade-offs," *Journal of labor economics*, 6, 254–276.

——— (1992): "Labor supply, hours constraints, and job mobility," *Journal of Human Resources*, 27, 256–279.

ALTUG, S. AND R. A. MILLER (1990): "Household choices in equilibrium," *Econometrica: Journal of the Econometric Society*, 543–570.

ANGRIST, J. D. (1991): "Grouped-data estimation and testing in simple labor-supply models," *Journal of Econometrics*, 47, 243–266.

ANGRIST, J. D., S. CALDWELL, AND J. V. HALL (2017): "Uber vs. Taxi: A Driver's Eye View," NBER Working Paper No. 23891.

APPS, P. F. AND R. REES (1988): "Taxation and the Household," *Journal of Public Economics*, 35, 355–369.

ATHEY, S. AND G. IMBENS (2016): "Recursive partitioning for heterogeneous causal effects," *Proceedings of the National Academy of Sciences*, 113, 7353–7360.

ATTANASIO, O., P. LEVELL, H. LOW, AND V. SÁNCHEZ-MARCOS (2018): "Aggregating Elasticities: Intensive and Extensive Margins of Female Labour Supply," *Econometrica*, forthcoming.

AUTEN, G. AND R. CARROLL (1999): "The Effect Of Income Taxes On Household Income," *The Review of Economics and Statistics*, 81, 681–693.

AUTOR, D., A. R. KOSTOL, M. MOGSTAD, B. SETZLER, ET AL. (2017): "Disability benefits, consumption insurance, and household labor supply," Working paper, MIT.

AUTOR, D. H., L. F. KATZ, AND M. S. KEARNEY (2006): "The polarization of the US labor market," *American economic review*, 96, 189–194.

AUTOR, D. H., F. LEVY, AND R. J. MURNANE (2003): "The skill content of recent technological change: An empirical exploration," *The Quarterly journal of economics*, 118, 1279–1333.

BALL, L. (1990): "Intertemporal substitution and constraints on labor supply: Evidence from panel data," *Economic Inquiry*, 28, 706–724.

BARRETT, G. F. AND D. S. HAMERMESH (2017): "Labor Supply Elasticities: Overcoming Nonclassical Measurement Error Using More Accurate Hours Data," *Journal of Human Resources*, 1216–8407R1.

BATTISTI, M., R. MICHAELS, AND C. PARK (2016): "Labor supply within the firm," Ifo Working Paper No. 222.

BECKER, G. S. (1965): "A Theory of the Allocation of Time," *The economic journal*, 493–517.

——— (1973): "A theory of marriage: Part I," *Journal of Political economy*, 81, 813–846.

——— (1976): *The economic approach to human behavior*, Chicago: University of Chicago Press.

BENHABIB, J., R. ROGERSON, AND R. WRIGHT (1991): "Homework in macroeconomics: Household production and aggregate fluctuations," *Journal of Political economy*, 99, 1166–1187.

BERTRAND, M., E. DUFLO, AND S. MULLAINATHAN (2004): "How much should we trust differences-in-differences estimates?" *The Quarterly journal of economics*, 119, 249–275.

BERTRAND, M., E. KAMENICA, AND J. PAN (2015): "Gender identity and relative income within households," *The Quarterly Journal of Economics*, 130, 571–614.

BIANCHI, M., B. R. GUDMUNDSSON, AND G. ZOEGA (2001): "Iceland's Natural Experiment in Supply-Side Economics," *American Economic Review*, 91, 1564–1579.

BIDDLE, J. E. (1988): "Intertemporal substitution and hours restrictions," *The Review of Economics and Statistics*, 347–351.

BITTMAN, M., P. ENGLAND, L. SAYER, N. FOLBRE, AND G. MATHESON (2003): "When does gender trump money? Bargaining and time in household work," *American Journal of sociology*, 109, 186–214.

BLAU, D. M. (1998): "Labor force dynamics of older married couples," *Journal of Labor Economics*, 16, 595–629.

BLAU, F. D. AND L. M. KAHN (2007): "Changes in the labor supply behavior of married women: 1980–2000," *Journal of Labor Economics*, 25, 393–438.

BLOMQUIST, S. AND H. SELIN (2010): "Hourly wage rate and taxable labor income responsiveness to changes in marginal tax rates," *Journal of Public Economics*, 94, 878–889.

BLUNDELL, R., M. COSTA DIAS, C. MEGHIR, AND J. SHAW (2016a): "Female labor supply, human capital, and welfare reform," *Econometrica*, 84, 1705–1753.

BLUNDELL, R. AND M. C. DIAS (2009): "Alternative approaches to evaluation in empirical microeconomics," *Journal of Human Resources*, 44, 565–640.

BLUNDELL, R. AND T. MACURDY (1999): "Labor supply: A review of alternative approaches," *Handbook of labor economics*, 3, 1559–1695.

BLUNDELL, R., C. MEGHIR, AND P. NEVES (1993): "Labour supply and intertemporal substitution," *Journal of Econometrics*, 59, 137–160.

BLUNDELL, R., L. PISTAFERRI, AND I. SAPORTA-EKSTEN (2016b): "Consumption inequality and family labor supply," *American Economic Review*, 106, 387–435.

——— (2018): "Children, time allocation and consumption insurance," *Journal of Political Economy*, 126, S73–S115.

BLUNDELL, R. AND I. WALKER (1986): "A life-cycle consistent empirical model of family labour supply using cross-section data," *The Review of Economic Studies*, 53, 539–558.

BORUSYAK, K. AND X. JARAVEL (2018): "Revisiting event study designs," Working paper, The London School Economics.

BREDEMEIER, C. AND F. JUESSEN (2013): "Assortative mating and female labor supply," *Journal of Labor Economics*, 31, 603–631.

BREIMAN, L. (2001): "Random forests," *Machine learning*, 45, 5–32.

BREWER, M., E. SAEZ, AND A. SHEPHARD (2010): "Means-testing and tax rates on earnings," *Dimensions of Tax Design: the mirrlees review*, 1.

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

BROWNING, M. (2005): "A working paper from April 1985: Which demand elasticities do we know and which do we need to know for policy analysis?" *Research in Economics*, 59, 293–320.

BROWNING, M., A. DEATON, AND M. IRISH (1985): "A profitable approach to labor supply and commodity demands over the life-cycle," *Econometrica: journal of the econometric society*, 503–543.

CAMERER, C., L. BABCOCK, G. LOEWENSTEIN, AND R. THALER (1997): "Labor supply of New York City cabdrivers: One day at a time," *The Quarterly Journal of Economics*, 112, 407–441.

CARRINGTON, W. J. (1996): "The Alaskan labor market during the pipeline era," *Journal of Political Economy*, 104, 186–218.

CENTRAL BANK OF ICELAND (2018): *The Economy of Iceland*, Central Bank of Iceland.

CESARINI, D., E. LINDQVIST, M. J. NOTOWIDIGDO, AND R. ÖSTLING (2017): "The effect of wealth on individual and household labor supply: evidence from Swedish lotteries," *American Economic Review*, 107, 3917–3946.

CHETTY, R. (2012): "Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply," *Econometrica*, 80, 969–1018.

CHETTY, R., J. N. FRIEDMAN, T. OLSEN, AND L. PISTAFERRI (2011): "Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records," *The quarterly journal of economics*, 126, 749–804.

CHETTY, R., A. GUREN, D. MANOLI, AND A. WEBER (2013): "Does indivisible labor explain the difference between micro and macro elasticities? A meta-analysis of extensive margin elasticities," *NBER Macroeconomics Annual*, 27, 1–56.

CHETTY, R., A. LOONEY, AND K. KROFT (2009): "Salience and taxation: Theory and evidence," *American economic review*, 99, 1145–77.

CHIAPPORI, P.-A. (1988): "Rational household labor supply," *Econometrica*, 63–90.

CHIAPPORI, P.-A. AND M. MAZZOCCO (2017): "Static and intertemporal household decisions," *Journal of Economic Literature*, 55, 985–1045.

CONESA, J. C., S. KITAO, AND D. KRUEGER (2009): "Taxing capital? Not a bad idea after all!" *American Economic Review*, 99, 25–48.

CULLEN, J. B. AND J. GRUBER (2000): "Does unemployment insurance crowd out spousal labor supply?"

Journal of labor Economics, 18, 546–572.

DICKENS, W. T. AND S. J. LUNDBERG (1993): "Hours Restrictions and Labor Supply," *International Economic Review*, 169–192.

DINARDO, J., N. M. FORTIN, AND T. LEMIEUX (1996): "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, 64, 1001–1044.

DONNI, O. AND P.-A. CHIAPPORI (2011): "Nonunitary models of household behavior: a survey of the literature," in *Household economic behaviors*, Springer, 1–40.

EUWALS, R. (2001): "Female labour supply, flexibility of working hours, and job mobility," *The Economic Journal*, 111, 120–134.

FARBER, H. S. (2015): "Why you can't find a taxi in the rain and other labor supply lessons from cab drivers," *The Quarterly Journal of Economics*, 130, 1975–2026.

FARRELL, D. AND F. GREIG (2016): "Paychecks, paydays, and the online platform economy: Big data on income volatility," *JP Morgan Chase Institute*.

FEHR, E. AND L. GOETTE (2007): "Do workers work more if wages are high? Evidence from a randomized field experiment," *American Economic Review*, 97, 298–317.

FELDSTEIN, M. (1995): "The effect of marginal tax rates on taxable income: a panel study of the 1986 Tax Reform Act," *Journal of Political Economy*, 103, 551–572.

FRENCH, E. (2004a): "The Labor Supply Response to (Mismeasured but) Predictable Wage Changes," *The Review of Economics and Statistics*, 86, 602–613.

——— (2004b): "The labor supply response to (mismeasured but) predictable wage changes," *Review of Economics and Statistics*, 86, 602–613.

GELBER, A. M. (2014): "Taxation and the earnings of husbands and wives: evidence from Sweden," *Review of Economics and Statistics*, 96, 287–305.

GIERTZ, S. H. (2007): "The Elasticity of Taxable Income over the 1980s and 1990s," *National Tax Journal*, 743–768.

GINÉ, X., M. MARTINEZ-BRAVO, AND M. VIDAL-FERNÁNDEZ (2017): "Are labor supply decisions consistent with neoclassical preferences? Evidence from Indian boat owners," *Journal of Economic Behavior & Organization*, 142, 331–347.

GOLDBERG, J. (2016): "Kwacha gonna do? Experimental evidence about labor supply in rural Malawi," *American Economic Journal: Applied Economics*, 8, 129–49.

GOLDIN, C. AND L. F. KATZ (2016): "A most egalitarian profession: pharmacy and the evolution of a family-friendly occupation," *Journal of Labor Economics*, 34, 705–746.

GOOLSBEE, A. (2000): "What happens when you tax the rich? Evidence from executive compensation," *Journal of Political Economy*, 108, 352–378.

