Essays on Behavioral Responses to Corporate and Personal Income Taxation

Dissertation

zur Erlangung des Doktorgrades

der Wirtschafts- und Sozialwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

vorgelegt von

Sean Mc Auliffe

aus Bad Dürkheim

Tübingen

2023

Tag der mündlichen Prüfung:	15.12.2023
Dekan:	Prof. Dr. Ansgar Thiel
1. Gutachter:	Prof. Dr. Georg Wamser
2. Gutachter:	Prof. Dr. Jörg Baten

Acknowledgements

I would like to thank Georg Wamser and Valeria Merlo for agreeing to supervise my thesis. I am particularly indebted to Georg Wamser for his guidance and much appreciated support. I strongly benefited from our close collaboration, which is reflected in the co-authorship of the Chapters 1 and 2. I am also indebted to Peter Egger and Valeria Merlo, who are co-authors of Chapter 2, and Georg U. Thunecke, who is co-author of Chapter 1. Grateful acknowledgements are made to the School of Business and Economics and the Research School of International Taxation (RSIT) for providing me with an excellent research environment. I also a owe special debt to Peter Egger for inviting me for a productive and engaging three month research stay at the ETH Zürich, where both he and his staff treated me as one of their own. I would also like to thank all my colleagues along the way for the pleasant working atmosphere and the many valuable contributions. Participants at various conferences and workshops helped improve the research with helpful comments and suggestions. I am grateful to the University of Wisconsin-Madison for the provision of the data used in Chapter 3. I am also grateful to the Deutsche Forschungsgemeinschaft for financial support. Finally, I would like to thank everyone who supported me along the way.

Contents

List of Tables

List of Figures

Introduction

This dissertation is a collection of three essays that investigate research questions related to corporate income taxation (Chapters 1 and 2) and personal income taxation (Chapter 3). In the following, the methodologies and key findings of each chapter are briefly discussed. In particular, the data contributions of the Chapters 1 and 2, both of which develop new approaches for computing forward-looking and backward-looking corporate income tax rates, respectively, are highlighted and compared to each other.

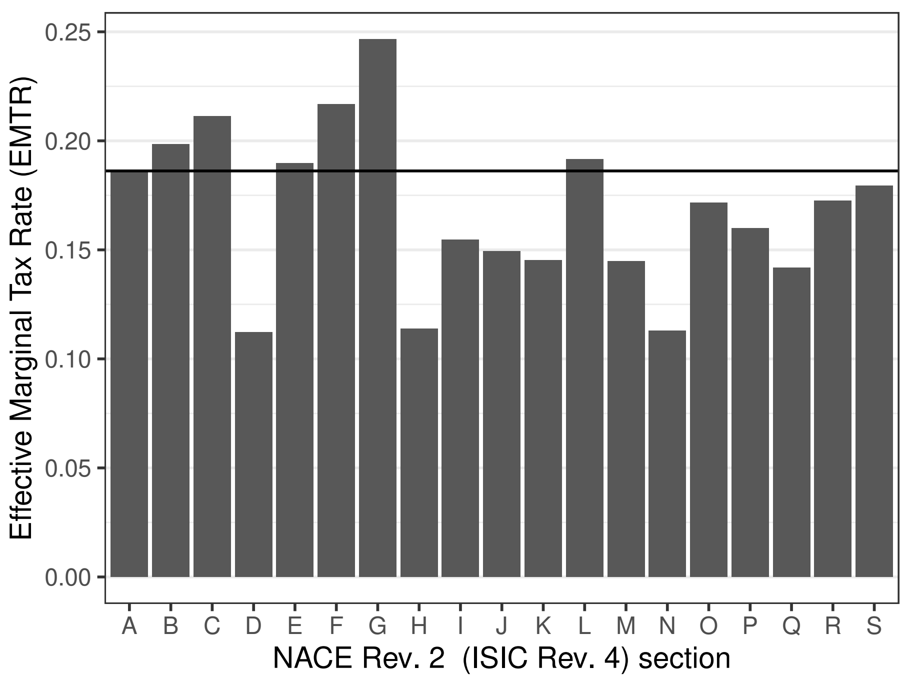
Chapter 1 estimates the tax elasticity of tangible fixed assets, a highly relevant policy parameter that measures the magnitude of investment responses to changes in the tax code. The key innovation lies in the methodology developed to calculate forward-looking effective tax rates (FLETRs), i.e., tax measures that depict the tax burden of hypothetical investment projects by combining statutory tax rates and tax base determinants. Unlike the FLETRs commonly used in the previous literature, the measures computed in Chapter 1 account for typical country-industry-specific financing structures as well as asset structures. Accounting for these characteristics is relevant, as interest payments on debt are generally tax-deductible and depreciation allowances differ between asset categories. Thereby, we ensure that the FLETRs adequately capture the variation in the tax incentives that a country's tax code implicitly grants to different industries. The importance of accounting for these country-industry-specific characteristics is illustrated in Figure 0.1 using the effective marginal tax rate (EMTR), an FLETR depicting the tax burden at the intensive margin (Devereux and Griffith, 1998).[1] In detail, the figure depicts the EMTR for a single country (France) and a single year (2015). Due to the industry-specific financing and asset structures, the EMTR substantially varies between different industries (depicted on the

[1] Besides the EMTR, effective average tax rates (EATRs) that capture the effective tax burden of all inframarginal units invested in a hypothetical investment project (Devereux and Griffith, 2003) are calculated in Chapter 1. Note that since the estimation of the tax-elasticity of tangible fixed assets is identified from intensive margin investments, the focus lies on the EMTR, which is the adequate measure in this context.

horizontal axis). In particular, for most industries, there is a large difference between the magnitude of the industry-specific EMTRs and the conventional EMTR that is calculated using identical financing and asset structures for all industries, depicted by the horizontal line. These discrepancies, which are analyzed in more detail in Chapter 1, suggest that not accounting for country-industry-specific heterogeneity in financing and asset structures introduces a substantial measurement error.

Figure 0.1: ***COUNTRY-INDUSTRY-SPECIFIC EMTRs FOR FRANCE IN 2015***

The figure depicts country-industry-specific effective marginal tax rates (EMTRs) for France and the year 2015. The industries that are depicted on the horizontal axis denote NACE Rev. 2 (ISIC Rev. 4) sections (for definitions of the sections, see Appendix 1.8.3 of Chapter 1). The horizontal line depicts the country-specific EMTR for France and the year 2015. For details on the calculation of the country-industry-year-specific EMTRs as well as the country-year-specific EMTRs, see Chapter 1.



Using the newly calculated FLETRs, we estimate the tax semi-elasticity of tangible fixed assets based on a panel of over 24 million firm-entity observations. The results suggest a statistically significant semi-elasticity with respect to the EMTR of -0.41, which is at the lower end of previous findings (see, e.g., Feld and Heckemeyer, 2011). An interesting – and highly policy relevant – additional result suggests that firm-entities are more responsive to changes in depreciation allowances than to changes in the statutory tax rate. Finally, motivated by various theoretical and empirical contributions to the literature, we estimate the tax semi-elasticity for various groups. The group-specific elasticity estimations confirm the hypotheses derived from the literature. For instance, we find that firm-entities that are either located in a low-tax country and/or are part of an MNE that has access to a low-tax country react less sensitively to tax incentives compared to firm-entities without access to a low-tax jurisdiction. This finding is in line with the profit-shifting literature that shows that large multinational corporations are able to shift profits to entities located in low-tax jurisdictions to avoid taxes, which also makes their firm-entities in high-tax locations less responsive.

Chapter 2 investigates the relationship between effective corporate income taxation and corruption in the EU. To this end, we compute region-industry-year-specific empirical effective income tax rates (EEITRs) using firm-entity-level income statement data. In detail, we define the EEITR as the marginal effect of a one unit increase in the Earnings Before Interest, Taxes, Depreciation, and Amortization (EBITDA) of profitable entities on their total tax liabilities. The estimation of this marginal effect is carried out via instrumental variable regressions using as instrument for the EBITDA of a given entity observation the average EBITDA of the other profitable entities operating in the same country and the same 3-digit industry. Unlike conventional backward-looking empirical tax rates that are typically calculated as the ratio of total tax liability relative to pre-tax profits (see, e.g., Janský, 2023), our novel approach "isolates" the corporate income taxes in the total tax variable by exploiting the fact that the corporate income tax payments are the only taxes that directly vary with and depend on profits.[2] In a next step, we analyze the within country, industry, and year variation in the region-industry-year-specific EEITRs by regressing them on proxies for deductions that could legally be claimed, namely depreciation allowances, deduction

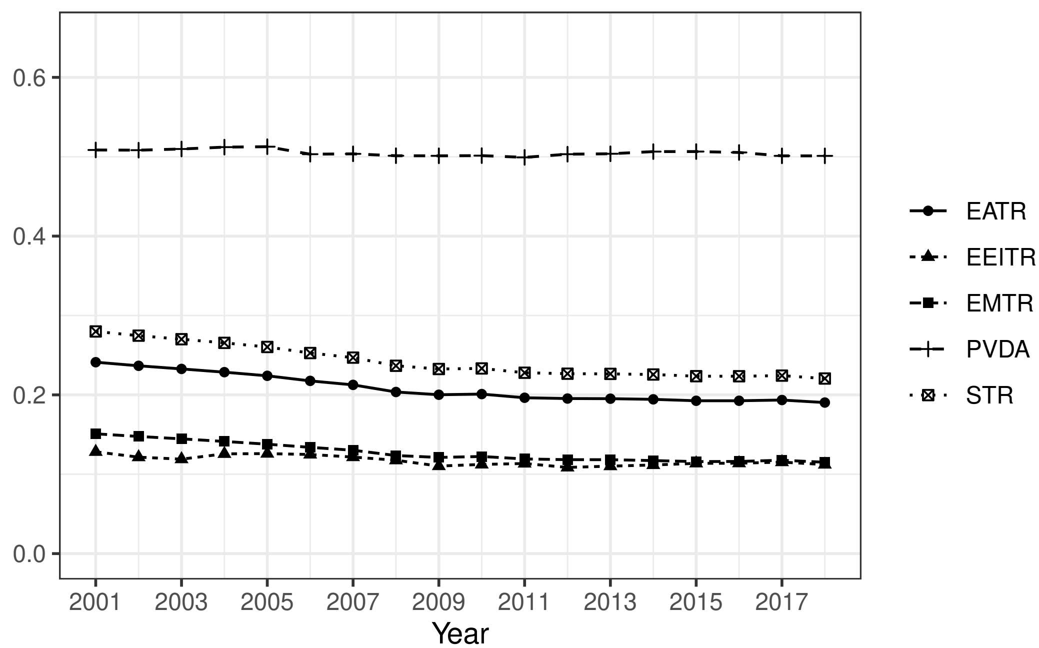
[2]Other taxes that are potentially included in the total tax liability item from entities' income statements include, e.g., carbon taxes and property taxes that have as tax base carbon emissions and property values, respectively.

of interest payments, potential for loss carryforwards, and preferential treatment of patent revenues, as well as additional controls such as regional GDP. In theory, in the absence of tax evasion, tax deductions that can legally be claimed should be the only variables that impact the EEITR. However, additionally controlling for a regional corruption measure, we find that the EEITR is lower in regions where citizens perceive corruption to be comparatively more prevalent. In detail, our findings suggest that a one standard deviation increase in corruption leads to a statistically significant decrease in EEITRs of approximately 0.4 percentage points. From an economic point of view, this effect is sizeable given that several countries in our sample exhibit between regions differences in corruption of more than one standard deviation. Since our EEITRs are based on the EBITDA from publicly available income statement data that therefore should not deviate from the figures reported to the tax authorities, our results suggest that the systematically lower EEITRs in high-corruption regions are likely the result of tax evasion via overstated deductions. Regarding this tax evasion channel, our findings differ from the existing literature on corporate tax evasion, which focuses on underreporting of profit or sales figures to tax authorities, but not on how the reported figures are transformed into the final tax base via deductions (see, e.g., Alm et al., 2016).

While both Chapters 1 and 2 present research related to corporate income taxation, the tax rates developed in the chapters fundamentally differ in nature, as they are tailored to the respective research questions. Using the EEITRs developed in Chapter 2 to analyze investment responses as done in Chapter 1 would be unsuitable, as empirical tax rates also capture tax avoidance and tax evasion, which are themselves responses to the statutory tax code. In the context of Chapter 2, however, we are interested in such a tax measure that captures these response margins, as this is precisely the firm behavior we want to study. Figure 0.2 provides a visual comparison between the forward-looking tax rates from Chapter 1 and the backward-looking EEITRs from Chapter 2 by plotting yearly averages across all

Figure 0.2: ***COMPARISON OF DIFFERENT TAX MEASURES OVER THE YEARS, ALL INDUSTRIES***

The figure depicts unweighted yearly means of different country-industry-year-specific tax measures, as well as of the statutory tax rate (STR). EATR denotes the country-industry-year-specific effective average tax rate. EMTR denotes the country-industry-year-specific effective marginal tax rate. PVDA denotes the country-industry-year-specific net present value of tax depreciation allowances per unit of investment. EEITR denotes the country-industry-year-specific empirical effective income tax rate. The total number of observations for each tax measure, except for the EEITR, is 66,956, spanning across 221 countries and 19 industries. The total number of observations for the EEITR is 7,194, spanning across 42 countries and 19 industries. The industries denote the NACE Rev. 2 (ISIC Rev. 4) sections. The two sections that are not included for all tax measures are T and U (for definitions of the sections, see Appendix 1.8.3 of Chapter 1). For details on the calculation of all measures except for the EEITR, see Chapter 1. For details on the calculation of the EEITR, see Chapter 2 (note that the calculation is identical to the one of the region-industry-specific EEITRs used in Chapter 1).



available observations for each measure.[3] In addition, we plot the statutory tax rate as well as the country-industry-year-specific net present value (NPV) of tax depreciation allowances per unit of investment. These two measures are the key components of both the EMTR and the EATR. The time series depicted in Figure 0.2 suggest that in the aggregate, the NPV of depreciation allowances remains fairly constant over time. The statutory tax rate, on the other hand, declines by several percentage points during the depicted time span, which is also reflected in the downward movement of both the EMTR and the EATR. Note that the level difference between the EMTR and EATR is due to mechanical differences in the computation of these measures.[4] The most interesting finding in Figure 0.2 is the fact that the EEITR seems to converge to the EMTR from below during the depicted time span. The EEITR being roughly comparable to the EMTR in magnitude, however, is expected, as the EEITR is itself defined as a marginal tax measure (see above). A caveat regarding the comparison of the measures from the different chapters is the fact that the time series averages are based on varying numbers of observations, with the forward-looking measures being based on 221 countries, while the EEITR is only available for 42 countries. Therefore, to allow for a meaningful comparison, we reduce the sample to country-industry combinations for which all measures are available for two selected industries and all years between 2001 to 2018, reducing the sample to 13, primarily European, countries.[5] The time series corresponding to this sample are shown in Figure 0.3. It shows that most of the patterns discussed above – in particular also the convergence between EEITR and EMTR – hold for both depicted industries. However, since these findings are merely descriptive, we refrain from speculating about the causes of the convergence pattern.[6] Instead, this is left for future research.

[3] Note that the EEITRs in Chapter 2 are calculated at the regional level. To allow for a sensible comparison to the measures from Chapter 1, the depicted EEITRs Figure 0.2 are calculated using the same approach, but at the country level.

[4] Devereux and Griffith (2003) show that the EATR equals a weighted average of the EMTR and an adjusted statutory tax rate.

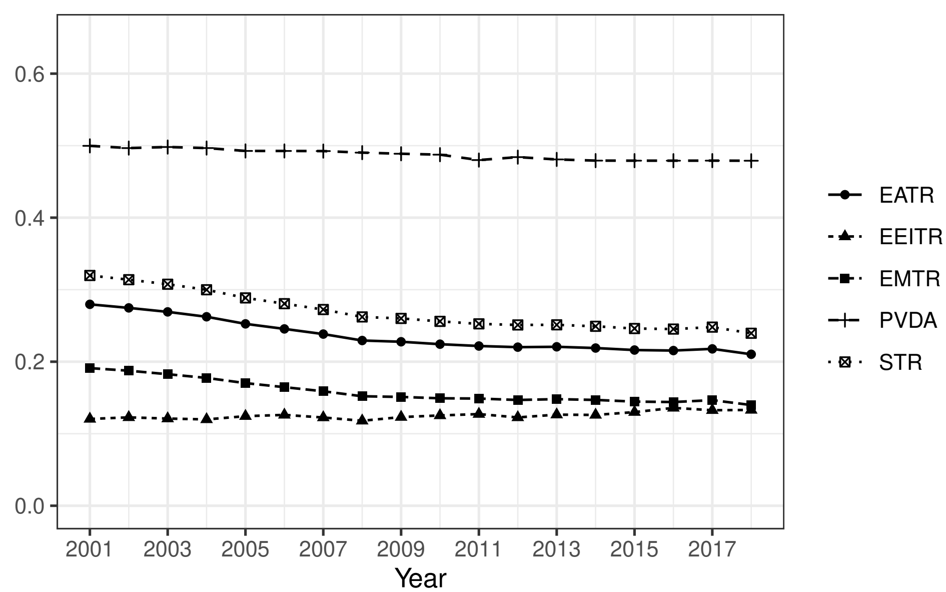
[5] The selected industries, Manufacturing and Construction, are chosen for the sole reason that they exhibit good data coverage.

[6] Note also that despite the sample used to plot the time series in Figure 0.3 being perfectly balanced, variation over time in the EEITR may also be caused by yearly changes in the sample of firm-entities used to calculate the EEITRs.

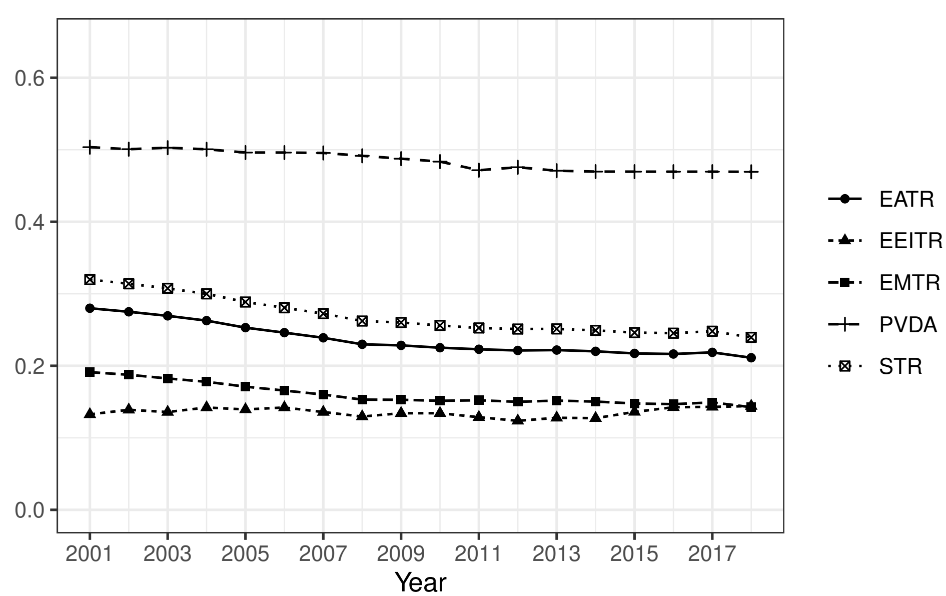
Figure 0.3: *COMPARISON OF DIFFERENT TAX MEASURES OVER THE YEARS, SELECTED INDUSTRIES*

The figure depicts unweighted yearly means of different country-industry-year-specific tax measures, as well as of the statutory tax rate (STR). EATR denotes the country-industry-year-specific effective average tax rate. EMTR denotes the country-industry-year-specific effective marginal tax rate. PVDA denotes the country-industry-year-specific net present value of tax depreciation allowances per unit of investment. EEITR denotes the country-industry-year-specific empirical effective income tax rate. Panel A depicts the tax measures for the NACE Rev. 2 (ISIC Rev. 4) section *C Manufacturing*. Panel A depicts the tax measures for the NACE Rev. 2 (ISIC Rev. 4) section *F Construction*. Both panels are based on perfectly balanced panels comprising the following 13 countries (denoted by ISO 3 codes): BEL, CZE, ESP, FIN, FRA, GBR, GRC, HRV, ITA, KOR, NLD, SRB, SWE. For details on the calculation of all measures except for the EEITR, see Chapter 1. For details on the calculation of the EEITR, see Chapter 2 (note that the calculation is identical to the one of the region-industry-specific EEITRs used in Chapter 1).

Panel A: Manufacturing



Panel B: Construction



The final Chapter 3 contrasts to the previous two chapters in that it does not relate to corporate income taxation but instead focuses on personal income taxation. More precisely, the chapter presents a welfare analysis of the Negative Income Tax (NIT), a transfer scheme that aims at reducing poverty. It provides households without any income with a certain transfer. However, this transfer declines linearly in family income at the take-back rate up until a break-even point where the transfer becomes zero (see, e.g., Saez, 2002). In our analysis, we conduct a welfare analysis to determine the optimal level of the take-back rate. The setup of the welfare analysis builds on the contribution by Kasy (2018) who illustrates the importance of allowing for nonlinearities in key behavioral response margins – in our case the labor supply – across different policy levels. In this regard, the implemented approach differs from conventional welfare analysis that often relies on "sufficient statistics", i.e., single parameter estimates, typically elasticities, that describe key behavioral relationships (Chetty, 2009).[7] In a first step of the analysis, a notion of social welfare is theoretically derived that takes into account the families' allocation decision regarding their disposable time which can either be used for work (which earns labor income that is used for consumption, but also reduces the NIT transfer) or leisure. Using Cobb-Douglas utility functions, our model predicts a negative and nonlinear relationship between the take-back rate and labor supply in terms of hours worked. Using data from two NIT experiments conducted in the US in the 1970s, we estimate the labor supply as a function of the take-back rate allowing for nonlinearities. We find that the empirical labor supply function for the most part matches the theoretical prediction. We then plug the labor supply function into our social welfare formula to determine the welfare optimizing take-back rate, which lies – depending on the precise parameterization – between 65% and 69%. However, due to issues regarding the data quality, the validity of this finding is likely low. A perhaps more interesting finding is the fact that the social welfare optimizing take-back rate differs substantially depending on whether one allows for nonlinearities in the labor supply estimation or not, suggesting

[7] In the context of the NIT, such sufficient statistics welfare formulas can be found in, e.g., Saez (2002).

that simple notions of social welfare that do not account for nonlinear responses, such as the aforementioned sufficient statistics approach, potentially yield very imprecise results.

References

Alm, J., Martinez-Vazquez, J., & McClellan, C. (2016). Corruption and firm tax evasion. *Journal of Economic Behavior & Organization*, 124, 146-163.

Chetty, R. (2009). Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annual Review of Economics*, 1(1), 451-488.

Devereux, M. P., & Griffith, R. (1998). The taxation of discrete investment choices. IFS working papers No. W98/16.

Devereux, M. P., & Griffith, R. (2003). Evaluating tax policy for location decisions. *International Tax and Public Finance*, 10(2), 107-126.

Feld, L. P., & Heckemeyer, J. H. (2011). FDI and taxation: A meta-study. *Journal of Economic Surveys*, 25(2), 233-272.

Janský, P. (2023). Corporate effective tax rates for research and policy. *Public Finance Review*, 51(2), 171-205.

Kasy, M. (2018). Optimal taxation and insurance using machine learning–Sufficient statistics and beyond. *Journal of Public Economics*, 167, 205-219.

Saez, E. (2002). Optimal income transfer programs: intensive versus extensive labor supply responses. *The Quarterly Journal of Economics*, 117(3), 1039-1073.

1. The Tax-Elasticity of Tangible Fixed Assets: Evidence from Novel Corporate Tax Data

Abstract* –This paper develops a new approach to calculate country-industry-year-specific forward-looking effective tax rates (FLETRs) based on a panel of 19 industries, 221 countries, and the years 2001 to 2020. Besides statutory corporate tax rate and tax base determinants, the FLETRs account for typical country-industry-specific financing structures as well as asset compositions. We show that FLETRs suffer from significant measurement error when the latter information is neglected, owing primarily to inappropriately assigned asset weights to statutory depreciation allowances. Our empirical analysis exploits the substantial variation in FLETRs over time to provide estimates of the tax semi-elasticity of corporate investment in tangible fixed assets. Based on more than 24 million firm-entity observations, our results suggest a statistically significant tax semi-elasticity of -0.41, which is at the lower end of previous findings. We further show that different subgroups of firms respond very heterogeneously to tax incentives.

*This chapter is based on joint work with Georg U. Thunecke and Georg Wamser. The corresponding paper is available as CESifo working paper No. 10628.

1.1. Introduction

The main objective of fundamental tax reforms such as the 2017 US "Tax Cuts and Jobs Act" is to increase economic outcomes such as employment and economic growth. The success of a reform in this regard depends in large parts on the extent to which reductions in the tax burden stimulate firms' real investments. A central measure of the relationship between corporate taxation and investment is the tax-elasticity of corporate investment. Despite its policy relevance, there is disagreement as to the responsiveness of real assets to changes in corporate taxation, which is underlined by the heterogeneity in previous estimates. In designing tax reforms, policymakers often rely on two main policy instruments to stimulate investment: statutory tax rate cuts and generous bonus depreciation rules for specific assets.[1] This paper contributes to the literature by providing estimates on the tax-elasticity of firms' tangible fixed assets that are based on a broad panel of novel forward-looking corporate effective tax rates (FLETRs), capturing changes in statutory tax rates as well as depreciation allowances.

The aim of FLETRs is to capture incentives of the corporate tax code (statutory tax rate and tax base determinants) by depicting the tax burden of a hypothetical investment project, which makes FLETRs particularly suitable for the analysis of investment responses (Sørensen, 2004).[2] The key distinction from the previous literature regarding the way we calculate FLETRs is that we account for typical country-industry-specific financing and asset structures. Accounting for these country-industry-specific characteristics plays an important role in determining the magnitude of the FLETRs, as interest payments on debt are generally tax-deductible and depreciation allowances differ between asset categories. In other words,

[1] Earlier policy reforms in the US, such as the 2005 Domestic Production Activities Deduction (DPAD), provided substantial corporate tax provisions, affecting firms' tax bases (see Ohrn, 2018).

[2] Note that quantifying the tax burden using forward-looking measures goes back to the seminal contribution of King and Fullerton (1984) and was substantially advanced by Devereux and Griffith (1998a) as well as Devereux and Griffith (2003). So-called backward-looking measures (calculated as taxes paid relative to pre-tax profit), in contrast, not only fail to capture current and future investment incentives but are also prone to severe endogeneity concerns as both taxes paid and pre-tax profit may be driven by tax-planning decisions of a firm (Devereux and Griffith, 2002). For an extensive overview over different backward-looking measures, see Janský (2023).

using country-industry-level financing and asset weights allows us to correctly capture the general heterogeneity in investment incentives that a country's tax code implicitly offers to the different industries. At the same time, as our FLETRs rely on typical country-industry-specific weights, we ensure that tax measures are exogenous and primarily capture incentive effects from tax law. Most previous studies do not account for country-industry heterogeneity, but instead calculate FLETRs using identical financing and asset structures for all countries and industries, often due to a lack of adequate data.[3] Our analysis shows that disregarding country-industry-level heterogeneity for calculating FLETRs leads to a systematic measurement error.