GOOS, M., A. MANNING, AND A. SALOMONS (2014): "Explaining job polarization: Routine-biased technological change and offshoring," *American Economic Review*, 104, 2509–26.

GOUX, D., E. MAURIN, AND B. PETRONGOLO (2014): "Worktime regulations and spousal labor supply," *American Economic Review*, 104, 252–76.

GRUBER, J. (1994): "The incidence of mandated maternity benefits," *The American economic review*, 622–641.

GRUBER, J. AND E. SAEZ (2002): "The elasticity of taxable income: evidence and implications," *Journal of public Economics*, 84, 1–32.

GRUBER, J. AND D. WISE (1999): "Social Security Programs and Retirement Around the World: Introduction and Summary," in *Social Security and Retirement Around the World*, University of Chicago Press.

GUSTMAN, A. L. AND T. L. STEINMEIER (2000): "Retirement in dual-career families: a structural model," *Journal of Labor economics*, 18, 503–545.

HALL, J. V. AND A. B. KRUEGER (2018): "An analysis of the labor market for Uber's driver-partners in the United States," *ILR Review*, 71, 705–732.

HAM, J. C. AND K. T. REILLY (2002): "Testing intertemporal substitution, implicit contracts, and hours restriction models of the labor market using micro data," *American Economic Review*, 92, 905–927.

HANSSON, Å. (2007): "Taxpayers' responsiveness to tax rate changes and implications for the cost of taxation in Sweden," *International Tax and Public Finance*, 14, 563–582.

HAVRÁNEK, T. (2015): "Measuring intertemporal substitution: The importance of method choices and selective reporting," *Journal of the European Economic Association*, 13, 1180–1204.

HEATHCOTE, J., K. STORESLETTEN, AND G. L. VIOLANTE (2017): "Optimal tax progressivity: An analytical framework," *The Quarterly Journal of Economics*, 132, 1693–1754.

HECKMAN, J. AND T. E. MACURDY (1980): "A Life Cycle Model of Female Labor Supply," *Review of Economic Studies*, 41, 47–74.

HECKMAN, J. J. (1993): "What has been learned about labor supply in the past twenty years?" *The American Economic Review*, 83, 116–121.

HECKMAN, J. J., H. ICHIMURA, AND P. E. TODD (1997): "Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme," *The review of economic studies*, 64, 605–654.

HEIM, B. T. (2007): "The incredible shrinking elasticities married female labor supply, 1978–2002," *Journal of Human resources*, 42, 881–918.

IACUS, S. M., G. KING, AND G. PORRO (2012): "Causal inference without balance checking: Coarsened exact matching," *Political analysis*, 20, 1–24.

IMAI, S. AND M. P. KEANE (2004): "Intertemporal labor supply and human capital accumulation," *International Economic Review*, 45, 601–641.

IMBENS, G. W., D. B. RUBIN, AND B. I. SACERDOTE (2001): "Estimating the effect of unearned income on labor earnings, savings, and consumption: Evidence from a survey of lottery players," *American Economic Review*, 778–794.

IMROHOROĞLU, S. AND S. KITAO (2012): "Social Security reforms: Benefit claiming, labor force participation, and long-run sustainability," *American Economic Journal: Macroeconomics*, 4, 96–127.

KAHN, S. AND K. LANG (1991): "The effect of hours constraints on labor supply estimates," *The Review of Economics and Statistics*, 605–611.

KATZ, L. F. AND A. B. KRUEGER (2016): "The rise and nature of alternative work arrangements in the United States, 1995-2015," NBER Working Paper No. 22667.

KEANE, M. AND R. ROGERSON (2015): "Reconciling micro and macro labor supply elasticities: a structural perspective," *Annual Review of Economics*, 7, 89–117.

KEANE, M. P. (2011): "Labor supply and taxes: A survey," *Journal of Economic Literature*, 49, 961–1075.

KILLINGSWORTH, M. R. (1983): *Labor supply*, Cambridge: Cambridge University Press.

KINDLUND, H. AND D. BITERMAN (2002): ". Work and the labour market," *International Journal of Social Welfare*, 11.

KLEVEN, H. J., M. B. KNUDSEN, C. T. KREINER, S. PEDERSEN, AND E. SAEZ (2011): "Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark," *Econometrica*, 79, 651–692.

KLEVEN, H. J. AND E. A. SCHULTZ (2014): "Estimating taxable income responses using Danish tax reforms," *American Economic Journal: Economic Policy*, 6, 271–301.

KOPCZUK, W. (2005): "Tax bases, tax rates and the elasticity of reported income," *Journal of Public Economics*, 89, 2093–2119.

——— (2012): "The Polish business "flat" tax and its effect on reported incomes: a Pareto improving tax reform," Working paper, Columbia University.

KOUSTAS, D. K. (2018): "Consumption Insurance and Multiple Jobs: Evidence from Rideshare Drivers," Job market paper, University of California, Berkely.

KRISHNAN, P. (1990): "The economics of moonlighting: A double self-selection model," *The review of economics and statistics*, 361–367.

LABANCA, C. AND D. POZZOLI (2018): "Coordination of Hours within the Firm," Working Paper.

LEE, C.-I. (2001): "Finite sample bias in IV estimation of intertemporal labor supply models: is the intertemporal substitution elasticity really small?" *Review of Economics and Statistics*, 83, 638–646.

LOONEY, A. AND M. SINGHAL (2006): "The effect of anticipated tax changes on intertemporal labor supply and the realization of taxable income," NBER Working Paper No. 12417.

LUCAS, R. E. AND L. A. RAPPING (1969): "Real wages, employment, and inflation," *Journal of political economy*, 77, 721–754.

LUNDBERG, S. (1985): "The added worker effect," *Journal of Labor Economics*, 3, 11–37.

MACURDY, T. E. (1981): "An empirical model of labor supply in a life-cycle setting," *Journal of political Economy*, 89, 1059–1085.

MANOLI, D. AND A. WEBER (2016): "Nonparametric evidence on the effects of financial incentives on retirement decisions," *American Economic Journal: Economic Policy*, 8, 160–182.

MARTINEZ, I. Z., E. SAEZ, AND M. SIEGENTHALER (2018): "Intertemporal Labor Supply Substitution? Evidence from the Swiss Income Tax Holidays," NBER Working Paper No. 24634.

MARTINEZ-GRANADO, M. (2005): "Testing labour supply and hours constraints," *Labour Economics*, 12, 321–343.

MAS, A. AND A. PALLAIS (2017): "Valuing alternative work arrangements," *American Economic Review*, 107, 3722–59.

MCCLELLAND, R. AND S. MOK (2012): "A review of recent research on labor supply elasticities," Working paper, Congressional Budget Office.

MEGHIR, C. AND D. PHILLIPS (2010): "Labour supply and taxes," *Dimensions of tax design: The Mirrlees review*, 202–74.

NEKOEI, A. AND D. SEIM (2018): "How do Inheritances Shape Wealth Inequality? Theory and Evidence from Sweden," CEPR Discussion Paper 13199.

O'CONNELL, J. F. (1979): "Multiple job holding and marginal tax rates," *National Tax Journal*, 32, 73–76.

OECD (1991): *OECD Economic Surveys: Iceland 1991*, Paris: OECD Publishing.

——— (2007): *OECD Economic Surveys: Iceland 2007*, Paris: OECD Publishing.

OETTINGER, G. S. (1999): "An empirical analysis of the daily labor supply of stadium venors," *Journal of political Economy*, 107, 360–392.

ÓLAFSDÓTTIR, T., B. HRAFNKELSSON, G. THORGEIRSSON, AND T. L. ÁSGEIRSDÓTTIR (2016): "The tax-free year in Iceland: A natural experiment to explore the impact of a short-term increase in labor supply on the risk of heart attacks," *Journal of health economics*, 49, 14–27.

OLGEIRSSON, F. G. (2013): *Í þágu þjóðar. Saga skatta og skattkerfisbreytinga á Íslandi 1877-2012*, Reykjavik, Iceland: Directorate of Internal Revenue.

PAUL, R. E. (1954): *Taxation in the United States*, Boston: Little, Brown and Company.

PAXSON, C. H. AND N. SICHERMAN (1996): "The dynamics of dual job holding and job mobility," *Journal of labor economics*, 14, 357–393.

PENCAVEL, J. (1986): "Labor supply of men: a survey," *Handbook of labor economics*, 1, 3–102.

PERLMAN, R. (1966): "Observations on overtime and moonlighting," *Southern Economic Journal*, 237–244.

PISTAFERRI, L. (2003): "Anticipated and unanticipated wage changes, wage risk, and intertemporal labor supply," *Journal of Labor Economics*, 21, 729–754.

ROGERSON, R. AND J. WALLENIUS (2009): "Micro and macro elasticities in a life cycle model with taxes," *Journal of Economic theory*, 144, 2277–2292.

ROSENBAUM, P. R. (2002): "Observational studies," in *Observational studies*, Springer, 1–17.

RUBIN, D. B. (2006): *Matched sampling for causal effects*, Cambridge University Press.

RUPERT, P., R. ROGERSON, AND R. WRIGHT (1995): "Estimating substitution elasticities in household production models," *Economic theory*, 6, 179–193.

——— (2000): "Homework in labor economics: household production and intertemporal substitution," *Journal of Monetary Economics*, 46, 557–579.

SAEZ, E. (2003): "The effect of marginal tax rates on income: a panel study of 'bracket creep'," *Journal of Public*

Economics, 87, 1231–1258.

SAEZ, E., J. SLEMROD, AND S. H. GIERTZ (2012): "The elasticity of taxable income with respect to marginal tax rates: A critical review," *Journal of economic literature*, 50, 3–50.

SAEZ, E. AND M. R. VEALL (2005): "The evolution of high incomes in Northern America: lessons from Canadian evidence," *American Economic Review*, 95, 831–849.

SAIA, A. (2017): "Why Do Taxi Drivers Seem To Be Target Earners? Insights From the Labor Supply Behavior of Pizza Vendors," Working Paper.

SHISHKO, R. AND B. ROSTKER (1976): "The economics of multiple job holding," *The American Economic Review*, 298–308.

SIGURDSSON, J. AND R. SIGURDARDOTTIR (2016): "Time-dependent or state-dependent wage-setting? Evidence from periods of macroeconomic instability," *Journal of Monetary Economics*, 78, 50–66.

SILLAMAA, M.-A. AND M. R. VEALL (2001): "The effect of marginal tax rates on taxable income: a panel study of the 1988 tax flattening in Canada," *Journal of Public Economics*, 80, 341–356.

STAFFORD, T. M. (2015): "What do fishermen tell us that taxi drivers do not? An empirical investigation of labor supply," *Journal of Labor Economics*, 33, 683–710.

STANTCHEVA, S. (2017): "Optimal taxation and human capital policies over the life cycle," *Journal of Political Economy*, 125, 1931–1990.

STEFÁNSSON, A. (2019): "Labor supply response to a tax holiday: The take-home from a large and salient shock," Mimeo, Uppsala University.