With respect to the country-industry-specific financing structures, we aggregate firm-entity level data from Bureau van Dijk's *Orbis* database. For the calculation of the country-industry-specific asset structures, we distinguish a total of seven different asset types for which we derive weights from the *EUKLEMS & INTANProd* database and *Orbis*. If these data sources are not available for country-industry combinations or do not provide sufficient information, we impute the financing and asset structures using *Predictive Mean Matching (PMM)*, which we run on extensive sets of both country- and industry-specific matching covariates. Combining the country-industry-specific financing and asset structures with statutory information on the tax rates and depreciation regimes yields FLETRs for virtually the entire world and all industries. More precisely, our almost perfectly balanced panel includes FLETRs for 221 countries,[4] 19 industries, and 20 years (2001 to 2020). For the time

[3]See, e.g., Da Rin et al. (2010), Devereux and Griffith (1998b), Egger et al. (2014), Overesch and Rincke (2009), Spengel et al. (2016a, 2016b, 2016c), Steinmüller et al. (2019). Note that many contributions use the constant financing and asset structures proposed in the seminal publication *Taxing Profits in a Global Economy: Domestic and International Issues* by the OECD from 1991. The Centre for European Economic Research (ZEW), the University of Mannheim, and PricewaterhouseCoopers (Spengel et al., 2021) compiled a comprehensive collection of effective tax rates for 35 (mainly EU) countries. Notable studies that are using varying weighting schemes are Egger et al. (2009) as well as Egger and Loretz (2010), who use *Orbis* to calculate firm-year-specific financing and asset structures, and Steinmüller et al. (2019), who use a combination firm-specific financing structures based on *Orbis* and industry-specific asset structures that are equal for all countries. Using time-varying and/or firm-level financing and asset weights would lead to endogenous FLETRs, though, as such structures capture endogenous responses to tax incentives.

[4]Our dataset primarily comprises UN member states but also non-member states, e.g., Taiwan, as well as self-governing territories that formally are part of other states, e.g., Greenland. For the sake of clarity and simplicity, we shall henceforth refer to all included tax-jurisdictions as "countries".

span covered by our panel, we observe a total of close to 700 reforms of national tax codes, i.e., changes to the statutory tax rate and/or depreciation allowances. To the best of our knowledge, while the approach we suggest for determining the industry-specific FLETRs is novel, this is by far the largest dataset of forward-looking tax measures that has ever been calculated (and made available) in terms of country and time coverage.

We exploit the substantial variation in our FLETRs to estimate tax semi-elasticities of firms' tangible fixed assets using the *Orbis* dataset with over 24 million firm-entity-level observations. Our preferred model yields a statistically significant tax semi-elasticity of -0.41, which is at the lower end of previous findings. We also show that our main result – a negative elasticity that is small in magnitude – is highly robust to alternative model specifications. Focusing on depreciation policy, we find that more generous depreciation allowances are even more effective than tax cuts in terms of boosting firm-level investment. We additionally conduct a set of group-specific tax-elasticity estimations. This part of the paper is motivated by different strands of both theoretical and empirical contributions to the literature suggesting that certain groups of firms respond particularly (in-)sensitively to changes in tax incentives. For instance, the tax-sensitivity of investment should depend on the degree to which a firm is financially constrained.[5]

While our group-specific results are consistently in line with the hypotheses derived from the literature, let us highlight one interesting finding, illustrating that firm entities' tax-sensitivity is highly correlated with firms' ability to avoid taxes through relocating profits to low-tax countries. The idea is to examine how firm entities respond to changes in the marginal tax rate given the group-wide (or multinational-firm-wide) minimum statutory corporate tax rate.[6] The reasoning behind this exercise is the following: firms that can access low-tax or tax haven countries may be able to shift profits there, and are thus less responsive

[5]See, e.g., Keuschnigg and Ribi (2013).

[6]That is, the minimum tax rate that they are exposed to within their firm group. This minimum tax rate is defined as the group-wide (or multinational-firm-wide) minimum statutory tax rate or, in the case of stand-alone entities, the tax rate of the country they are located in. We then group all observations according to the respective minimum tax rate.

to tax incentives at the other locations. While we generally find negative elasticities, the tax response becomes steadily stronger for those groups where the firm-group-specific minimum tax rate is high. In fact, we find that only after a threshold of a minimum tax rate of 24%, the tax-elasticities turn statistically significant. This finding allows for an interpretation in light of the profit-shifting literature, as large multinational corporations are able to shift profits to entities located in low-tax jurisdictions to avoid taxes. The latter makes these firms less responsive to tax incentives at all other locations.

Besides the new approach of calculating FLETRs, this paper adds to the literature in several ways. We primarily contribute to previous research on investment responses to tax incentives. Some recent estimates suggest quite substantial tax effects in this context. For example, Ohrn (2018) shows that a 1 percentage point reduction in firms' effective corporate tax rate (through additional tax-base deductions) is associated with a 4.7 percent increase in installed capital. Based on a large sample of US firms, Zwick and Mahon (2017) investigate the impact of temporary bonus depreciation rules on firms' investments, distinguishing between eligible and non-eligible capital and industries. They find a substantial increase of investment into eligible equipment. A seminal empirical paper quantifying investment responses to taxes at the level of firms is that of Chirinko et al. (1999). This study suggests a user cost of capital elasticity of about -0.25. Earlier work of Cummins et al. (1994, 1995, 1996) exploits tax reforms to learn about the consequences of changes in the user cost of capital. Their findings indicate substantial investment effects of tax policy. Using German data, Harhoff and Ramb (2001) estimate a long-run user cost of capital elasticity of -0.56.[7] Our

[7] There is also a large literature that uses international investment data (foreign direct investments) to identify tax effects from country-year-specific variation in taxes. De Mooij and Ederveen (2003) perform a meta-analysis to estimate the tax-elasticity of corporate investment. They find substantial heterogeneity in elasticities across studies with a median value of -3.3. Similarly, De Mooij and Ederveen (2008) illustrate that corporate taxation has a substantial impact on the choice of the legal form, financing structure, profit shifting, as well as (intensive and extensive margin) investment decisions. Additionally, they demonstrate that the tax elasticities along these different decision margins vary substantially. Feld and Heckemeyer (2011) conduct a meta-analysis and estimate a median tax semi-elasticity of corporate investment of -2.49 and illustrate that employing firm-level data and (country-specific) effective tax measures yields more accurate estimations of the semi-elasticity. In contrast to this literature, our novel FLETR data allow us to adequately capture and exploit variation within and across different industries and countries in a unified estimation and data context.

semi-elasticity estimates, based on our large sample including about 24 million observations, are substantially smaller.

Some previous contributions study investment responses and specific channels through which investment is affected. For example, Chaney et al. (2012) analyze the effect of real estate price shocks that affect the collateral value of firms, which is a significant driver of investment. Edgerton (2010) accounts for financing constraints as well as loss carrybacks and carryforwards leading to asymmetries in tax responses. Early work of Fazzari et al. (1987) highlights that investment responses depend on the extent to which firms are financially constrained. As mentioned above, in Section 1.6.5, we add to these findings by providing group-specific estimates on the tax-elasticity of the tangible fixes assets. While this paper is not primarily interested in one particular heterogeneity, it illustrates that our approach of measuring tax incentives is very consistent with how we expect taxes to affect heterogeneous firms' decisions.

The remainder of the paper is structured as follows. Section 1.2 derives the country-industry-year-specific FLETR measure. Section 1.3 describes the various data sources that are used for the calculation of the FLETRs and the estimation of the tax-elasticities of investment. The calculation of country-industry-specific financing and asset structures is detailed in Section 1.4. Section 1.5 describes the country-industry-year-specific FLETRs. The tax semi-elasticities of investment are estimated in Section 1.6. Finally, Section 1.7 concludes and presents policy implications.

1.2. Forward-looking Effective Tax Measures

For the empirical estimation of the tax semi-elasticity of investment, it is crucial to model the incentive effects of the corporate tax code in an adequate way. The literature recommends the use of forward-looking measures in such a context, as the incentive to invest depends on current and expected taxation (Sørensen, 2004). In this paper, we distinguish two different kinds of FLETRs: the effective marginal tax rate (EMTR) and the effective average tax rate

(EATR). The EMTR captures incentives of the tax code at the intensive margin – i.e., the tax burden a firm would face on a marginal investment that just breaks even. This property makes it particularly suitable for the calculation of tax semi-elasticities. The EATR, on the other hand, depicts the effective tax burden of all infra-marginal units invested. It is primarily used for the analysis of discrete investment choices, such as location decisions (Devereux and Griffith, 2003). Consequently, the EATR plays only a minor role in our analysis. Since our estimation of the tax semi-elasticities is based primarily on the EMTR, this section focuses on providing some intuition for this measure and the role that country-industry-specific financing and asset structures play for its computation. A brief discussion of the EATR is provided in Appendix 1.8.1.

The theoretical framework of the EMTR is developed in the seminal contributions by Devereux and Griffith (1998a), Hall and Jorgenson (1967), King (1974), King and Fullerton (1984), and OECD (1991). A simple formal representation of the EMTR is

$$EMTR = \frac{(\tau - \tau\delta)}{(1 - \tau\delta)}, \tag{1.1}$$

where τ, $\tau \in [0, 1]$, and δ, $\delta \in [0, 1]$, denote the statutory tax rate and the net present value (NPV) of depreciation allowances, respectively. A detailed derivation of this formula is provided in Appendix 1.8.2. Note that one goal of the simple representation in (1.1) is that it measures tax incentives in a tractable way and allows us to observe all statutory tax code determinants and incentives for as many countries as possible.[8]

From (1.1), it is evident that the marginal investment is not affected (i.e., $EMTR = 0$) if $\delta = 1$, i.e., in the case where the tax law allows a firm to immediately deduct the full purchase price of an asset (e.g., a machine) for tax purposes. The EMTR may even

[8] Calculating FLETRs involves a number of trade-offs. On the one hand, our goal is to capture incentive effects from tax law in a very detailed way, so we aim at including both the tax rate and tax base determinants. On the other hand, the more details we model, the more assumptions we need to accept. We believe that the parsimonious EMTR shown in (1.1) is ideal for the purpose of this paper, as it accounts for the most important tax code information and heterogeneity in tax-base effects. In fact, the major advantage we see is that, with some assumptions, we are able to observe all parameters that allow us to calculate adequate EMTRs.

become negative, which means that the tax system effectively subsidizes investments, e.g., when governments allow for generous deductions and allowances (including interest) such as investment tax credits or bonus depreciation (Zwick and Mahon, 2017). Given the way we calculate the $EMTR$, however, this is not possible and the minimum $EMTR$ is bounded at zero ($EMTR = 0$ if $\delta = 1$ or $\tau = 0$, or both).

For the sake of illustration, let us look at a specific example for the EMTR. In 2010, France levied a corporate tax rate of 34.4%, and granted a NPV of depreciation allowance for equity financed machinery of 0.81. Plugging these values into equation (1.1) yields an EMTR of 9.1% (for marginal investments in machinery).[9] Our goal, however, is not to calculate FLETRs of investments in a single asset type that are purely equity financed. Instead, our goal is to depict the tax burden of country-industry-typical investments in different asset categories that are financed using a combination of equity and debt.[10] In total, we distinguish between seven asset categories: *Buildings, Machinery, Office equipment, Computer equipment, Intangible fixed assets, Vehicles* and *Inventory.* The distinction of different asset categories is important, as different assets are subject to varying depreciation allowances, e.g., buildings depreciate over a substantially longer time period than computer equipment. The distinction between equity and debt financing is relevant as interest payments on debt are usually tax deductible, which results in higher NPVs of depreciation allowances for debt financing compared to financing through retained earnings.[11] Let us denote the NPV of depreciation allowances per unit of investment in asset type a in country c in year t by A^E_{act} and A^D_{act}, with the superscripts E and D indicating financing through retained earnings and debt, respectively.[12] It is important to note that the NPVs of depreciation allowances are purely determined by national tax codes, and tax law applies equally to all industries in a

[9]The corresponding EATR equals 27.4%.

[10]Note that using time-constant rather than time-varying financing and asset weights for the empirical analysis of investment avoids endogeneity issues that may arise due to changes in the financing and asset structures in response to changes in the tax code. Assuming that these weights are constant over a relatively short period of time seems to be an acceptable assumption.

[11]Note that we disregard the possibility of issuing new equity.

[12]Note that since we disregard inflation, inventories are not depreciable, i.e., $A^E_{invent,ct} = A^D_{invent,ct} = 0 \; \forall \; c, t$ (see Hanappi, 2018).

country. Hence, the only reason why different industries located in the same country have a different overall NPV of depreciation allowances is that they use different financing and asset compositions when carrying out investment projects. This is reflected in the formal depiction of the country-industry-year-specific NPV:

$$\delta_{cit} = ES_{ci} \sum_{a=1}^{7} w_{aci} \cdot A_{act}^{E} + DS_{ci} \sum_{a=1}^{7} w_{aci} \cdot A_{act}^{D}, \tag{1.2}$$

where w_{aci} denotes the share of asset a in a typical investment carried out in industry i in country c.[13] ES_{ci} and DS_{ci} denote the country-industry-specific shares of retained earnings and debt used to finance the investment, respectively, which add up to unity. The procedures to obtain w_{aci} as well as ES_{ci} and DS_{ci} are explained in greater detail in Section 1.4.