STEPHENS, JR, M. (2002): "Worker displacement and the added worker effect," *Journal of Labor Economics*, 20, 504–537.

STEWART, M. B. AND J. K. SWAFFIELD (1997): "Constraints on the desired hours of work of British men," *The Economic Journal*, 107, 520–535.

TAZHITDINOVA, A. (2017): "When tax incentives, hour constraints and job availabilities clash: Moonlighting responses to a large tax reform," Working paper, UC Santa Barbara.

THOMSEN, P. M. (2018): "Ragnarök: Iceland's Crisis, its Successful Stabilization Program, and the Role of the IMF," Speech, International Monetary Fund.

WAGER, S. AND S. ATHEY (2017): "Estimation and inference of heterogeneous treatment effects using random forests," *Journal of the American Statistical Association*.

ZILIAK, J. P. AND T. J. KNIESNER (1999): "Extimating Life Cycle labor Supply Tax Effects," *Journal of Political Economy*, 107, 326–359.

——— (2005): "The effect of income taxation on consumption and labor supply," *Journal of Labor Economics*, 23, 769–796.

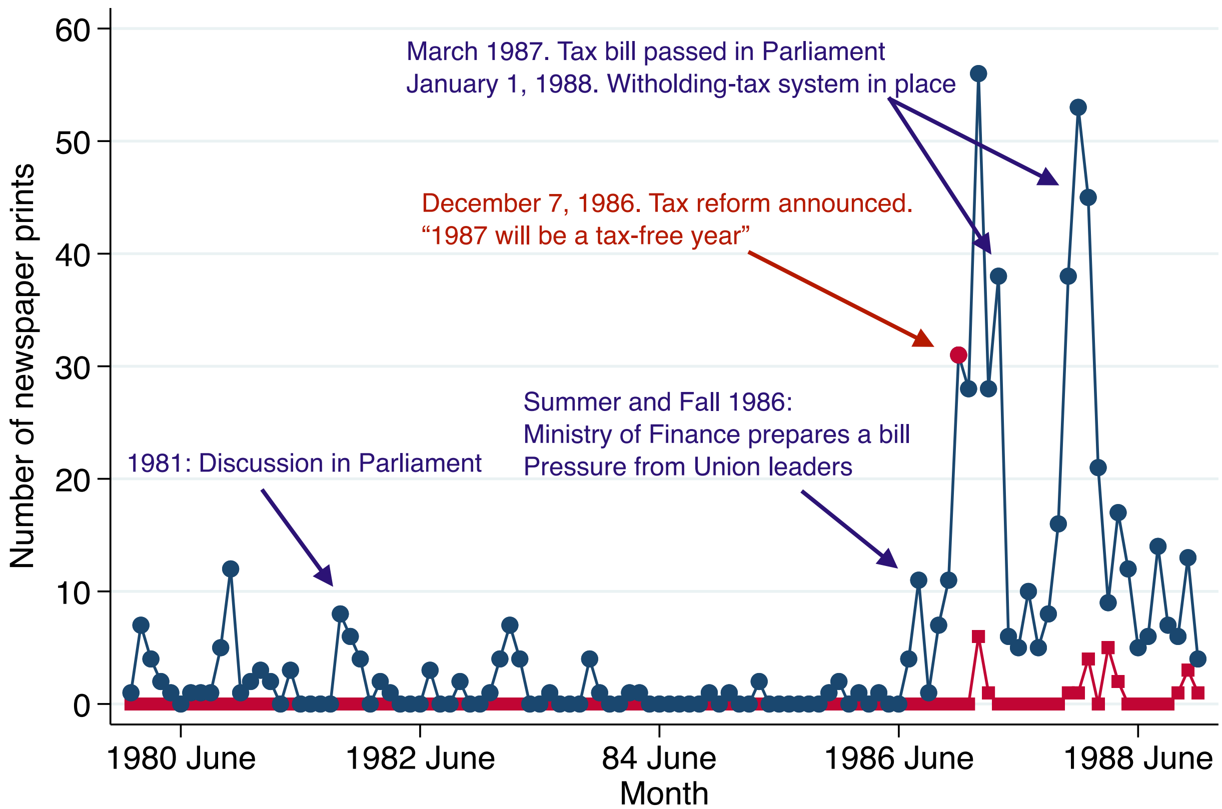
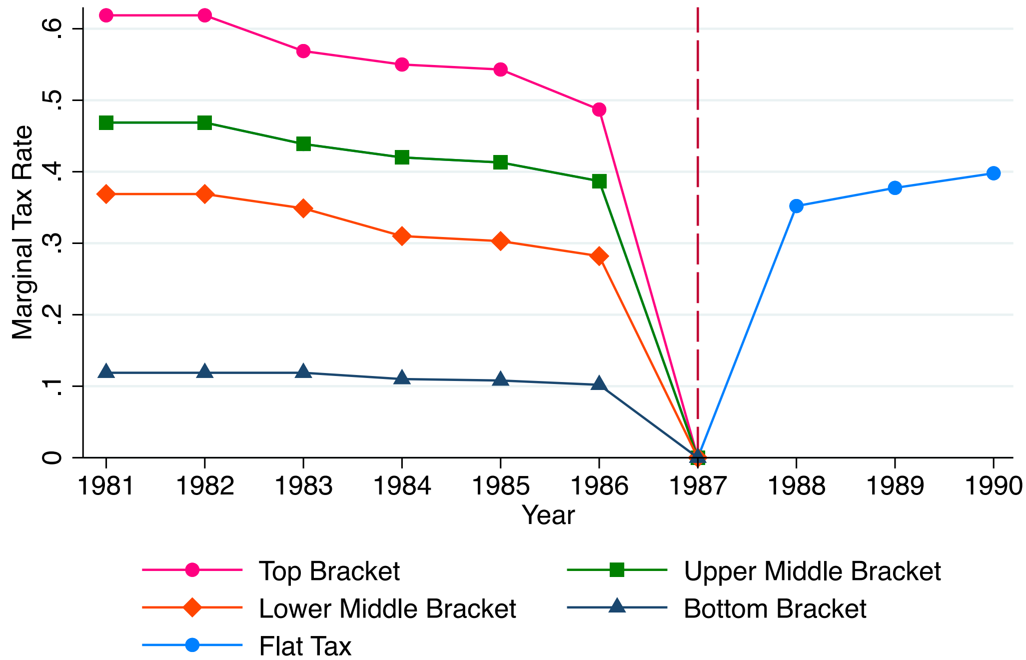


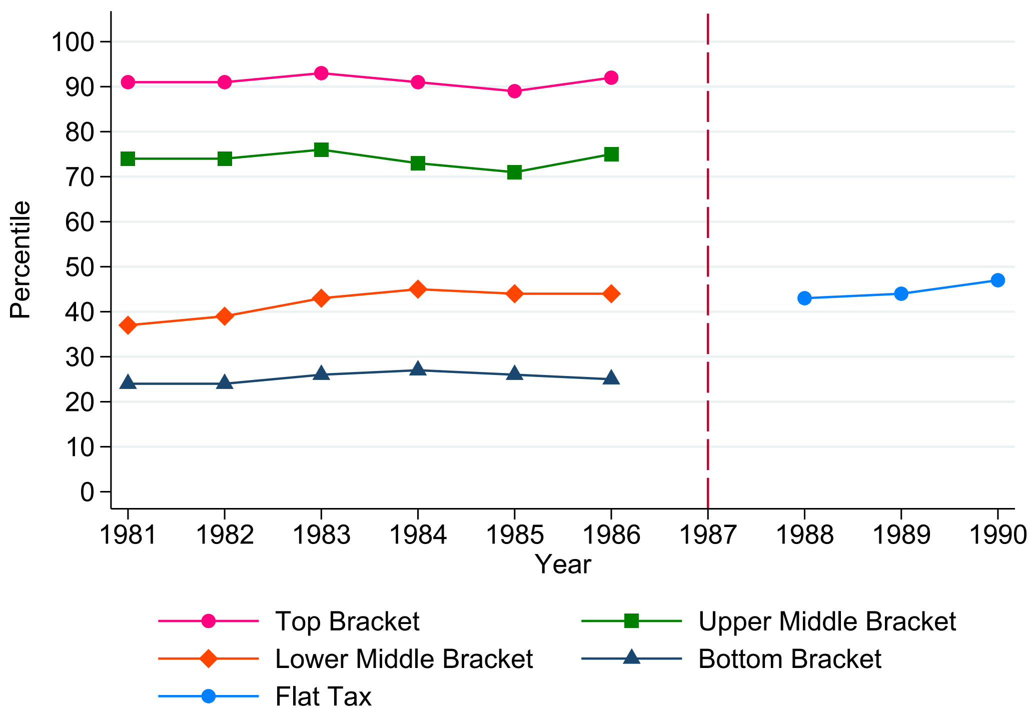
Figure 4: Number of Printed Newspapers Mentioning Withholding-Tax

Notes: The figure plots in blue dots the number of printed newspapers mentioning a withholding-based pay-as-you-earn tax system per month during the period January 1980 to December 1988. Appendix B provides a detailed time-line of events. The keywords searched for were “Staðgreiðsla skatta” and “Staðgreiðslukerfi skatta”. In red squares I plot a similar count of newspapers mentioning a flat tax system, which was adopted in 1988. The keywords searched for were “eitt skatthlutfall”,“eitt skattþrep” and “flatur skattur”. The count is based on searches in the Icelandic newspaper database Tímarit.is in the six main newspapers (*Alþýðublaðið, Dagblaðið Vísir (DV), Dagur, Morgunblaðið, Tíminn, Þjóðviljinn*). The total number of printed newspapers per month is about 145 on average.



(a) Marginal tax rate by tax brackets

Notes: The figure shows the evolution of statutory marginal tax rates by tax-brackets, where local-level tax is the average across municipalities. Small lump-sum and flat income taxes, such as health insurance contribution, cemetery charge, church tax and contribution to the construction fund for the elderly, are excluded in the graph.

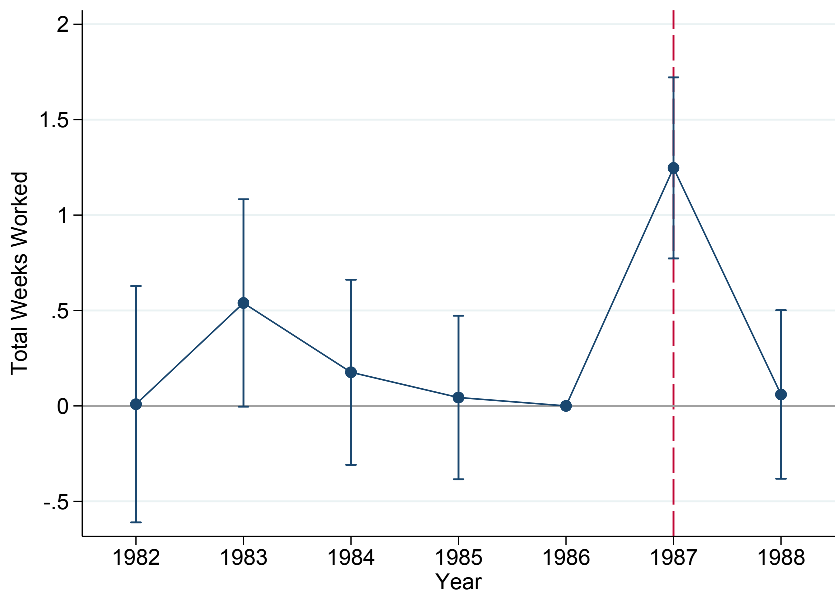


(b) Tax-bracket thresholds in percentiles of income

Notes: The figure shows the evolution of tax-bracket thresholds, set in nominal values and updated regularly by the Icelandic Parliament to account for changes in prices and wages. The thresholds are presented as the percentile of the taxable income distribution each year. Calculations assume that workers deduct the statutory minimum of 10% from their national-level income tax base each year. For more details on the Icelandic tax system and tax deductions, see Appendix A.

Figure 5: Marginal tax rates and tax-bracket thresholds

(a) Reduced Form: Log Labor Earnings



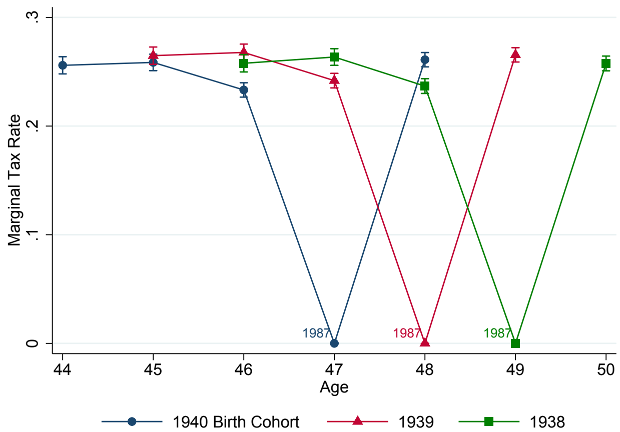
(b) Reduced Form: Total Weeks Worked

Figure 6: Dynamic Difference-in-Difference — Placebo Tests

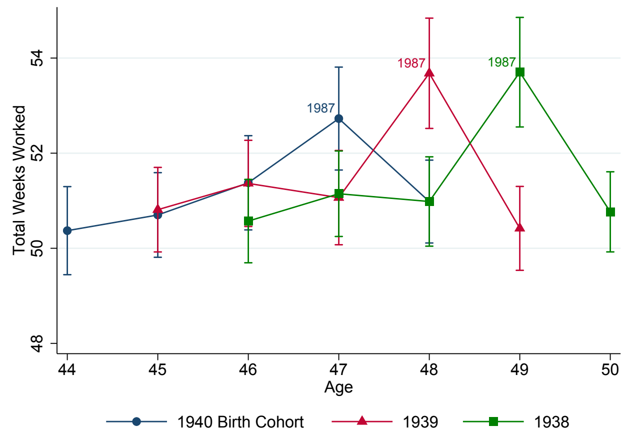
Notes: The figures present estimates from a dynamic DD version of equation (1), estimated in the following regression

$$y_{it} = bracket_{i,t-1} + \delta_t + \eta_t \cdot B_{i,t-1} \times \delta_t + \boldsymbol{X'}_{it}\gamma + \mu_{it},$$

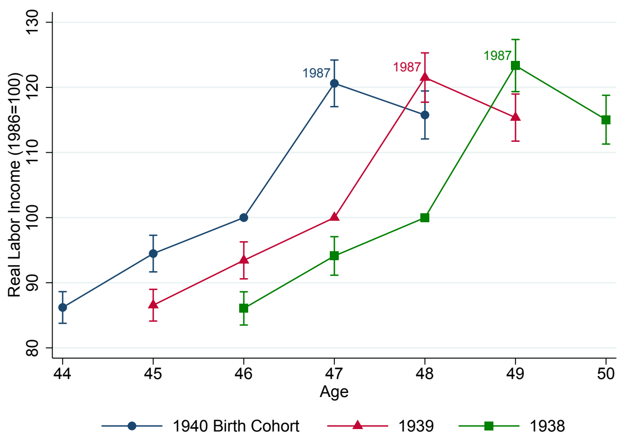
where the outcome variable in panel (a) is log labor earnings and in panel (b) total weeks worked. They plot the coefficients η_t, where $B_{i,t-1} \times \delta_{t=1986}$ is normalized to zero. Standard errors are clustered at the individual level and the vertical bars plot the 95%-confidence intervals. Figures A.5, A.6 and A.7 in the Appendix provide a graphical presentation of the reduced form evidence and the first stage.



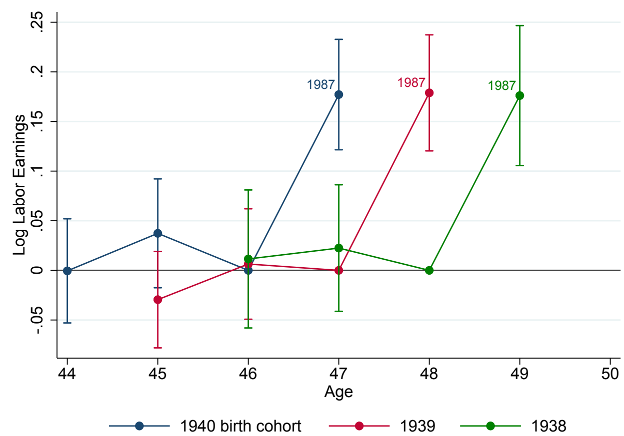
(a) Marginal Tax Rate



(b) Weeks Worked



(c) Labor Earnings (levels)



(d) Labor Earnings (differences)

Figure 7: Graphical Evidence: Life-Cycle Difference-in-Difference Research Design

Notes: The figure shows the evolution of marginal tax rates, weeks worked and labor earnings for birth cohorts born in 1940, 1939 and 1938. In (a) I plot the average marginal tax rate, in (b) the average weeks worked, in (c) the average labor earnings in real terms, normalized to 100 in 1986, and in (d) the average difference in log real labor earnings relative to the cohort born one year earlier, which detrends the series plotted in panel (c). Each graph is based on a matched sample based on the procedure described in the main text in Section 5.1. The vertical bars plot the 95%-confidence intervals.

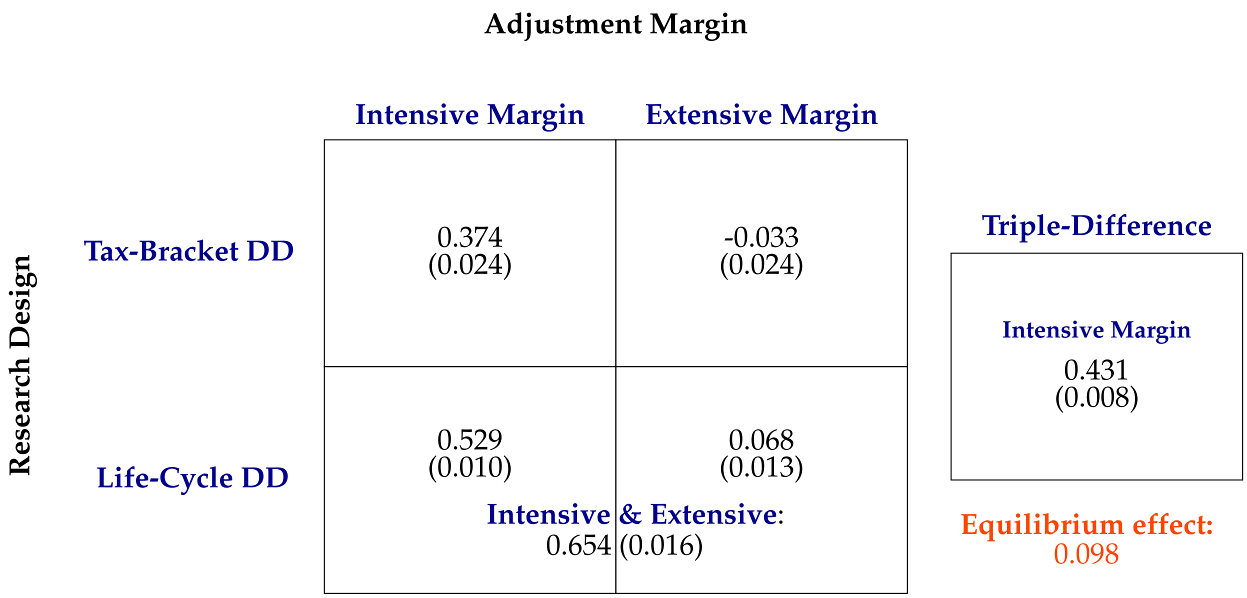
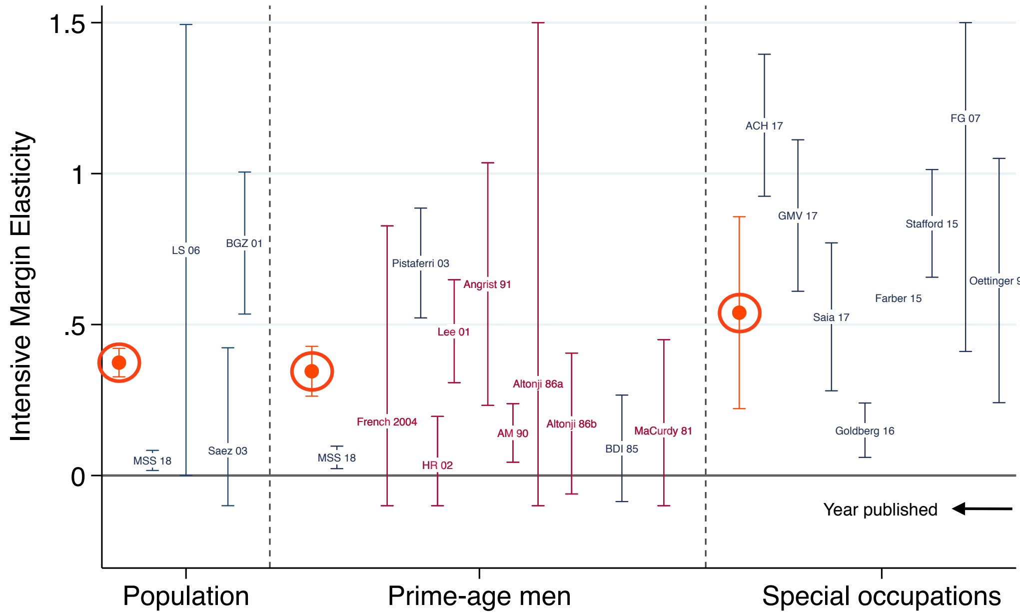