Finally, using (1.2), we obtain country-industry-year-specific EMTRs:

$$EMTR_{cit} = \frac{(\tau_{ct} - \tau_{ct}\delta_{cit})}{(1 - \tau_{ct}\delta_{cit})}. \tag{1.3}$$

Expression (1.3) is used to calculate the EMTRs we present in Section 1.5.

1.3. Data

Throughout this paper, we use data from a total of five different databases to (i) calculate and impute country-industry-specific financing and asset weights (see Section 1.4); (ii) calculate FLETRs (see Section 1.5); and to (iii) estimate tax semi-elasticities of investment (see Section 1.6). In the following, we briefly describe the databases, which data they provide, and how we use the data for our purposes.

1.3.1. RSIT International Tax Institutions (ITI) Database

The statutory corporate tax regime data that we use to calculate FLETRs is taken from the Research School of International Taxation's (RSIT) *International Tax Institutions (ITI)*

[13] The sum of the asset weights equals one for each country-industry pair, i.e., $\sum_{a=1}^{7} w_{aci} = 1$.

database (Wamser et al., 2023). More precisely, we use the data on statutory corporate tax rates (τ_{ct}) as well as NPVs of depreciation allowances for the six asset categories *Buildings, Machinery, Office equipment, Computer equipment, Intangible fixed assets,* and *Vehicles.* This panel includes 3,954 year-specific data points that span over a total of 221 countries over the years 2001 to 2020.[14] Additionally, we obtain the count variable of the number of double taxation treaties that a country has in a given year ($NDTT_{ct}$), which serves as a control variable in the estimation of the tax semi-elasticity of investment (see Section 1.6).

1.3.2. EUKLEMS & INTANProd

The country-industry-specific asset weights that we derive in this paper are – with the exception of the asset type inventory – based on the 2021 release of the *EUKLEMS & INTANProd* database provided by the Luiss Lab of European Economics. For our purpose, we use the net capital stocks at current replacement costs in million units of the respective national currency that the database provides at the NACE Rev. 2 (ISIC Rev. 4) section level.[15] In detail, we use the capital stock variables for *Dwellings, Other buildings and structures, Computer hardware, Research and development, Computer software and databases, Other machinery equipment and weapon systems, Telecommunications equipment*, and *Transport equipment*. Note that these asset categories, which are based on the European System of Accounts (ESA) 2010, do not directly match the ones for which the *ITI* database provides the depreciation allowances which we use for the calculation of the FLETRs. In Section 1.4.2, we therefore have to regroup the ESA-2010-based variables from *EUKLEMS & INTANProd* to obtain industry-specific net stock values of all considered asset categories for 18 EU countries, as well as for the UK, Japan, and the US.[16] The data coverage ranges from 1995 to 2019, though 2019 is scarcely covered. With the exception of Japan and the US,

[14] For a detailed description of the dataset and data sources, see Wamser et al. (2023).

[15] For the descriptions of all sections, see Appendix 1.8.3. Note that since NACE Rev. 2 was created based on ISIC Rev. 4, these two classification systems are identical at the section level.

[16] Note that, since in Section 1.4.2 we obtain asset weights by summing up all capital stock variables and then taking shares, we only consider observations for which all variables are non-missing.

data for 19 NACE Rev. 2 (ISIC Rev. 4) sections are provided.[17]

1.3.3. Orbis Dataset

For our firm-level analysis of the tax semi-elasticity of investment in Section 1.6, we use Bureau van Dijk's *Orbis* database. *Orbis* contains yearly balance sheet and income statement data as well as general information on the firm entities, such as industry affiliation, year of incorporation, and ownership structure. The definition of variables and a more detailed description of the data for the purpose of the investment elasticity estimation is provided in Section 1.6.2.

We further use *Orbis* information to obtain the financing structure and the weight of the asset type inventory at the country-industry level, both of which we need for the calculation of the FLETRs. More precisely, for the calculation of the financing structure (see Section 1.4.1), we use the variables non-current liabilities ($NCLI_{jt}$, with j and t denoting firm entity and year, respectively) and total assets ($TOAS_{jt}$). The calculation of the inventory shares is based on the stocks of current assets (i.e., inventories) (INV_{jt}), tangible fixed assets ($TFAS_{jt}$), and intangible fixed assets ($IFAS_{jt}$),

1.3.4. World Development Indicators and Worldwide Governance Indicators

In our analysis of the tax semi-elasticity of investment (see Section 1.6), we condition on a number of country-level factors that possibly influence investment behavior. Furthermore, we feed the matching algorithm for the imputation of missing financing and asset weights with country-level variables (Section 1.4.3). Our sources for the country-level controls are the World Bank's *World Development Indicators (WDI)* and *Worldwide Governance Indicators (WGI)* databases.

From the *WDI* database, we obtain the GDP measures GDP in constant PPP US$

[17] Note that the sections T and U are not covered in the *EUKLEMS & INTANProd* database. For Japan, additionally the sections M and N are not covered and for the US the sections *D*, *E*, and *O*. For the descriptions of all sections, see Appendix 1.8.3.

(GDP_{ct}), GDP per capita in constant PPP US\$ ($GDP\ p.c._{ct}$), and GDP growth ($GDP\ growth_{ct}$). Additional variables taken from the *WDI* data are inflation ($Inflation_{ct}$), domestic credit to the private sector in percent of a country's GDP ($DCPS_{ct}$), and the real interest rate ($Real\ interest\ rate_{ct}$).[18]

From the *WGI* database, we use the *Rule of Law* indicator (ROL_{ct}), which captures "perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence", as well as a *Control of Corruption* variable ($Corruption_{ct}$) measuring "the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as 'capture' of the state by elites and private interests" (Kaufmann et al., 2011, p. 223).[19] Note that both measures are varying in an interval of -2.5 to 2.5. The Worldwide Governance indicators are defined such that a higher value corresponds to better governance, i.e., a higher value of $Corruption_{ct}$ indicates less corruption (Kaufmann et al., 2011).

1.3.5. Eora Global Supply Chain Database

Finally, we obtain a set of industry-specific variables from the *Eora26* database, which is part of the *Eora Global Supply Chain* database (Lenzen et al., 2012; Lenzen et al., 2013). These variables are solely used for imputing financing and asset weights (see Section 1.4.3). Note that the *Eora26* industry classification system is different from the NACE Rev. 2 (ISIC Rev. 4) classification that we use throughout this paper. Appendix 1.8.4 provides detailed information on how we convert the *Eora26* classification along with general information on the structure of the database.

[18]Note that the real interest rate variable is unfortunately not as well covered as the other variables we control for in our analysis of the tax-elasticity in Section 1.6. To be able to control for the interest rate without reducing the sample size, we impute missing observations as follows. If for a given country one or more years are covered, then missing values of that country are imputed with the mean over these observed values. For countries without any coverage in the *WDI* database, we impute using a mean over all values of the countries for which values are observed.

[19]Note that since the *WGI* database was only updated biennially between 1996 and 2002, we impute the missing 2001 values by taking the mean of the respective variables of the years 2000 and 2002.

The industry-specific variables (in basic prices in $1,000$ current year US$) that we take from the *Eora26* database are: Gross output (GO_{cit}), gross input (GI_{cit}), compensation of employees (COE_{cit}), net taxes on production (calculated as the difference between taxes on production and subsidies on production) ($net\ TOP_{cit}$), net operating surplus (NOS_{cit}), net mixed income (NMI_{cit}), and consumption of fixed capital ($COFC_{cit}$). Additionally, the *Eora26* database records sector-specific information on greenhouse gas emissions associated with production (Kanemoto et al., 2014; Kanemoto et al., 2016). For our purpose, we use the variable total CO_2 emissions in gigagrams ($CO2_{cit}$). In total, all variables are covered for 189 countries over the time span 1990 to 2016.

1.4. Calculating Country-Industry-Specific Weights

In this section, we detail how we calculate the country-industry-specific financing (i.e., DS_{ci} ES_{ci}) and asset weights (i.e., w_{aci}) that are needed to compute the country-industry-year-specific FLETRs that we use for the estimation of the tax semi-elasticity of investment. More precisely, the industry levels we distinguish are the NACE Rev. 2 (ISIC Rev. 4) *sections*.[20]

The derivation of the weights is undertaken in two steps. First, we compute (or impute) country-industry-year-specific weights for years within the time horizon for which we want to calculate the FLETRs.[21] Second, we obtain the time-constant weights by taking averages over all year-specific weights belonging to a certain country-industry combination.

Depending on the respective data availability, a certain country-industry-specific weight may either be obtained (i) directly from data, (ii) by imputation using a matching algorithm, or (iii) by imputation using weights from countries in geographical proximity. The preferred approach is (i); approach (ii) is only implemented in the case where (i) does not yield a single

[20] Note that we can only calculate weights and therefore FLETRs for 19 of the 21 sections. We cannot calculate weights for the sections *T activities of households as employers; undifferentiated goods- and services-producing activities of households for own use*, and *U activities of extraterritorial organisations and bodies*, due to lack of data.

[21] Note that we generally only consider data for the years 2001 to 2018 for the calculation of the weights. The year 2001 is chosen as first year as it is also the first year for which we calculate FLETRs. The year 2018 is the latest year for which sufficient capital stock information from the *EUKLEMS & INTANProd* is available.

year-specific weight; approach (iii) is used only when also approach (ii) does not yield a single year-specific weight due to lack of data for matching. Note that by requiring an approach to only yield a minimum of one year-specific data point, we obtain final time-constant weights that are averages over varying numbers of years.[22]

1.4.1. Financing Structure

For the derivation of the country-industry-specific financing structures, we start by calculating debt ratios at the firm-entity level using data from *Orbis*. Following Steinmüller et al. (2019), we define the debt ratio of firm-entity j in year t, DS_{jt}, as long-term debt over total assets.[23] More specifically, for long-term debt we use the *Orbis* variable non-current liabilities ($NCLI_{jt}$).[24] Formally, we have

$$DS_{jt} = \frac{NCLI_{jt}}{TOAS_{jt}}. \tag{1.4}$$

We proceed by aggregating the entity-level data points from (1.4) to the final country-industry-specific debt shares in two steps. First, we create year-specific debt shares for country c and industry i, DS_{cit}, by taking unweighted means over all firm entities belonging to a given country-industry-year bin.[25] Second, we obtain the final time-constant debt shares, DS_{ci}, by taking unweighted means over all available year-specific debt shares, DS_{cit}, corresponding to the given country-industry pair. Respective equity shares are then obtained by subtracting these debt shares from unity, i.e., $ES_{ci} = 1 - DS_{ci}$.

[22] The reason for not combining the different approaches to maximize the number of year-specific observations is that we perceive having possibly few but very precise yearly weights preferable to having a larger number of yearly weights out of which some are imputed with less precision.

[23] Steinmüller et al. (2019) argue that only long-term debt can be harnessed to finance investment projects and is therefore the relevant measure to be considered when assessing an entity's investment opportunities, even if it underestimates its actual (total) debt ratio.

[24] Note that we exclude observations with non-positive total assets and set ratios that are negative due to negative long-term debt equal to zero. As we do not allow weights to exceed unity, we set debt ratios exceeding unity to one.

[25] To obtain meaningful values, we set the minimum number of firm entities per bin to five.

1.4.2. Asset Structure

For the calculation of the country-industry-specific asset weights, we use data from two different sources. For the asset categories *Buildings, Machinery, Office equipment, Computer equipment, Intangible fixed assets,* and *Vehicles*, we use data from the *EUKLEMS & INTANProd* database. Information on the asset category *Inventory* is obtained from *Orbis*. Since the coverage of these two data sources differs, we first calculate time-constant asset weights using only the *EUKLEMS & INTANProd* data, without taking inventory into account, i.e., the weights for *Buildings, Machinery, Office equipment, Computer equipment, Intangible fixed assets,* and *Vehicles* initially sum up to unity without inventory. Then, we determine time-constant inventory weights and rescale the weights of the other asset types such that the weights of all assets – including inventory – add up to unity. The advantage of separating the calculations this way is that we are not limiting the data usage to years that are covered by both sources, but instead are able to use all available information. As noted in the data section above, we regroup the capital stock variables that we obtain from *EUKLEMS & INTANProd* to match the ones that we distinguish for the calculation of the FLETRs.[26]

Next, for each country-industry-year combination, we sum up the six asset stock figures and take shares for the individual asset types denoted by w^*_{acit}. The superscript asterisk indicates that the weights are not yet re-scaled with the inventory share. We then obtain the respective time-constant, country-industry-specific weights, w^*_{aci}, by taking unweighted means over all available year-specific weights w^*_{acit}.

For the calculation of the inventory shares we follow Egger et al. (2009) and Egger and Loretz (2010), who define the firm-entity j-specific inventory share in year t as

$$w_{\text{invent},jt} = \frac{INV_{jt}}{TFAS_{jt} + IFAS_{jt} + INV_{jt}}, \tag{1.5}$$

[26]For details on the mapping of the capital stock variables to the asset categories used in this paper, see Appendix 1.8.5.

where $TFAS_{jt}$, $IFAS_{jt}$, and INV_{jt} denote tangible fixed assets, intangible fixed assets, and stocks of current assets (i.e., inventories), respectively. The aggregation to country-industry-year-specific inventory weights, $w_{\text{invent},cit}$, and then the final time-constant weights, $w_{\text{invent},ci}$, is identical to the one used for the debt shares in Section 1.4.1.

Finally, we re-scale the time-constant asset weights obtained from *EUKLEMS & INTANProd* by multiplying each of them with the factor $(1 - w_{\text{invent},ci})$. So, for instance, the final weights for the asset type *Buildings*, $w_{\text{build},ci}$, are obtained as $w_{\text{build},ci} = w^{*}_{\text{build},ci} \cdot (1 - w_{\text{invent},ci})$. This ensures that the sum over all seven asset types equals unity.