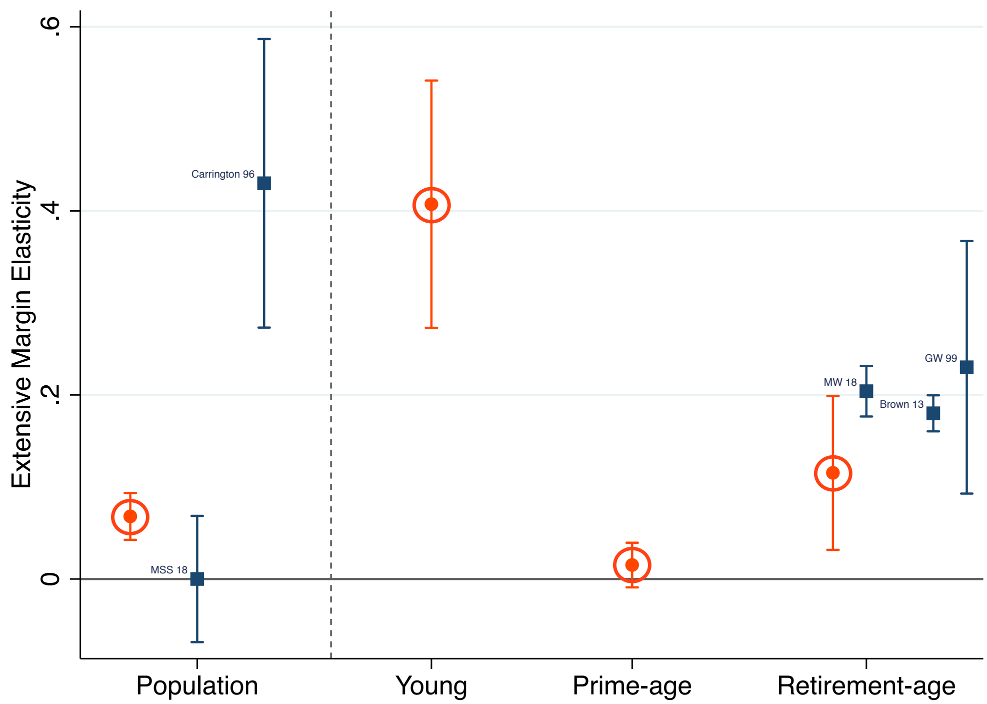
Figure 8: Summary of Frisch Elasticity Estimates

Notes: The figure summarizes estimates of intensive and extensive margin Frisch elasticities using different research designs. The cells reporting estimates of intensive and extensive margin elasticities based on tax-bracket DD are estimates from Tables 1 and 3, respectively. I interpret the earnings elasticity as an intensive margin as the sample is restricted to workers in the labor force prior to the reform and I find no extensive margin responses. Estimates of the extensive margin elasticity based on life-cycle DD are from Table 6. The earnings elasticity using the life-cycle DD, reported in Table 4, combines both intensive and extensive margin responses. Estimates of the intensive margin elasticity using the life-cycle DD are based on the same restricted sample as used applying the tax-bracket DD. The results are reported in Appendix Table A.27. The triple-difference estimate uses the same sample and therefore corresponds to the intensive margin elasticity. Standard errors are in parentheses. The equilibrium effect is the difference between the life-cycle DD and triple-difference intensive-margin estimates $(0.529 - 0.431)$.



(a) Intensive-Margin Frisch Elasticity

Notes: The figure plots estimates of intensive margin Frisch elasticity by subgroup or population studied. Point estimates refer to the authors' main, representative, or preferred specification. 95% confidence intervals are either based on reported standard errors or computed using the delta method, and are censored at 1.5 and -0.10 for visual purposes. My estimates are in orange and circled. The labels are as follows."MMS 18": Martinez, Saez, and Siegenthaler (2018), "LS 06": Looney and Singhal (2006), "Saez 03": Saez (2003), , "BGZ 01": Bianchi, Gudmundsson, and Zoega (2001) "French 04": French (2004b), "Pistaferri 03": Pistaferri (2003), "HR 02": Ham and Reilly (2002), "Lee 01": Lee (2001), "Angrist 01": Angrist (1991), "AM 90": Altug and Miller (1990), "Altonji 86a" and "Altonji 86b": Altonji (1986), "BDI 85": Browning, Deaton, and Irish (1985), "MaCurdy 81": MaCurdy (1981), "ACH 17": Angrist, Caldwell, and Hall (2017), "GMV 17": Giné, Martinez-Bravo, and Vidal-Fernández (2017), "Saia 17": Saia (2017), "Goldberg 16": Goldberg (2016), "Farber 15":Farber (2015), "Stafford 15": Stafford (2015), "FG 07": Fehr and Goette (2007), "Oettinger 99": Oettinger (1999). Estimates in MaCurdy (1981) of 6.25, as reported in Keane (2011), and negative elasticities in Camerer, Babcock, Loewenstein, and Thaler (1997), are excluded for visual purposes.



(b) Extensive-Margin Frisch Elasticity

Notes: The figure plots estimates of extensive margin Frisch elasticity by subgroup or population studied. The point estimates refer to the authors' main, representative, or preferred specification. 95% confidence intervals are either based on reported standard errors or computed using the delta method. My estimates are in orange and circled. The labels are as follows."Carrington 96": Carrington (1996), "MMS 18": Martinez, Saez, and Siegenthaler (2018), "MW 16": Manoli and Weber (2016), "Brown 13": Brown (2013), "GW 99": Gruber and Wise (1999).

Figure 9: Summary of Frisch elasticity estimates

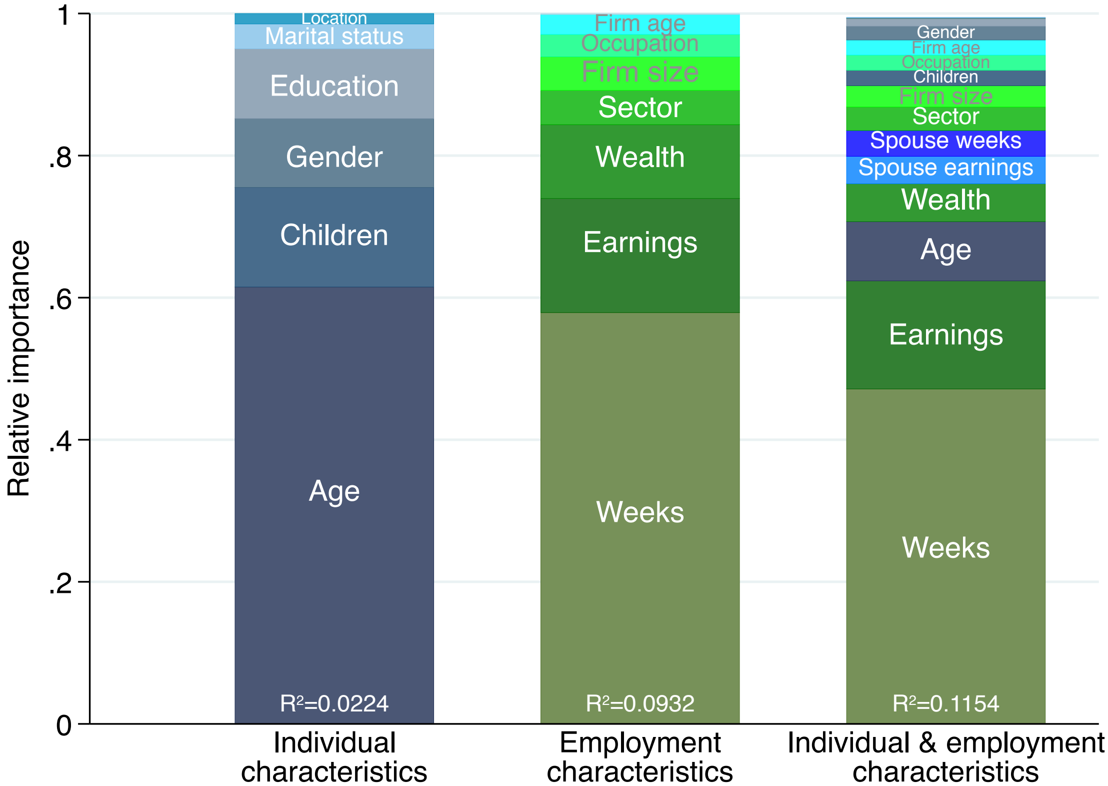
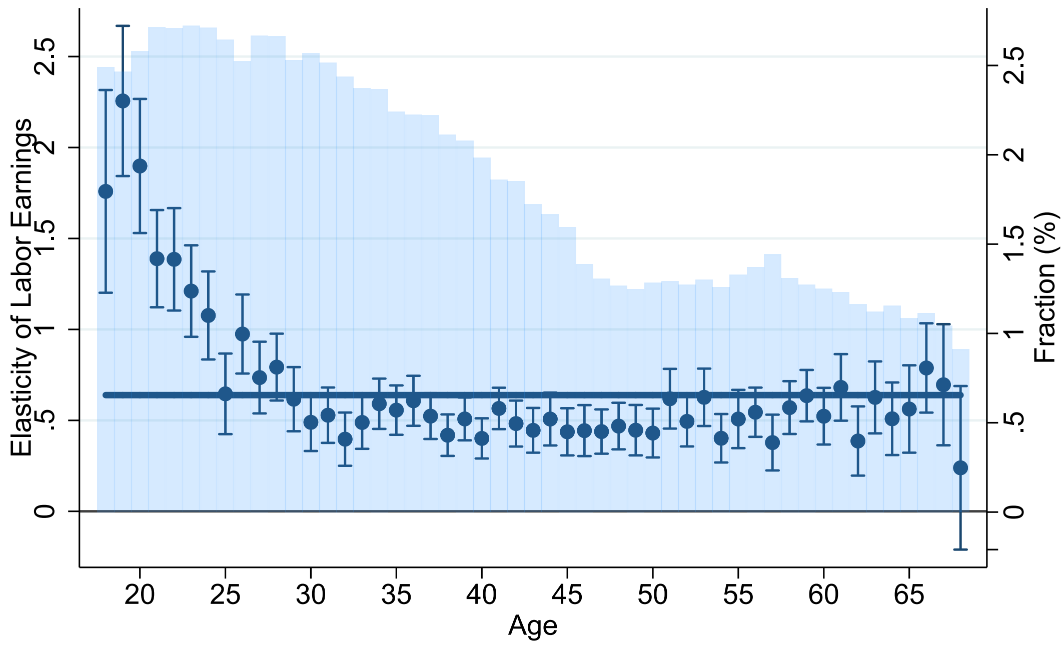
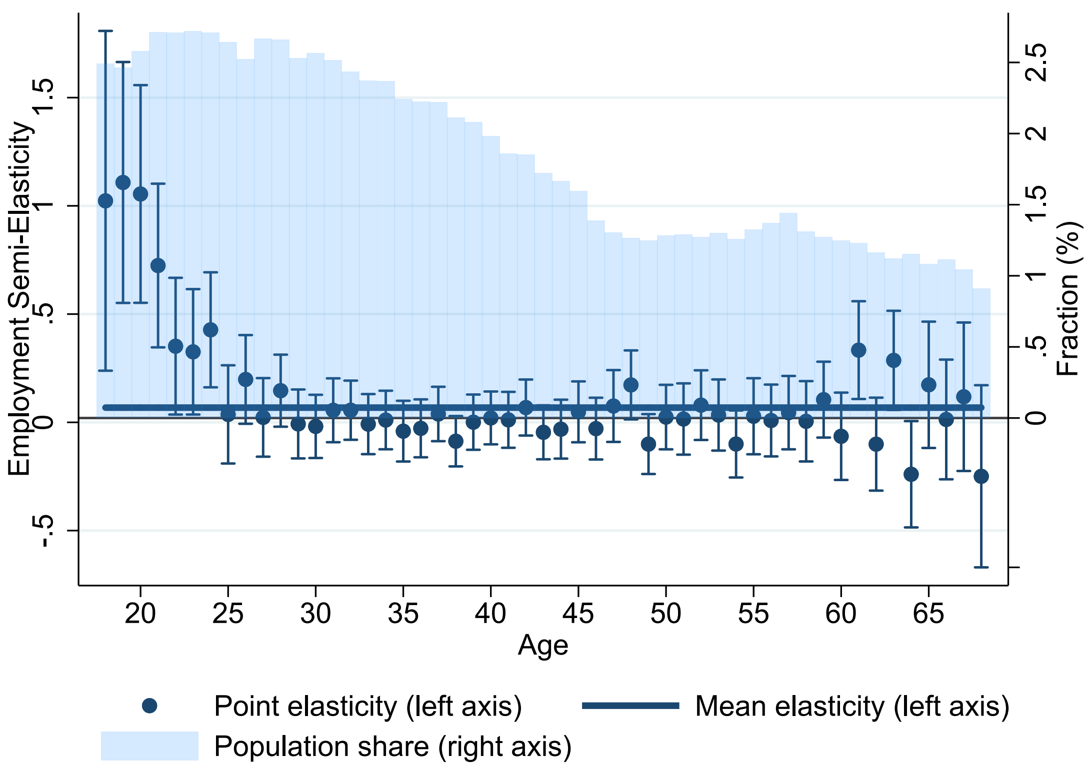