1.4.3. Imputation

Using the *Orbis* and *EUKLEMS & INTANProd* databases does not yield financing and asset structures for all country-industry combinations for which we intend to calculate FLETRs. Therefore, we implement an imputation strategy that matches observed weights from country-industry pairs that are covered in the data to those that are missing.

The matching algorithm that we use for the imputation is *Predictive Mean Matching* (PMM) (Little, 1988; Rubin, 1986). The PMM-based imputation of a single missing weight corresponding to country k, industry l, and year m, denoted by $y^{miss}_{c=k,i=l,t=m}$, is carried out as follows.[27] In a first step, we estimate a linear model, using all observations corresponding to the same industry l. Formally, this model can be written as

$$\boldsymbol{y}^{obs}_{c,i=l,t} = \boldsymbol{\beta}_{i=l} \boldsymbol{X}^{obs}_{c,i=l,t} + \boldsymbol{\varepsilon}^{obs}_{c,i=l,t}. \tag{1.6}$$

$\boldsymbol{y}^{obs}_{c,i=l,t}$ denotes the vector of all observed weights for industry l. $\boldsymbol{X}^{obs}_{c,i=l,t}$ denotes a matrix of covariates that are used for the matching (including a vector of ones, i.e., a constant is always included) and $\boldsymbol{\beta}_{i=l}$ denotes the corresponding coefficient vector. The model errors are collected in the vector $\boldsymbol{\varepsilon}^{obs}_{c,i=l,t}$. Estimating (1.6) yields the coefficient estimate vector $\widehat{\boldsymbol{\beta}}_{i=l}$ that is then used to form predictions for all complete cases that were used to estimate (1.6),

[27] We follow the notation of van Buuren (2018).

i.e.,

$$\widehat{\boldsymbol{y}}^{obs}_{c,i=l,t} = \widehat{\boldsymbol{\beta}}_{i=l}\boldsymbol{X}^{obs}_{c,i=l,t}. \tag{1.7}$$

Furthermore, $\widehat{\boldsymbol{\beta}}_{i=l}$ is used to calculate an estimate for the case we want to impute, i.e.,

$$\widehat{y}^{miss}_{c=k,i=l,t=m} = \widehat{\boldsymbol{\beta}}_{i=l}\boldsymbol{X}^{miss}_{c=k,i=l,t=m}, \tag{1.8}$$

with $\boldsymbol{X}^{miss}_{c=k,i=l,t=m}$ denoting the covariates for the missing observation. The missing weight $y^{miss}_{c=k,i=l,t=m}$ is imputed with the observed weight (the so-called donor), $y^{obs}_{c=o,i=l,t=p}$, for which

$$|\widehat{y}^{miss}_{c=k,i=l,t=m} - \widehat{y}^{obs}_{c=o,i=l,t=p}| \tag{1.9}$$

is minimal. We require the donor to be from the same industry that we are looking to impute (here industry l). However, the donor must not necessarily stem from the same year of the data point we are looking to impute, i.e., m and p in (1.9) may be different.[28]

An advantage that PMM holds over other so-called "hot deck" imputation methods, i.e., methods that use values observed elsewhere for imputation, is that the covariates are summarized into one matching metric using a weighting scheme, i.e., the $\widehat{\boldsymbol{\beta}}_i$, that reflects the importance of the different covariates for predicting financing and asset weights.[29] Another advantage of PMM is that it is implicit (Little and Rubin, 2019), i.e., there is no need to define distributions from which noise components for imputed values are randomly drawn from.[30] Instead, the only assumption that has to be invoked is that the distribution of a missing entry is identical to the observed data of the donor (Van Buuren, 2018).

[28] Alternatively, instead of using just the donor for which the corresponding prediction is closest to the prediction of the missing data point, the mean of the d-closest matches can be considered for imputation. As robustness check, we graphically provide imputation results for $d = 5, 10$, and 15 in Appendix 1.8.6.

[29] In contrast, for instance with the widely used k-Nearest Neighbor (k-NN) matching, all included covariates are assigned the same importance for finding a match (Hastie et al., 2009, ch. 13.3). In Appendix 1.8.6, we provide graphical evidence, using k-NN for imputation and compare the results to the ones obtained using PMM.

[30] Methods that involve random noise components are discussed in Van Buuren (2018, ch. 3.2).

We first impute the financing weights. The dependent variable is the observed yearly debt share, which we get from Orbis, DS_{cit} (see Section 1.4.1). The country-level covariates used for the matching largely follow the ones used by Goldbach et al. (2021), and broadly aim at capturing the condition of a country's financial market, the strength of its institutions, as well as its overall economic development. More specifically, we control for the *Rule of Law* indicator (ROL_{ct}), the *Control of Corruption* indicator ($Corruption_{ct}$), the logarithm of the variable measuring domestic credit provided to the private sector relative to a country's GDP ($log\ DCPS_{ct}$), annual inflation ($Inflation_{ct}$), as well as GDP growth ($GDP\ growth_{ct}$). As described above, all these variables are taken from the World Bank's WDI database. Furthermore, we include the statutory tax rate τ_{ct} as a proxy for a country's corporate tax code. Additionally, we condition on a set of country-industry-level variables to account for the size and characteristics of industries. These variables are the logarithm of gross output ($log\ GO_{cit}$), gross input ($log\ GI_{cit}$), compensation of employees ($log\ COE_{cit}$), net operating surplus ($log\ NOS_{cit}$), net mixed income ($log\ NMI_{cit}$), paid net taxes on production ($log\ net\ TOP_{cit}$), and consumption of fixed capital ($log\ COFC_{cit}$). As described above, all country-industry-level variables are taken from the *Eora26* database. Finally, we include year indicators to control for year-specific effects that are common to all countries.[31] Once we have imputed the yearly debt ratios, DS_{cit}, for country-industry combinations that are not covered in *Orbis*, we proceed to compute time-constant debt and retained earnings shares as described in Section 1.4.1.

Next, we proceed to impute the asset weights. As discussed above in Section 1.4.2, we calculate the asset weights for inventory and the six other asset types separately using two different databases with different coverage. As a result, in many cases, we only need to impute the inventory share or the composition of the other asset types, but not both. To optimally use all available data and to be able to sensibly combine imputed and observed

[31] Descriptive statistics of the matching covariates are presented in Appendix 1.8.7. Note that we can only impute yearly financing weights for the years 2001 to 2016, as 2016 is the last year for which the industry-specific matching covariates are available.

asset weights into one structure, we disregard inventory when imputing the asset categories *Buildings*, *Machinery*, *Office equipment*, *Computer equipment*, *Intangible fixed assets*, and *Vehicles*. That is, we use the yearly weights derived from the *EUKLEMS & INTANProd* database that have not yet been rescaled with the inventory share (denoted by w^*_{acit} in Section 1.4.2). By construction, these w^*_{acit}'s add up to unity in each year for a given country-industry combination. For the imputed equivalents, however, this is not necessarily the case, as different donors may be drawn for each asset type. We therefore rescale the imputed w^*_{acit}'s such that they add up to unity at the year level for each country-industry combination. Thereafter, the derivation of the final time-constant asset structures is identical to the procedure described in Section 1.4.2.

For the imputation of the asset weights, we again use a combination of country-specific and country-industry-specific covariates to control for market size and market conditions, economic development, as well as the industry-specific structure of primary inputs and production.[32] At the country level, we control for the logarithm of GDP ($log\ GDP_{ct}$) and GDP per capita ($log\ GDP\ p.c._{ct}$). At the industry level, we control for the logarithm of the compensation of employees ($log\ COE_{cit}$), the net operating surplus ($log\ NOS_{cit}$), the net mixed income ($log\ NMI_{cit}$), the consumption of fixed labour ($log\ COFC_{cit}$), as well as the logarithm of CO_2 emissions ($log\ CO2_{cit}$). Finally, we control for year-specific effects by including time dummies.[33]

Due to a lack of data on covariates, missing financing structures in 53 countries and asset structures in 56 countries cannot be imputed using the PMM procedure. In order to calculate FLETRs for these countries, they are assigned the time-constant observed and/or PMM-imputed weights of their geographical neighbors. For instance, San Marino is assigned the weights from Italy and Andorra is assigned the mean of the weights of France and Spain. More than half of the countries that we are missing are small islands in the Caribbean

[32]Note that we use the same covariates for the imputation of each asset type.

[33]Descriptive statistics of the covariates used for the matching of asset weights are presented in Appendix 1.8.7. Again, note that we can only impute yearly weights for the years 2001 to 2016 due to covariate coverage.

region or Oceania. In these cases, missing asset and/or financing weights are replaced by the region-specific mean of all non-missing weights.[34] [35]

1.4.4. Descriptive Statistics of Asset and Financing Weights

Table 1.1 presents summary statistics of the time-constant, country-industry-specific financing and asset structures that we use for the calculation of the FLETRs. The summary statistics are grouped by the different approaches we use obtain the weights. Panel A only depicts information on weights that are directly derived from the primary data sources. The Panels B and C describe weights we have imputed using the PMM procedure or values from geographically proximate countries, respectively. A key result that holds for each panel is that there is substantial variation between the mean values of the debt ratios and in particular the asset weights of the different industries. Intuitively, this heterogeneity seems plausible. For instance, section *C manufacturing* exhibits a noticeably higher share of *machinery* in its mean asset composition than the service industries, e.g., *P education*. Additionally, there is substantial variation in every weight within each industry, irrespective of the method used to derive it, as indicated by the standard deviations. The fact that the variation is strong not only for the weights derived directly from the data but also for the PMM imputed weights (Panel B) indicates that for the latter approach, a wide range of observed values was drawn for the matching.[36] Overall, the strong variation both between and within industries corroborates our approach of estimating country-industry-specific financing and asset compositions for the calculation of FLETRs. Assuming symmetric asset and financing structures across all countries and industries, as done by most of the previous literature, most likely leads to imprecise tax measures and introduces measurement error. We finally provide a number of plausibility checks also by looking at single data points in Appendix 1.8.6.

[34] The exact imputation using geographically close countries is detailed in Appendix 1.8.8.

[35] Note that we have collected statutory tax rates and tax base rules for these countries. As the variation in EMTRs is largely driven by statutory tax determinants, we prefer to make somewhat stricter assumptions on asset and financing weights, but instead are able to keep these countries in our sample.

[36] In Appendix 1.8.6, we illustrate graphically that this result is robust to increasing the number of donors considered for the imputation of a single year-specific data point.

Table 1.1: *DESCRIPTIVES ON IMPUTED COUNTRY-INDUSTRY-SPECIFIC FINANCING AND ASSET WEIGHTS*

The table depicts means (expressed in %) and standard deviations (in brackets) of the financing and asset weights by NACE Rev. 2 (ISIC Rev. 4) sections. Panel A depicts weights that are derived directly from data sources (see Sections 1.4.1 and 1.4.2). Panels B and C depict weights that are imputed using PMM with $d = 1$ donor and weight values of geographically proximate countries, respectively (see Section 1.4.3). The last row of each panel gives the number of the country-industry-specific weights across all industries. Descriptions for the different NACE Rev. 2 (ISIC Rev. 4) sections are provided in Appendix 1.8.3.

Panel A: Weights derived directly from data sources

	DS_{ci}	$w_{\text{build},ci}$	$w_{\text{comp},ci}$	$w_{\text{ifas},ci}$	$w_{\text{mach},ci}$	$w_{\text{office},ci}$	$w_{\text{vehic},ci}$	$w_{\text{invent},ci}$
A	18.0 (0.098)	51.6 (0.116)	0.2 (0.002)	0.5 (0.007)	18.7 (0.088)	0.3 (0.007)	5.0 (0.030)	26.4 (0.096)
B	17.8 (0.093)	47.7 (0.148)	0.3 (0.003)	1.9 (0.020)	28.3 (0.135)	0.5 (0.009)	2.9 (0.025)	21.0 (0.089)
C	15.1 (0.096)	24.4 (0.102)	0.5 (0.003)	10.9 (0.086)	25.5 (0.057)	0.6 (0.008)	1.4 (0.007)	37.4 (0.102)
D	24.4 (0.120)	64.9 (0.166)	0.3 (0.002)	1.7 (0.022)	23.9 (0.135)	0.9 (0.015)	0.8 (0.011)	11.6 (0.107)
E	17.7 (0.110)	73.0 (0.100)	0.2 (0.002)	0.9 (0.008)	9.8 (0.075)	0.4 (0.007)	1.9 (0.010)	16.5 (0.102)
F	14.6 (0.110)	35.6 (0.115)	0.7 (0.006)	1.8 (0.021)	18.8 (0.096)	0.5 (0.005)	9.7 (0.048)	33.2 (0.145)
G	11.8 (0.091)	29.1 (0.066)	0.9 (0.006)	2.6 (0.021)	10.2 (0.031)	0.6 (0.006)	4.5 (0.024)	54.3 (0.129)
H	18.3 (0.106)	56.3 (0.140)	0.4 (0.002)	1.3 (0.015)	7.4 (0.042)	0.9 (0.010)	26.8 (0.143)	9.8 (0.077)
I	20.4 (0.124)	59.9 (0.096)	0.7 (0.006)	1.3 (0.023)	15.4 (0.073)	1.1 (0.011)	2.5 (0.024)	18.8 (0.142)
J	12.7 (0.086)	40.9 (0.136)	3.1 (0.022)	15.1 (0.103)	10.1 (0.065)	11.1 (0.093)	3.3 (0.082)	18.2 (0.099)
K	14.0 (0.082)	54.4 (0.188)	3.9 (0.025)	15.4 (0.114)	8.0 (0.063)	1.7 (0.027)	4.4 (0.047)	11.8 (0.096)
L	22.8 (0.127)	85.9 (0.080)	0.0 (0.001)	0.1 (0.001)	0.4 (0.005)	0.0 (0.001)	0.1 (0.002)	18.1 (0.144)
M	11.6 (0.072)	36.8 (0.128)	3.1 (0.026)	27.9 (0.152)	11.4 (0.041)	1.7 (0.024)	4.4 (0.017)	16.9 (0.091)
N	14.3 (0.085)	28.5 (0.144)	2.1 (0.014)	4.5 (0.061)	19.1 (0.082)	2.3 (0.029)	30.7 (0.131)	15.7 (0.087)
O	14.3 (0.124)	76.0 (0.087)	0.5 (0.003)	1.9 (0.012)	8.0 (0.066)	0.3 (0.003)	1.7 (0.021)	14.9 (0.107)
P	14.0 (0.115)	68.1 (0.136)	0.8 (0.005)	14.3 (0.104)	5.4 (0.032)	0.5 (0.006)	0.8 (0.007)	10.1 (0.080)
Q	16.3 (0.110)	66.1 (0.139)	1.1 (0.009)	4.1 (0.040)	17.7 (0.103)	1.0 (0.018)	1.7 (0.007)	11.7 (0.077)
R	16.8 (0.120)	68.7 (0.143)	1.1 (0.010)	3.0 (0.038)	12.6 (0.119)	1.5 (0.018)	1.7 (0.016)	12.4 (0.079)
S	14.7 (0.103)	50.7 (0.100)	1.5 (0.016)	4.9 (0.069)	12.1 (0.075)	1.2 (0.017)	3.8 (0.028)	25.6 (0.127)
Obs	1,321	394	394	394	394	394	394	1,261

Panel B: Weights imputed using PMM

	DS_{ci}	$w_{\text{build},ci}$	$w_{\text{comp},ci}$	$w_{\text{ifas},ci}$	$w_{\text{mach},ci}$	$w_{\text{office},ci}$	$w_{\text{vehic},ci}$	$w_{\text{invent},ci}$
A	14.7 (0.050)	33.7 (0.108)	0.2 (0.001)	2.3 (0.013)	30.4 (0.080)	2.5 (0.013)	4.4 (0.025)	26.1 (0.050)
B	14.9 (0.048)	57.9 (0.085)	0.3 (0.003)	1.3 (0.011)	13.6 (0.082)	0.5 (0.010)	2.3 (0.016)	25.0 (0.045)
C	14.3 (0.042)	33.0 (0.057)	0.6 (0.004)	1.7 (0.034)	24.3 (0.041)	0.9 (0.009)	2.4 (0.005)	36.5 (0.026)
D	18.9 (0.048)	64.0 (0.071)	0.2 (0.002)	1.3 (0.007)	19.5 (0.053)	0.3 (0.010)	0.5 (0.006)	14.6 (0.037)
E	9.4 (0.050)	65.9 (0.106)	0.4 (0.003)	0.8 (0.007)	12.6 (0.081)	0.4 (0.008)	3.1 (0.010)	16.8 (0.059)
F	12.9 (0.050)	40.7 (0.092)	0.7 (0.003)	1.1 (0.009)	17.9 (0.052)	0.3 (0.004)	7.2 (0.048)	31.9 (0.037)
G	12.0 (0.041)	27.3 (0.060)	0.4 (0.006)	1.2 (0.015)	8.7 (0.037)	0.8 (0.003)	4.4 (0.015)	58.7 (0.046)
H	12.1 (0.055)	52.4 (0.121)	0.4 (0.001)	0.4 (0.004)	4.7 (0.024)	0.6 (0.005)	27.1 (0.142)	16.1 (0.065)
I	13.5 (0.072)	56.4 (0.092)	0.4 (0.004)	1.0 (0.013)	14.7 (0.060)	2.4 (0.010)	1.8 (0.010)	25.0 (0.052)
J	9.1 (0.036)	42.4 (0.104)	1.5 (0.017)	8.4 (0.049)	9.5 (0.054)	8.4 (0.049)	3.6 (0.074)	29.5 (0.096)
K	10.6 (0.037)	29.2 (0.202)	6.8 (0.046)	28.0 (0.162)	9.7 (0.034)	5.2 (0.067)	8.8 (0.064)	12.6 (0.071)
L	17.3 (0.080)	72.5 (0.145)	0.0 (0.001)	0.2 (0.003)	1.2 (0.008)	0.1 (0.001)	0.2 (0.001)	28.7 (0.129)
M	9.4 (0.033)	48.9 (0.113)	1.4 (0.023)	12.4 (0.117)	13.5 (0.041)	2.2 (0.031)	3.3 (0.017)	18.7 (0.054)
N	12.6 (0.038)	44.4 (0.119)	2.1 (0.010)	1.8 (0.013)	16.5 (0.064)	2.7 (0.029)	14.1 (0.113)	19.1 (0.078)
O	8.5 (0.049)	68.6 (0.089)	0.7 (0.003)	0.7 (0.008)	11.6 (0.079)	0.3 (0.005)	1.8 (0.028)	16.0 (0.050)
P	9.3 (0.056)	70.9 (0.091)	1.7 (0.006)	9.9 (0.091)	6.4 (0.027)	0.7 (0.004)	0.8 (0.004)	9.4 (0.040)
Q	11.6 (0.040)	39.8 (0.179)	1.7 (0.008)	4.5 (0.032)	37.2 (0.177)	2.9 (0.018)	2.3 (0.007)	11.0 (0.070)
R	8.5 (0.061)	52.2 (0.141)	1.5 (0.009)	2.0 (0.014)	24.0 (0.120)	1.9 (0.024)	6.7 (0.040)	11.4 (0.053)
S	9.9 (0.056)	41.3 (0.113)	0.9 (0.007)	1.7 (0.029)	12.0 (0.073)	1.0 (0.006)	7.6 (0.033)	38.7 (0.112)
Obs	1,921	2,741	2,741	2,741	2,741	2,741	2,741	1,930

Panel C: Weights imputed using values of geographically proximate countries

	DS_{ci}	$w_{\text{build},ci}$	$w_{\text{comp},ci}$	$w_{\text{ifas},ci}$	$w_{\text{mach},ci}$	$w_{\text{office},ci}$	$w_{\text{vehic},ci}$	$w_{\text{invent},ci}$
A	17.1 (0.046)	39.7 (0.096)	0.1 (0.001)	2.1 (0.008)	26.8 (0.055)	2.2 (0.010)	4.6 (0.013)	24.3 (0.058)
B	16.0 (0.035)	58.4 (0.060)	0.3 (0.001)	1.2 (0.012)	15.9 (0.055)	0.3 (0.001)	1.4 (0.010)	22.5 (0.047)
C	14.1 (0.054)	30.3 (0.061)	0.6 (0.002)	4.0 (0.060)	24.3 (0.031)	0.7 (0.004)	2.2 (0.005)	37.8 (0.043)
D	21.9 (0.058)	62.1 (0.078)	0.3 (0.001)	1.6 (0.017)	21.1 (0.070)	0.5 (0.009)	0.7 (0.008)	13.8 (0.039)
E	13.5 (0.054)	70.8 (0.080)	0.3 (0.002)	0.7 (0.005)	10.0 (0.055)	0.2 (0.003)	2.2 (0.007)	15.0 (0.057)
F	14.8 (0.048)	42.6 (0.076)	0.6 (0.002)	1.0 (0.008)	15.3 (0.028)	0.4 (0.002)	7.5 (0.029)	31.9 (0.053)
G	13.7 (0.047)	25.9 (0.034)	0.5 (0.003)	1.5 (0.011)	8.7 (0.020)	0.7 (0.002)	4.4 (0.008)	58.3 (0.060)
H	14.1 (0.064)	46.7 (0.088)	0.4 (0.001)	0.7 (0.006)	6.2 (0.026)	0.6 (0.004)	34.0 (0.102)	11.4 (0.049)
I	16.1 (0.070)	57.3 (0.060)	0.5 (0.002)	0.9 (0.010)	16.2 (0.035)	1.9 (0.006)	2.0 (0.006)	21.2 (0.068)
J	10.2 (0.051)	42.2 (0.073)	3.0 (0.016)	10.3 (0.078)	9.4 (0.028)	9.6 (0.043)	3.1 (0.020)	22.4 (0.067)
K	13.1 (0.043)	28.4 (0.162)	9.4 (0.037)	32.7 (0.128)	10.4 (0.037)	3.4 (0.024)	6.0 (0.040)	9.6 (0.053)
L	21.2 (0.066)	76.7 (0.103)	0.1 (0.000)	0.2 (0.001)	1.2 (0.005)	0.0 (0.001)	0.2 (0.001)	21.2 (0.083)
M	10.6 (0.040)	44.9 (0.076)	1.6 (0.014)	19.4 (0.114)	13.7 (0.034)	1.5 (0.011)	3.2 (0.007)	15.5 (0.049)
N	13.7 (0.038)	41.1 (0.092)	2.2 (0.005)	2.3 (0.019)	18.9 (0.052)	2.4 (0.012)	17.5 (0.096)	15.5 (0.058)
O	9.6 (0.033)	68.6 (0.050)	0.6 (0.002)	1.4 (0.008)	11.2 (0.042)	0.2 (0.003)	1.8 (0.015)	16.2 (0.049)
P	11.8 (0.044)	68.8 (0.052)	1.5 (0.004)	13.9 (0.057)	6.1 (0.014)	0.7 (0.003)	0.9 (0.003)	8.3 (0.034)
Q	14.0 (0.052)	42.8 (0.133)	1.8 (0.006)	4.5 (0.015)	36.9 (0.104)	2.3 (0.009)	2.4 (0.005)	9.1 (0.049)
R	12.4 (0.058)	54.1 (0.109)	1.5 (0.005)	2.0 (0.009)	23.8 (0.085)	1.7 (0.010)	6.0 (0.025)	10.7 (0.038)
S	12.5 (0.053)	43.4 (0.066)	1.3 (0.007)	3.0 (0.045)	12.9 (0.034)	1.0 (0.008)	6.1 (0.021)	32.2 (0.084)
Obs	957	1,064	1,064	1,064	1,064	1,064	1,064	1,008

1.5. Country-Industry-Year-specific FLETRs

In this section, we calculate and describe the new country-industry-year-specific FLETRs using the time-constant, country-industry-specific financing and asset weights we have calculated and estimated in the previous section. In a first step, we compute the NPV of depreciation allowances, δ_{cit}, by plugging the financing shares, ES_{ci} and DS_{ci}, as well as the asset shares, w_{aci}, into equation (1.2). The country-industry-year-specific EMTRs are then obtained by inserting δ_{cit} as well as the statutory tax rate, τ_{ct}, into (1.3). For the calculation of the country-industry-year-specific EATRs we use the same NPV of depreciation allowances, δ_{cit}. For details on the calculation of the EATR as well as descriptions, refer to Appendix 1.8.1 and Appendix 1.8.9, respectively.

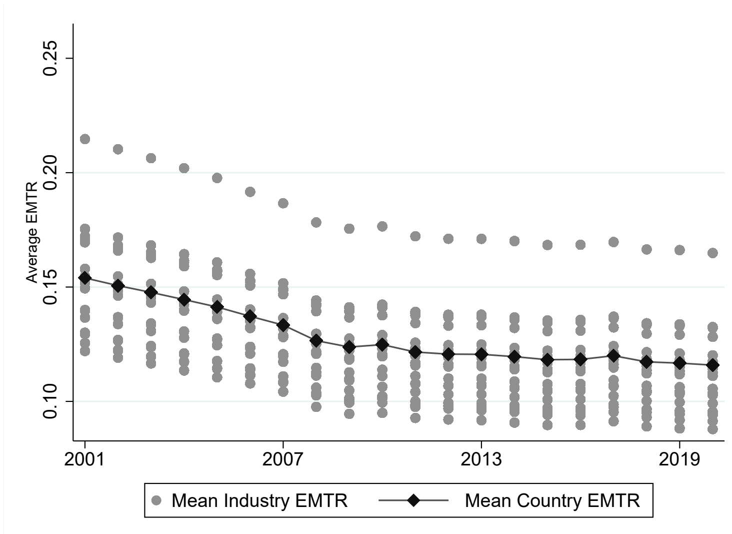
For the sake of comparison, we additionally calculate EMTRs that are based on symmetric financing and asset weights for all countries and industries, which is the common approach in the existing literature. More precisely, we use the parameterization by Steinmüller et al. (2019) that matches the asset types that are also used in the paper at hand.[37] We denote the EMTRs and NPVs of depreciations allowances based on these symmetric weights as $EMTR_{ct}$ and δ_{ct}, respectively.

In a first step of analyzing the new country-industry-specific EMTRs, we plot year-specific means over all countries for each industry. For the sake of comparison, we add year-specific means of the country-year specific EMTRs over all countries to the plot. The resulting Figure 1.1 suggests that the country-industry-year-specific EMTRs follow, on average, the same downward trend as their country-year-specific counterparts. There is, however, substantial variation in the average EMTRs across industries implying that the country-year-specific average EMTRs significantly over-/underestimate the tax burden for certain industries. For example, firms operating in the sections *Construction*, *Manufacturing*, as well as *Wholesale*

[37] In detail, the asset structure is composed as follows: *Buildings* 38%, *Computer equipment* 2%, *Intangible fixed assets* 11%, *Inventory* 26%, *Machinery* 2%, *Office equipment* 1%, *Vehicles* 2%. The debt-financing share and the equity-financing share are assumed to amount to 1/3 and 2/3, respectively (Steinmüller et al., 2019).