Figure 10: Feature Importance in Explaining Variation in Elasticities

Notes: The figure plots the relative contribution of each feature in predicting labor supply elasticity. This is measured by first estimating labor supply elasticity at the individual level using the life-cycle DD design, matching each individual to a counterfactual constructed from all individuals with the exact same set of characteristics. Then, I predict labor supply elasticity using the available set of characteristics using the random forest algorithm. The importance of each feature is then measured with the gain in prediction achieved over all trees through splits using a given feature. The total gain is normalized to 1, giving the relative importance of each characteristic in each model. R^2 is calculated through cross-validation, where model predictions using the training data are compared to actual values. All employment and job characteristics are pre-reform values as of 1986, except *weeks* which bundles the prediction gain using weeks worked in the three pre-reform years. This measure (*weeks*) serves as my measure of labor-market attachment.



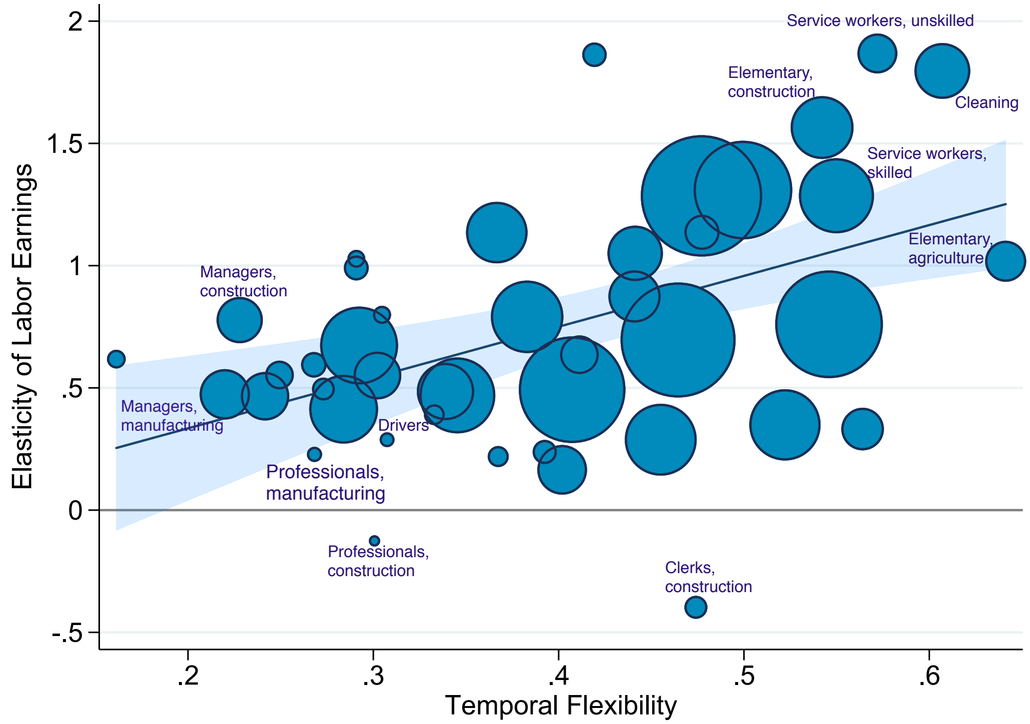
(a) Elasticity of labor earnings



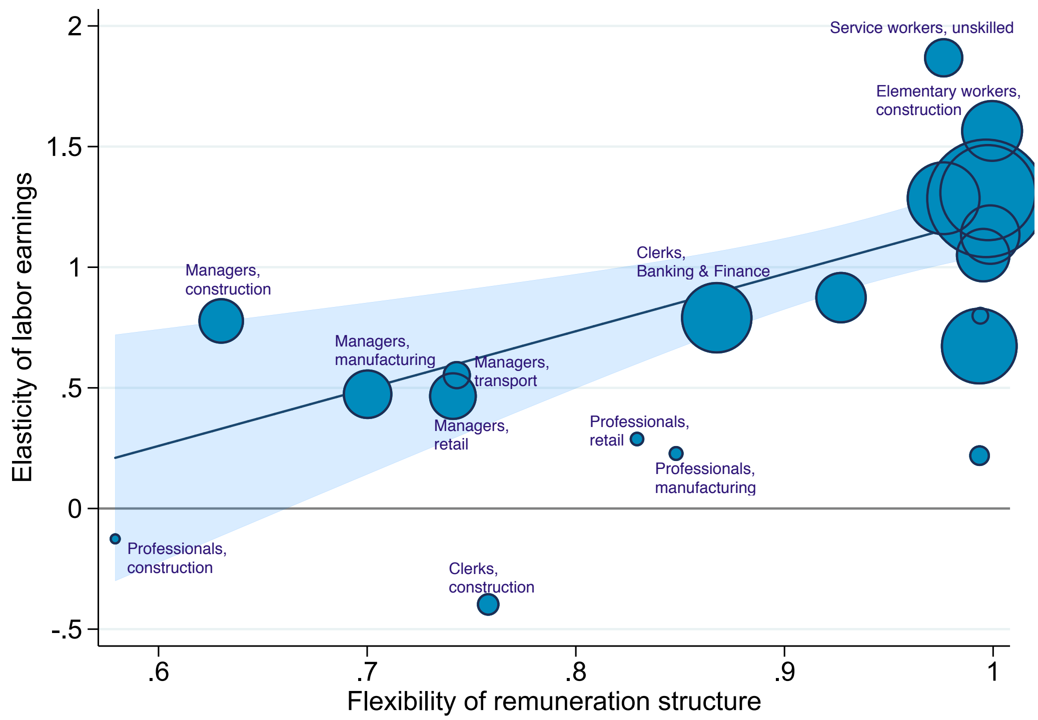
(b) Employment semi-elasticity (extensive margin)

Figure 11: Labor supply elasticity by age

Notes: Panel (a) plots the elasticity of labor earnings for each cohort of age 18-68 in 1987. Each point in the graph is a separate estimate from equation (4), where the dependent variable is the logarithm of labor earnings and the treatment group is of the age denoted on the x-axis in 1987. Panel (b) plots employment semi-elasticity for each cohort estimated in separate regressions according to equation (4), where the dependent variable is an employment indicator. The vertical bars plot the 95%-confidence intervals. The horizontal line plots the average elasticity, as reported in Tables 4 and 6. The shaded area (bars) is the population distribution, where each bar corresponds to the fraction of the working-age population (in %).



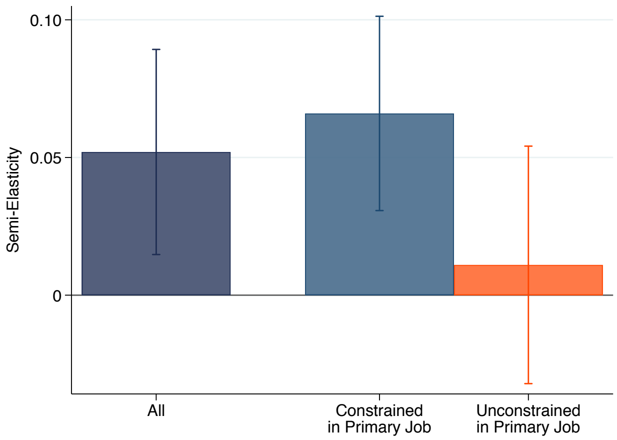
(a) Elasticity by temporal flexibility



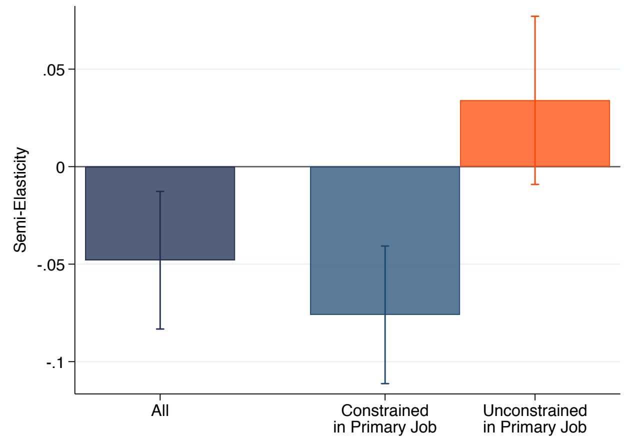
(b) Elasticity by flexibility of remuneration structure

Figure 12: How important are adjustment frictions in shaping labor supply elasticities?

Notes: Each panel plots labor earnings elasticity estimates by groups against a measure of adjustment frictions. In Panel (a), "*temporal flexibility*" is measured with the coefficient of variation in weeks worked, i.e. the occupation-level dispersion in working time. In Panel (b), "*flexibility of remuneration structure*" is measured as one minus the occupation-share of workers with fixed salary contracts. Elasticities are estimated using the life-cycle DD design, after conducting the matching procedure described in Section 5.1 within the set of workers employed in each group pre-reform. This enables me to compare elasticities across occupations without the difference being driven by compositional differences in other characteristics. The size of the dots on the graphs is proportional to the number of workers in each group.