Figure 1.1: ***DEVELOPMENT OF MEAN COUNTRY-YEAR AND COUNTRY-INDUSTRY-YEAR-SPECIFIC EMTRs***

The figure depicts the development of the mean country-year and country-industry-year-specific EMTRs calculated in Section 1.5. The grey dots represent the mean country-industry-year-specific EMTRs across countries for each year. The black dots that are connected by black lines represent the mean country-year-specific EMTRs across countries for each year. Calculations are based on a sample of 75,126 observations.



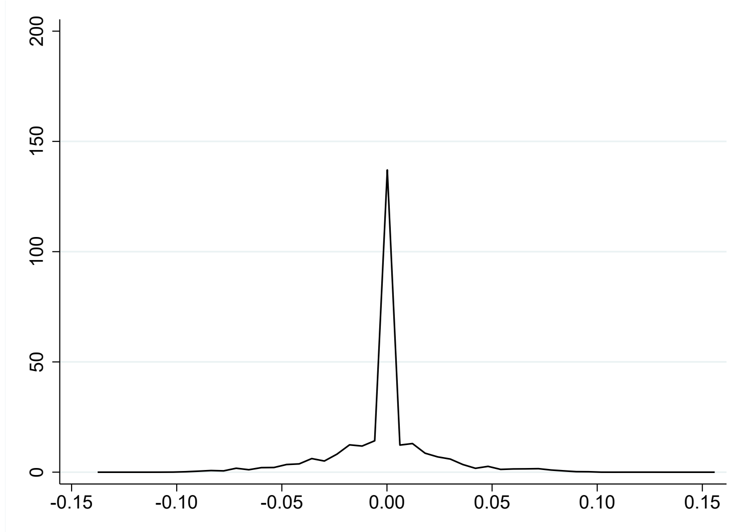
and retail trade face among highest average EMTRs.[38] On the other hand, firms engaged in *Arts, entertainment, and recreation*, *Financial and insurance activities*, as well as *Human health and social work activities* face the lowest effective tax burden. Overall, the findings in Figure 1.1 suggest that disregarding country-industry-level heterogeneity for calculating EMTRs leads to a systematic measurement error.

To further explore the heterogeneity from using country-industry-year-specific versus country-year-specific EMTRs, we take the difference between the levels of these two measures ($EMTR_{cit} - EMTR_{ct}$) and plot the corresponding distribution (Figure 1.2). It can be seen that most of the mass of the density plot is located on an interval of plus/minus five percentage points with a steep spike on the interval of plus/minus one percentage point. This suggests that the additional variation in $EMTR_{cit}$ that is introduced by the country-industry-specific financing and asset weights does not lead to a large structural deviation

[38]Note that the largest EMTRs are the ones for *Wholesale and retail trade*. This can in large parts be explained with the high inventory shares that we find for this industry and the fact that inventories are not subject to depreciation.

Figure 1.2: *DISTRIBUTION OF DEVIATIONS FROM THE COUNTRY-YEAR EMTRs*

The figure depicts the distribution of the differences between country-industry-year-specific and country-year-specific EMTRs calculated in Section 1.5. The distribution is calculated based on 75,126 observations using a triangle kernel with a bandwidth of 0.0005.



from EMTR measures where this heterogeneity is neglected. In other words, the finding suggests that the magnitude in the country-industry-specific EMTRs is mainly determined by the national tax code and only to a comparably smaller part by the country-industry-specific characteristics. This finding can be confirmed by performing a simple linear regression of $EMTR_{cit}$ on country-year fixed effects. These fixed effects reflect the margin at which changes to the tax code impact the $EMTR_{cit}$. Keeping in mind that the financing and asset structures are time-constant, the R^2 of such a regression can be interpreted as the share in variance of the $EMTR_{cit}$ that is attributable to national tax codes. Performing such a regression yields a high adjusted R^2 of 83.4%.

In contrast to the EMTR, the EATR exhibits much less industry-specific variation when applying the same country-industry-year-specific NPVs of depreciation allowances. A detailed analysis of the EATR can be found in Appendix 1.8.9.

1.6. Tax Semi-Elasticity of Firms' Tangible Fixed Assets

1.6.1. Empirical Approach

In this section, we calculate the tax semi-elasticity of investment, using the EMTRs calculated in Section 1.5. Following Steinmüller et al. (2019), we use the logarithm of firm-entity j's tangible fixed assets ($log\ TFAS_{jt}$) as dependent variable to capture real investment behavior. This outcome has been used regularly in the literature and is also common in studies examining the effect of (corporate) taxation on foreign investments. We provide more discussion on this measurement and the empirical specification below. We implement the following estimation equation

$$log\ TFAS_{jt} = \gamma EMTR_{cit} + \boldsymbol{\psi} \boldsymbol{X_{jt-1}} + \boldsymbol{\zeta} \boldsymbol{X_{ct}} + c_j + \theta_t + \varepsilon_{jt}. \tag{1.10}$$

The coefficient γ measures the semi-elasticity of investment[39] with respect to the marginal tax rate, $EMTR_{cit}$. We control for a set of lagged affiliate-specific variables, denoted by $\boldsymbol{X_{jt-1}}$, and a set of country-specific variables, denoted by $\boldsymbol{X_{ct}}$, both of which are described in more detail below. The corresponding parameter estimates are contained in the vectors $\boldsymbol{\psi}$ and $\boldsymbol{\zeta}$, respectively. Furthermore, we control for firm-entity and year-specific effects, which we denote by c_j and θ_t, respectively.[40] Finally, ε_{jt} denotes the error component.

Note that, given specification (1.10), the role of the financial and asset weights becomes less important (but of course not irrelevant) since we condition on firm-entity-j-specific heterogeneity c_j, and focus on the substantial variation in the EMTRs over time. This allows us to identify changes in investment behavior, which are driven exclusively by changes in the tax code.

[39] Note that we use "investment" in our micro-level panel data approach interchangeably for "investment in tangible fixed assets".

[40] Note also that we provide an extensive discussion on different types of fixed effects we might include in the estimations (see Section 1.6.4).

1.6.2. Sample and Control Variables

The control variables that we use largely follow Steinmüller et al. (2019). At the firm-entity level (indicated by index j), we include the one-period lag of the logarithm of sales ($log\ SALES_{jt-1}$) and cost of employees ($log\ STAF_{jt-1}$). Additionally, we include three entity-level ratios proposed by Liu (2020): the cash flow rate ($CF\ rate_{jt}$), defined as the cash flow in year t divided by the sum of tangible and intangible fixed assets in $t-1$; the one period lag of the sales growth rate ($SALES\ growth_{jt-1}$), i.e., the sales growth rate from $t-2$ to $t-1$; and the one period lag of the profit margin ($Profit\ margin_{jt-1}$), with the profit margin in t being defined as $EBIT_{jt}/SALES_{jt}$. To minimize the influence of obvious outliers, we winsorize all three ratio variables at the top and bottom 1 percentiles.

At the country level, we control for host country c's GDP ($log\ GDP_{ct}$), GDP per capita ($log\ GDP\ p.c._{ct}$), and the GDP growth rate ($GDP\ growth_{ct}$) as proxies for market size, the state of a country's economic development, and the general economic situation, respectively. Additionally, we control for inflation ($Inflation_{ct}$) to capture investment risk. In particular, following the arguments in Aggarwal and Kyaw (2008), as well as Huizinga et al. (2008), countries with higher inflation usually exhibit a higher risk premium and higher general business risk. Furthermore, we include the real interest rate ($Real\ interest\ rate_{ct}$) to control the cost of debt financing.[41] The variable domestic credit to private sector relative to a country's GDP ($log\ DCPS_{ct}$) is included as a measure for capital market depth. The corruption ($Corruption_{ct}$) and rule of law (ROL_{ct}) indicators capture the strength of institutions such as creditor rights. We finally control for the number of double taxation treaties ($NDTT_{ct}$) that a country has.

For our sample, we consider *Orbis* observations for the time span 2001, i.e., the first year for which we calculate FLETRs, to 2018, which is the last year for which all control variables are available. Following the literature (e.g., Liu, 2020; Steinmüller et al., 2019), we exclude a

[41] Note that, depending on the theoretical concept of expressing the EMTR formally, a benchmark interest rate might also feature in the tax formula. However, for the sake of measurability, we aim at keeping the EMTR formula as simple as possible, but condition on the interest rate.

number of industries from our analysis (as tax treatment of these industries differs from the standard one).[42] We finally impose the requirement that a firm entity must be observed at least twice in the sample period. Descriptive statistics for our final sample of over 24 million observations as well as a correlation matrix for selected variables are provided in Table 1.2.

Table 1.2: ***DESCRIPTIVES ON DATA SET USED FOR THE ESTIMATION OF THE TAX-ELASTICITY OF CORPORATE INVESTMENT***

The table depicts descriptive statistics on all the variables used for the estimation of the tax-elasticity of investment. Panel A reports descriptives on the different tax measures applied. Panel B reports descriptives on the firm-entity level variables. Panel C reports descriptives on the country level variables. Panel D depicts Pearson correlation coefficients for key variables. Definitions of the variables are provided in Section 1.6.2.

Panel A: Tax measures

	Observations	Mean	(sd)
$EMTR_{cit}$	24,144,916	0.160	(0.063)
$EATR_{cit}$	24,144,916	0.234	(0.068)
τ_{ct}	24,144,916	0.266	(0.076)
δ_{cit}	24,144,916	0.473	(0.148)

Panel B: Firm-entity level variables

	Observations	Mean	(sd)
$log\ TFAS_{jt}$	24,144,916	10.780	(2.462)
$log\ SALES_{jt-1}$	24,144,916	13.192	(1.889)
$log\ STAF_{jt-1}$	24,144,916	11.514	(1.938)
$CF\ rate_{jt}$	24,144,916	1.808	(11.006)
$SALES\ growth_{jt-1}$	24,144,916	0.278	(1.507)
$Profit\ margin_{jt-1}$	24,144,916	-0.025	(0.541)

Panel C: Country level variables

	Observations	Mean	(sd)
$log\ GDP_{ct}$	24,144,916	27.605	(1.146)
$log\ GDP\ p.c._{ct}$	24,144,916	10.475	(0.333)
$GDP\ growth_{ct}$	24,144,916	1.471	(2.408)
$Inflation_{ct}$	24,144,916	1.700	(2.338)
$Real\ interest\ rate_{ct}$	24,144,916	2.857	(1.864)
$log\ DCPS_{ct}$	24,144,916	4.450	(0.462)
$Corruption_{ct}$	24,144,916	0.779	(0.726)
ROL_{ct}	24,144,916	0.934	(0.591)
$NDTT_{ct}$	24,144,916	90.172	(22.611)

Panel D: Correlation matrix (24,144,916 observations)

	$log\ TFAS_{jt}$	$EMTR_{cit}$	$EATR_{cit}$	τ_{ct}	δ_{cit}
$log\ TFAS_{jt}$	1.000				
$EMTR_{cit}$	0.052	1.000			
$EATR_{cit}$	0.054	0.873	1.000		
τ_{ct}	0.051	0.793	0.989	1.000	
δ_{cit}	0.000	-0.495	-0.038	0.098	1.000

[42] In detail, these industries are denoted by the section codes A, B, K, O, P, Q, T, and U. For a description on these sections, see Appendix 1.8.3.

1.6.3. Basic Results

Table 1.3 presents the basic estimation results of the tax semi-elasticity of investment.[43] The results presented in columns (1) to (4) are based on the largest possible sample of more than 24 million observations with 4,787,866 individual firm entities in 70 countries.

Our benchmark specification in column (1) suggests an EMTR semi-elasticity of about -0.41, i.e., a 1 percentage point higher EMTR results in 0.41% less investment in tangible fixed assets. The corresponding elasticity equals -0.065 which is a very small effect compared to the previous literature.

Column (2) illustrates that the effect of the EATR is not only smaller but also slightly less statistically significant. This result is in line with expectations as discrete investment decisions may be less responsive to changes in tax incentives in the short-run. Moreover, the fixed effects approach removes all cross-sectional variation in the tangible fixed assets and identification is based on changes in the EMTR over time. In this sense, the EMTR should be the best measure to explain changes in outcome. Column (3) employs the statutory tax rate (τ_{ct}) as an alternative tax measure, which neither accounts for tax base effects nor for appropriate asset and financing weights. While the coefficient is still negative and statistically significant, it is substantially smaller compared to the EMTR. Column (4) distinguishes between τ_{ct} and the weighted δ_{cit} to differentiate between tax rate and tax base effects. The coefficients are both statistically significant and have the expected signs. An interesting finding here is that the corresponding elasticity for δ_{cit} is relatively high (approx. 0.13), suggesting that depreciation rates may be even more effective as a policy instrument to boost investments. Given these results, the newly calculated EMTRs capture tax incentives in the most appropriate way by incorporating both statutory tax policy changes and country-industry-specific firm characteristics.

Let us briefly discuss the findings for the other control variables. We may distinguish

[43] Note that we report robust standard errors that are clustered at the country-industry-year level, i.e., the level at which we merge the tax measures to the firm-entity-level data (Moulton, 1990).

Table 1.3: *BENCHMARK ESTIMATES*

The table presents OLS estimates. The dependent variable is the logarithm of firm-entity j's tangible fixed assets, $log\ TFAS_{jt}$. Robust standard errors are reported in parentheses (clustered at the country-industry-year level). *** denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level. The last rows report elasticities corresponding to the tax measure(s) used in the respective model. Corresponding standard errors are obtained using the Delta method. Definitions and descriptive statistics on the explanatory variables are provided in Section 1.6.2.