(a) Secondary-job holding



(b) Primary-job change

Notes: The dependent variable in Panel (a) is an indicator that equals one if holding a secondary job, measured by working at least one week on a job other than the primary job within the year, but zero otherwise. The pre-reform mean of this dependent variable is 0.297. The dependent variable in Panel (b) is an indicator that equals one if the primary job is different from the primary job in the previous year, but zero otherwise. The pre-reform mean of this dependent variable is 0.232. *"Constrained in primary job"* is an indicator that equals one if working 52 weeks in the primary job in the prior year, but zero otherwise. The figure presents results from a 2SLS estimation of equation (2), where the net-of-tax rate is instrumented with an interaction between indicators of treatment status and tax-free year. Separate semi-elasticities for those constrained and unconstrained in primary jobs are obtained by interacting the net-of-tax rate and the instrument in equation (2) with the indicator for being constrained. Controls are gender, age, education, marital status, whether living in the capital area or not, and the number of children at age 0-18. The figure shows 95% confidence intervals based on robust standard errors clustered by individual.

(c) Decomposition of treatment effect into primary and secondary jobs

Notes: The figure presents a decomposition of the total treatment effects on labor earnings and weeks worked into subcomponents, as described by equation (8). Calculations are based on estimates of equation (2) in levels of each outcome and the numbers presented are the contribution of each component to the total effect.

Figure 13: Secondary-job holding and primary-job change

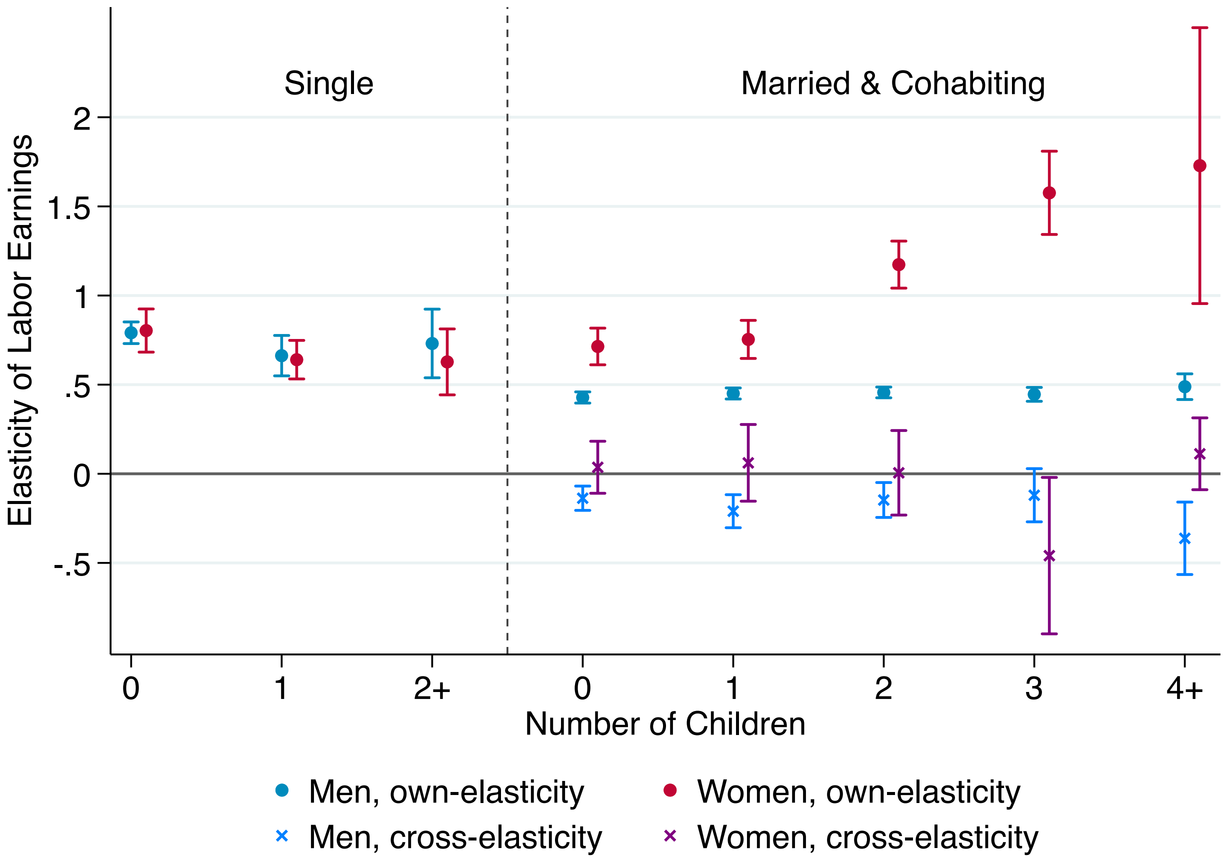


Figure 14: Elasticity of Labor Earnings by Number of Children and Marital Status

Notes: The figure presents estimates of own-elasticities (in circles) and cross-elasticities (in crosses) for men and women depending on marital status and the number of children. Marital status and the number of children are defined as of the previous year. Own-elasticities are estimated using a 2SLS estimation of equation (4), where the dependent variable is log labor earnings. Regressions include match-strata fixed effects, which refer to group fixed effects, where each group is a cell used in coarsened exact matching on age, gender and pre-treatment marital status, the number of children, education, location indicator and percentile of income. Cross-elasticities are estimated using a 2SLS estimation of the following modification of equation (2):

$$y_{it} = bracket_{i,t-1} + \delta_t + \varepsilon^{own} \cdot \log(1 - \tau_{it}) + bracket^{spouse}_{i,t-1} + \varepsilon^{cross} \cdot \log(1 - \tau^{spouse}_{it}) + \boldsymbol{X}'_{it}\gamma + \nu_{it}$$

where the dependent variable is the logarithm of the individual's labor earnings and the two endogenous variables, the individual's log net-of-tax rate and his spouse's log net-of-tax rate, are instrumented with an interaction between indicators of treatment status and tax-free year for the individual and his spouse separately. The coefficient ε^{cross} identifies the cross-elasticity. Estimates by subgroups are obtained by interacting group indicators with the log of the net-of-tax rate of the individual and his spouse as well as the respective instrumental variables. Regressions control for gender, age, education, marital status, and whether living in the capital area or not. The figure shows 95% confidence intervals based on clustered robust standard errors.

Table 1: Tax-Bracket DD: Effect of a Tax-Free Year on Labor Earnings

	(1)	(2)	(3)
2SLS DD estimate ($\frac{d\log y}{d\log(1-\tau)}$)	0.374***	0.330***	0.401***
	(0.024)	(0.024)	(0.032)
Reduced form estimate ($d\log y$)	0.077***	0.069***	0.077***
	(0.005)	(0.005)	(0.006)
First stage estimate ($d\log(1-\tau)$)	0.207***	0.208***	0.193***
	(0.001)	(0.001)	(0.001)
Controls	Yes	Yes	Yes
Occupation Fixed Effects	No	Yes	No
Sector Fixed Effects	No	Yes	No
Matching	No	No	Yes
Observations	526,955	526,955	526,458

Notes: The table presents results from difference-in-differences (DD) regressions, where each row and column entry corresponds to one regression estimate. The top row presents results from a 2SLS estimation of equation (2), where the dependent variable is the logarithm of labor earnings and the net-of-tax rate is instrumented with an interaction between indicators of treatment status and tax-free year. The middle row presents results from a reduced-form DD estimation of equation (1), where the outcome variable is the logarithm of labor earnings. The bottom row presents results from a first-stage DD estimation of equation (1), where the outcome variable is the logarithm of one minus the marginal tax rate. Controls are gender, age, education, marital status, whether living in the capital area or not, and the number of children at age 0-18. Occupation and sector fixed effects are group dummies for occupation and sector groups. "Matching" refers to weighted regressions after coarsened exact matching on age and pre-treatment marital status, the number of children and education. Robust standard errors clustered by individual in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2: Tax-Bracket DD: Effect of a Tax-Free Year on Total Weeks Worked

	(1)	(2)	(3)
2SLS DD estimate ($\frac{dy}{d\log(1-\tau)}$)	4.926***	4.818***	6.549***
	(0.784)	(0.765)	(1.074)
Reduced form estimate (dy)	1.023***	1.006***	1.267***
	(0.162)	(0.159)	(0.207)
First stage estimate ($d\log(1-\tau)$)	0.207***	0.208***	0.193***
	(0.001)	(0.001)	(0.001)
Mean of outcome variable	48.43	48.43	48.43
Controls	Yes	Yes	Yes
Occupation Fixed Effects	No	Yes	No
Sector Fixed Effects	No	Yes	No
Matching	No	No	Yes
Observations	520,438	520,438	519,941

Notes: The table presents results from difference-in-differences (DD) regressions, where each row and column entry corresponds to one regression estimate. The top row presents results from a 2SLS estimation of equation (2), where the dependent variable is the total number of weeks worked and the net-of-tax rate is instrumented with an interaction between indicators of treatment status and tax-free year. The middle row presents results from a reduced-form DD estimation of equation (1), where the outcome variable is the total number of weeks worked. The bottom row presents results from a first-stage DD estimation of equation (1), where the outcome variable is the logarithm of one minus the marginal tax rate. Controls are gender, age, education, marital status, whether living in the capital area or not, and the number of children at age 0-18. Occupation and sector fixed effects are group dummies for occupation and sector groups. "Matching" refers to weighted regressions after coarsened exact matching on age and pre-treatment marital status, number of children and education. Robust standard errors clustered by individual in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Tax-Bracket DD: Effect of a Tax-Free Year on Employment

	(1)	(2)
2SLS DD estimate ($\frac{dP}{d\log(1-\tau^a)}$)	-0.033	0.030
	(0.024)	(0.030)
Reduced form estimate (dP)	-0.004	0.004
	(0.003)	(0.002)
First stage estimate ($d\log(1-\tau^a)$)	0.127***	0.119***
	(0.001)	(0.001)
Mean of outcome variable	0.914	0.914
Controls	Yes	Yes
Matching	No	Yes
Observations	530,900	530,397

Notes: The table presents results from difference-in-differences (DD) regressions, where each row and column entry corresponds to one regression estimate. Employment is defined as earning more than a base income threshold, defined in terms of guaranteed income; see the main text of Section 4 for discussion and details. The top row presents results from a 2SLS estimation of equation (2), where the dependent variable is the total number of weeks worked and the log of net-of-average-tax rate, $(1-\tau^a)$, is instrumented with an interaction between indicators of treatment status and tax-free year. The middle row presents results from a reduced-form DD estimation of equation (1), where the outcome variable is the total number of weeks worked. The bottom row presents results from a first-stage DD estimation of equation (1), where the outcome variable is the logarithm of one minus the marginal tax rate. Controls are gender, age, education, marital status, whether living in the capital area or not, and the number of children at age 0-18. "Matching" refers to weighted regressions after coarsened exact matching on age and pre-treatment marital status, the number of children and education. Robust standard errors clustered by individual in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 4: Life-Cycle-DD: Effect of a Tax-Free Year on Labor Earnings

	(1)	(2)	(3)
2SLS DD estimate ($\frac{d\log y}{d\log(1-\tau)}$)	0.654***	0.655***	0.639***
	(0.016)	(0.016)	(0.016)
Reduced form estimate ($d\log y$)	0.145***	0.145***	0.143***
	(0.003)	(0.003)	(0.003)
First stage estimate ($d\log(1-\tau)$)	0.209***	0.209***	0.209***
	(0.002)	(0.002)	(0.002)
Match-strata fixed effects	Yes	Yes	No
Individual fixed effects	No	No	Yes
Occupation fixed effects	No	Yes	No
Sector fixed effects	No	Yes	No
Number of matched observations	546,434	546,434	542,768

Notes: The table presents results from difference-in-differences (DD) regressions, where each row and column entry corresponds to one regression estimate. The top row presents results from a 2SLS estimation of equation (4), where the dependent variable is the logarithm of labor earnings and the net-of-tax rate is instrumented with an interaction between indicators of treatment status and tax-free year. The middle row presents results from a reduced-form DD estimation of equation (3), where the outcome variable is the logarithm of labor earnings. The bottom row presents results from a first-stage DD estimation of equation (3), where the outcome variable is the logarithm of one minus the marginal tax rate. "Match-strata Fixed Effects" refers to group fixed effects, where each group is a cell used in coarsened exact matching on age, gender and pre-treatment marital status, the number of children, education, location indicator and percentile of income. Occupation and sector fixed effects are group dummies for occupation and sector groups. The number of matched observations corresponds to observations for the treatment group. Robust standard errors clustered at the match-strata level are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 5: Life-Cycle-DD: Effect of a Tax-Free Year on Weeks Worked

	(1)	(2)	(3)
2SLS DD estimate ($\frac{dy}{d\log(1-\tau)}$)	3.014***	2.740***	2.469***
	(0.345)	(0.339)	(0.325)
Reduced form estimate (dy)	0.670***	0.609***	0.555***
	(0.077)	(0.075)	(0.073)
First stage estimate ($d\log(1-\tau)$)	0.209***	0.209***	0.209***
	(0.002)	(0.002)	(0.002)
Mean dependent variable	38.37	38.37	38.37
Match-strata fixed effects	Yes	Yes	No
Individual fixed effects	No	No	Yes
Occupation fixed effects	No	Yes	No
Sector fixed effects	No	Yes	No
Number of matched observations	537,774	537,774	536,369

Notes: The table presents results from difference-in-differences (DD) regressions, where each row and column entry corresponds to one regression estimate. The top row presents results from a 2SLS estimation of equation (4), where the dependent variable is total weeks worked and the net-of-tax rate is instrumented with an interaction between indicators of treatment status and tax-free year. The middle row presents results from a reduced-form DD estimation of equation (3), where the outcome variable is total weeks worked. The bottom row presents results from a first-stage DD estimation of equation (3), where the outcome variable is the logarithm of one minus the marginal tax rate. "Match-strata Fixed Effects" refers to group fixed effects, where each group is a cell used in coarsened exact matching on age, gender and pre-treatment marital status, number of children, education, location indicator and percentile of income. Occupation and sector fixed effects are group dummies for occupation and sector groups. The number of matched observations corresponds to observations for the treatment group. Robust standard errors clustered at the match-strata level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Life-Cycle-DD: Effect of a Tax-Free Year on Employment

	(1)	(2)
2SLS DD estimate ($\frac{dP}{d\log(1-\tau^a)}$)	0.068***	0.058***
	(0.013)	(0.014)
Reduced form estimate (dP)	0.008***	0.006***
	(0.001)	(0.001)
First stage estimate ($d\log(1-\tau^a)$)	0.110***	0.110***
	(0.001)	(0.001)
Mean dependent variable	0.672	0.672
Match-strata fixed effects	Yes	No
Individual fixed effects	No	Yes
Number of matched observations	587,332	586,321

Notes: The table presents results from difference-in-differences (DD) regressions, where each row and column entry corresponds to one regression estimate. The top row presents results from a 2SLS estimation of equation (4), where the dependent variable is total weeks worked and the net-of-tax rate is instrumented with an interaction between indicators of treatment status and tax-free year. The middle row presents results from a reduced-form DD estimation of equation (3), where the outcome variable is total weeks worked. The bottom row presents results from a first-stage DD estimation of equation (3), where the outcome variable is the logarithm of one minus the marginal tax rate. "Match-strata Fixed Effects" refers to group fixed effects, where each group is a cell used in coarsened exact matching on age, gender and pre-treatment marital status, the number of children, education, location indicator and percentile of income. Occupation and sector fixed effects are group dummies for occupation and sector groups. The number of matched observations corresponds to observations for the treatment group. Robust standard errors clustered at the match-strata level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 7: Triple-Differences Estimation: Earnings, Weeks and Employment

	Earnings (1)	Weeks (2)	Employment (3)
2SLS DD estimate	0.431***	2.439***	-0.007
	(0.008)	(0.291)	(0.004)
Reduced form estimate	0.144***	0.816***	-0.002
	(0.003)	(0.098)	(0.001)
First stage estimate	0.335***	0.335***	0.335***
	(0.002)	((0.002))	((0.002))
Mean dependent variable	–	48.85	0.917
Match-strata fixed effects	Yes	Yes	Yes
Number of matched observations	398,033	390,959	401,491

Notes: The table presents results from difference-in-differences (DD) regressions, where each row and column entry corresponds to one regression estimate. The top row presents results from a 2SLS estimation of equation (4), where the dependent variable is the logarithm of labor earnings and the net-of-tax rate is instrumented with an interaction between indicators of treatment status and tax-free year. The middle row presents results from a reduced-form DD estimation of equation (3), where the outcome variable is the logarithm of labor earnings. The bottom row presents results from a first-stage DD estimation of equation (3), where the outcome variable is the logarithm of one minus the marginal tax rate. "Match-strata Fixed Effects" refers to group fixed effects, where each group is a cell used in coarsened exact matching on age, gender and pre-treatment marital status, the number of children, education, location indicator and percentile of income. Occupation and sector fixed effects are group dummies for occupation and sector groups. The robust standard errors clustered at the match-strata level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 8: Heterogeneous Labor Supply Responses by Flexibility of Employment Arrangement

	Temporal flexibility		Constrained in in primary job		Hours flexibility	
	Low (1)	High (2)	Yes (3)	No (4)	Low (5)	High (6)
	A. Labor Earnings					
2SLS DD estimate	0.556***	0.842***	0.451***	0.641***	0.642***	0.946***
	(0.017)	(0.035)	(0.009)	(0.016)	(0.019)	(0.028)
	B. Weeks Worked					
2SLS DD estimate	6.630***	5.368***	4.692***	3.947***	9.936***	1.087*
	(0.493)	(0.873)	(0.319)	(0.500)	(0.639)	(0.599)
Mean weeks pre reform	52.79	36.13	53.05	42.01	53.05	40.73

Notes: The table presents results from a 2SLS estimation of equation (4), where each row and column entry corresponds to one regression estimate. The dependent variable is indicated above each panel. Estimates by subgroups are obtained by interacting group indicators with the log of net-of-tax rate and the instrument in regression (4). *Temporal flexibility* splits the sample by a measure of relative variability in weeks worked within an occupation; see the main text for details. "Low" flexibility refers to workers in the bottom quartile of the distribution over the job flexibility measure, but "High" refers to the top quartile. *"Constrained in primary job"* is an indicator that equals one (*"Yes"*) if working 52 weeks in the primary job in the prior year, but zero (*"No"*) for those working between 26 and 51 weeks in the previous year. *Hours flexibility* splits the sample by occupations based on the share of workers with fixed-salary contracts, where "Low" share refers to occupations where less than 5% of the workers have a fixed salary and "High" share refers to occupations where more than 15% are salaried. All regressions include match-strata fixed effects, which are the cells used in coarsened exact matching on age, gender and pre-treatment marital status, the number of children, education, location indicator and percentile of income. Robust standard errors clustered at the match-strata level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 9: Cross-Elasticities of Earned Income: Husbands and Wives

	All		Children age 0-6		Age		Constrained in primary job	
			0	$\geqslant 1$	$60 <$	$\geqslant 60$	Yes	No
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Husbands				
Cross-elasticity	-0.172***	-0.150***	-0.121***	-0.206***	-0.193***	-0.044	-0.199***	-0.088*
	(0.022)	(0.022)	(0.037)	(0.028)	(0.024)	(0.051)	(0.053)	(0.045)
IHS(spouse income)	–	-0.015***	–	–	–	–	–	–
	–	(0.001)	–	–	–	–	–	–
Observations	223,919	223,919	223,919		223,919		223,919	
				Wives				
Cross-elasticity	0.025	0.014	0.042	0.006	0.014	0.082	-0.184	0.208*
	(0.054)	(0.053)	(0.080)	(0.065)	(0.059)	(0.103)	(0.112)	(0.109)
IHS(spouse income)	–	0.032***	–	–	–	–	–	–
	–	(0.009)	–	–	–	–	–	–
Observations	102,283	102,283	102,283		102,283		102,283	

Notes: The table presents estimates of the earnings elasticity of married and cohabiting individuals to their spouse's net-of-tax rate. Cross-elasticities are estimated using the 2SLS estimation of the following modification of equation (2):

$$y_{it} = bracket_{i,t-1} + \delta_t + \varepsilon^{own} \cdot \log(1 - \tau_{it}) + bracket^{spouse}_{i,t-1} + \varepsilon^{cross} \cdot \log(1 - \tau^{spouse}_{it}) + \boldsymbol{X}'_{it}\gamma + \nu_{it}$$

where the dependent variable is the logarithm of the individual's labor earnings and the two endogenous variables, the individual's log net-of-tax rate and his spouse's log net-of-tax rate, are instrumented with an interaction between indicators of treatment status and the tax-free year for the individual and his spouse separately. The coefficient ε^{cross} identifies the cross-elasticity. Estimates by subgroups are obtained by interacting group indicators with the log of the net-of-tax rate of the individual and his spouse as well as the respective instrumental variables. *"Constrained in primary job"* is an indicator that equals one (*"Yes"*) if working 52 weeks in a primary job pre-reform. All regressions control for age, education, whether living in the capital area or not and the number of children at age 0-18. Column (2) includes the inverse hyperbolic sin (IHS) function of spouse's income, instead of in logs, to account for the possibility of the spouse's income being zero. Robust standard errors clustered by individual in parentheses. *** p<0.01, ** p<0.05, * p<0.1