	(1)	(2)	(3)	(4)
$EMTR_{cit}$	-0.405**			
	(0.166)			
$EATR_{cit}$		-0.298**		
		(0.138)		
τ_{ct}			-0.248**	-0.253**
			(0.122)	(0.122)
δ_{cit}				0.273**
				(0.113)
$log\ SALES_{jt-1}$	0.266***	0.266***	0.266***	0.266***
	(0.005)	(0.005)	(0.005)	(0.005)
$log\ STAF_{jt-1}$	0.087***	0.087***	0.087***	0.087***
	(0.002)	(0.002)	(0.002)	(0.002)
$CF\ rate_{jt}$	-0.004***	-0.004***	-0.004***	-0.004***
	(0.000)	(0.000)	(0.000)	(0.000)
$SALES\ growth_{jt-1}$	-0.010***	-0.010***	-0.010***	-0.010***
	(0.000)	(0.000)	(0.000)	(0.000)
$Profit\ margin_{jt-1}$	-0.067***	-0.067***	-0.067***	-0.067***
	(0.003)	(0.003)	(0.003)	(0.003)
$log\ GDP_{ct}$	1.602***	1.603***	1.603***	1.603***
	(0.285)	(0.286)	(0.286)	(0.286)
$log\ GDP\ p.c._{ct}$	-0.504*	-0.508*	-0.508*	-0.509*
	(0.261)	(0.261)	(0.262)	(0.262)
$GDP\ growth_{ct}$	-0.007***	-0.007***	-0.007***	-0.007***
	(0.002)	(0.002)	(0.002)	(0.002)
$Inflation_{ct}$	-0.008***	-0.008***	-0.008***	-0.008***
	(0.001)	(0.001)	(0.001)	(0.001)
$Real\ interest\ rate_{ct}$	-0.004**	-0.004**	-0.004**	-0.004**
	(0.002)	(0.002)	(0.002)	(0.002)
$log\ DCPS_{ct}$	0.101***	0.100***	0.099***	0.099***
	(0.030)	(0.030)	(0.030)	(0.030)
$Corruption_{ct}$	-0.067***	-0.068***	-0.068***	-0.068***
	(0.024)	(0.024)	(0.024)	(0.024)
ROL_{ct}	-0.067*	-0.069*	-0.068*	-0.067*
	(0.039)	(0.040)	(0.040)	(0.040)
$NDTT_{ct}$	0.008***	0.008***	0.008***	0.008***
	(0.001)	(0.001)	(0.001)	(0.001)
Entity fixed effects	YES	YES	YES	YES
Year fixed effects	YES	YES	YES	YES
Adjusted R^2	0.898	0.898	0.898	0.898
Observations	24,144,916	24,144,916	24,144,916	24,144,916
Elasticities				
$EMTR_{cit}$	-0.065**			
	(0.027)			
$EATR_{cit}$		-0.070**		
		(0.032)		
τ_{ct}			-0.066**	-0.067**
			(0.032)	(0.032)
δ_{cit}				0.129**
				(0.053)

between different groups of variables. First, *log SALES* and *log STAF* are positively related to investments in fixed assets. These two variables, thus, seem to capture size effects. Second, *CF ratio*, *SALES growth*, and *Profit margin* are all negatively associated with the outcome variable. All three variables may be interpreted as proxies for investment opportunities. In fact, all three variables may be positively correlated with firm age, as well as firm and industry maturity, which explains the negative effect on investment in fixed assets. Third, of the different GDP indicators, it is mainly *log GDP* that has a positive and economically significant impact on investment. Fourth, the negative coefficient on *Inflation* is in line with the investment risk argument presented above. Fifth, the negative coefficient on *Real interest rate* suggests that a high cost of debt financing hampers investment. Sixth, *log DCPS*, a measure of capital market depth, facilitates investment, which is what we expect. We may finally highlight the positive impact of *NDTT*, which confirms earlier findings (see, e.g., Egger and Merlo, 2011).[44]

To test for robustness, we also also run (i) dynamic regressions, (ii) regressions using the gross investment rate as an alternative outcome following the setup used by Liu (2020), as well as (iii) specifications that are based on a balanced panel. This does not substantially change the EMTR effects.[45]

1.6.4. Alternative Fixed Effects Specifications

We now estimate equation (1.10) for alternative fixed effects specifications to test the robustness and sensitivity of the benchmark results. Table 1.4 demonstrates that we find a negative and highly significant tax effect, irrespective of the choice of alternative fixed effects.

The estimates closest to our benchmark result in Table 1.3 are those that condition on

[44] As a general remark, let us add that the estimates are not biased through time-constant country- or industry-specific effects per se, as these are captured by the entity-j-specific fixed effects.

[45] Note that the respective results are available upon request. The estimated coefficients are relatively close to the ones of the preferred model: (i) suggests a short-run effect of -0.217 (std. err.: 0.103) and a long-run effect of -0.49; (ii) a number of results following the specification of Liu (2020) are shown in more detail in Appendix 1.8.10; (iii) leads to a substantially smaller sample and a coefficient of -0.307 (std. err.: 0.174).

Table 1.4: *FIXED EFFECTS SPECIFICATIONS*

The table presents OLS estimates. The dependent variable is the logarithm of firm-entity j's tangible fixed assets, $log\ TFAS_{jt}$. The firm-entity level and country level control variables are the same that are used in Table 1.3, column (1). Robust standard errors are reported in parentheses (clustered at the country-industry-year level). *** denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level. The last rows report elasticities corresponding to the EMTR. Corresponding standard errors are obtained using the Delta method.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$EMTR_{cit}$	-0.402***	-0.485***	-1.523***	-4.752***	-1.847***	-1.449***	
	(0.151)	(0.142)	(0.299)	(0.495)	(0.384)	(0.301)	
$EMTR^A_{cit}$							-0.328***
							(0.037)
Firm-entity level controls	YES	YES	YES	YES	YES	YES	YES
Country level controls	YES	YES	YES	NO	YES	YES	YES
Entity fixed effects (fe)	YES	NO	NO	NO	NO	NO	YES
Group fe	NO	YES	YES	NO	NO	YES	NO
Country fe	NO	NO	NO	NO	YES	YES	NO
Year fe	NO	NO	YES	NO	NO	YES	NO
Industry-year fe	YES	YES	NO	NO	YES	NO	NO
Country-year fe	NO	NO	NO	YES	NO	NO	NO
Country-industry-year fe	NO	NO	NO	NO	NO	NO	YES
Adjusted R^2	0.898	0.855	0.849	0.395	0.447	0.850	0.899
Observations	24,144,916	24,205,343	24,205,343	25,332,567	25,332,663	24,205,341	24,144,274
Elasticity $EMTR_{cit}$	-0.064***	-0.077***	-0.243***	-0.757***	-0.294***	-0.231***	-0.038***
	(0.024)	(0.023)	(0.048)	(0.079)	(0.061)	(0.048)	(0.004)

firm entity (group) as well as industry-year effects, see columns (1) and (2).[46] The largest coefficient is found in specification (4), which is conditional on a country-year-specific fixed effect. Note that country-year-specific EMTRs would not be identified in this specification, but the country-industry-year specific ones ($EMTR_{cit}$) are.

A last, but very powerful test (last column in Table 1.4), relates to an estimate including entity-specific as well as country-industry-year-specific fixed effects.[47] The effect of the EMTR is then only identified when using an interaction term between a time-varying entity-j-specific variable and the EMTR. We thus suggest an alternative firm-specific effective tax measure, which we define as $EMTR^A_{cit} = EMTR_{cit} \times NOLOSS_{jt}$. $NOLOSS_{jt}$ is a time-varying j-specific binary variable indicating whether entity j suffers a loss or not

[46] Note that we identify groups using the information on the global ultimate owner (GUO) of a firm entity that is provided by *Orbis* for a subset of our sample. In the case where no information on the GUO is available, we treat an observed entity as a stand-alone firm.

[47] Of course, this set of fixed effects nests country- and group fixed effects.

($NOLOSS_{jt} = 1$ if a positive value for $EBIT$ is observed, 0 otherwise).

The logic behind this approach is that the EMTR should only have an effect when profits are positive, so that an interaction allows us to identify the EMTR effect. Assuming that tax incentives apply only to firms with positive profits, $EMTR^{A}_{cit}$ is just a version of a firm-specific effective tax measure. The estimate on $EMTR^{A}_{cit}$ equals -0.328 (std. err.: 0.037), which is relatively close to our benchmark estimate and highly statistically significant.

Altogether, the alternative fixed effects specifications suggest the following: It is important to condition on entity-specific heterogeneity; the country-industry-specific EMTRs offer substantial value-added compared to country-year-specific measures; the findings are very robust to various fixed effects specifications.

1.6.5. Heterogeneous Tax Responses

Finally, we exploit the substantial cross-country and industry variation of our new EMTRs to analyze the heterogeneous impact of statutory tax policy changes on the investment behavior of different subgroups of firms. Note that we basically motivate the heterogeneity analysis as well as the definition of subgroups along different contributions to the literature, providing arguments or evidence for heterogeneous tax responses. Let us start with a literature suggesting that *(i)* the tax-responsiveness of investment should be reduced if firms make losses (for similar arguments in the context of financial choices, see Goldbach et al., 2021, or MacKie-Mason, 1990); *(ii)* Egger et al. (2014) show that a small group of *tax-avoiding* multinationals do not respond to taxes at all. This result is consistent with the argument that the tax-sensitivity of investment declines in the extent to which firms are able to avoid being taxed (Goldbach et al., 2019, as well as Overesch, 2009, provide evidence on such effects); *(iii)* the theoretical contribution of Keuschnigg and Ribi (2013) argues that the tax-sensitivity of investment depends on the degree to which a firm is financially constrained; *(iv)* Zwick and Mahon (2017) empirically show that small firms respond more to depreciation incentives compared to large firms, which is in line with the hypothesis

Table 1.5: ***TAX-RESPONSIVENESS FOR DIFFERENT SUBGROUPS***

The table presents OLS estimates. The dependent variable is the logarithm of firm-entity j's tangible fixed assets, $log\ TFAS_{jt}$. The point estimates correspond to firm-entity j-specific subgroups and are estimated using the approach described in Section 1.6.5. In terms of control variables and fixed effects, the setup is identical to Table 1.3, column (1). Robust standard errors are reported in parentheses (clustered at the country-industry-year level). *** denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level. Note that except for the specification *Profitable firm entity*, we use samples that exclude firm entities j with a non-positive EBIT in more than 25% of the time in our panel. The subgroups are defined as follows. *Profitable firm entity*: Firm-entity j reports strictly positive EBIT in at least 75% over time. *Stand-alone firm*: j is not part of a group (note that only firm entities with information on the global ultimate owner are considered for this regression, which reduces the sample size considerably). *Young firm-entity age*: j's age (age is calculated as difference between current year and the year of incorporation) is lower than the 25 percentile of the age variable of the overall sample. *Manufacturing*: j operates in the section *C Manufacturing*. *Transportation and storage*: j operates in the section *H Transportation and storage*. *Construction*: j operates in the section *F Construction*. *Wholesale and retail*: j operates in the section *G Wholesale and retail trade; repair of motor vehicles and motorcycles*. *High GDP growth*: j is located in a country where half or more than half of the country-specific entity-year observations exhibit a GDP growth rate that is equal to or higher than the 75 percentile of the GDP growth rate of the overall sample. *Weak capital market*: j is located in a country where more than half of the country-specific entity-year observations exhibit a logarithm of the domestic credit to the private sector as share of the GDP that is lower than the 25 percentile of this variable of the overall sample. *Low GDP per capita*: j is located in a country where more than half of the country-specific entity-year observations exhibit a GDP per capita that is lower than the 25 percentile of this variable of the overall sample.

	Subgroup semi-elasticity	(se)
Profitable firm entity	-1.904***	(0.331)
Stand-alone firm	-1.243***	(0.265)
Young firm-entity age	-1.418***	(0.238)
Manufacturing	-1.226**	(0.567)
Transportation and storage	-2.628***	(0.798)
Construction	-1.130**	(0.474)
Wholesale and retail	-0.165	(0.218)
High GDP growth	-1.715***	(0.352)
Weak capital market	-1.543***	(0.321)
Low GDP per capita	-1.508***	(0.327)

that small firms are often financially constrained; *(v)* Overesch and Wamser (2009) suggest that the tax-elasticity of foreign direct investment (FDI) depends on the type of FDI, the underlying business model, as well as the internationalization of a multinational group (see also Stöwhase, 2005). We may thus focus on different industries, which we expect to be more or less tax-responsive.

Note that it is not a particular goal of our analysis to learn about a specific heterogeneity. We want to document, however, that we can adequately capture heterogeneous tax-responses using our new tax data in combination with a large micro-level dataset. As the following will show, the heterogeneous effects we find are consistent with what has been shown before in the above-mentioned literature. The subgroups that we use for the heterogeneity analysis are defined according to industry-, country- or firm-characteristics. For the purpose of this heterogeneity analysis, we introduce indicator variables for a specific subgroup and then, based on our large sample, report only the results from the interaction terms for the specific group we are interested in.[48]

Table 1.5 depicts the results for the different subgroups. For the precise definitions of the subgroup indicators, see also the table notes. We find a semi-elasticity of -1.90 for firm entities that report strictly positive profits in most years. This estimate is significantly larger than the benchmark estimate of -0.41, which is consistent with *(i)*.[49] Next, we find that stand-alone firms are more responsive, which is in line with the arguments presented in *(ii)* and *(iv)*. Furthermore, we find that comparably younger firm entities as well as firm entities located in countries with weak capital markets, high GDP growth, and countries with low GDP per capita are more responsive to tax policy changes compared to the benchmark result. These results may be explained with the financial constraint arguments *(iii)* and *(iv)*. Finally, we find that different industries respond differently to changes in the EMTR, as suggested in *(v)*. In detail, we find that the most responsive industry is *Transportation and storage* with a statistically significant EMTR semi-elasticity of about -2.6. The *Manufacturing*- and *Construction*-industry entities are less than half as sensitive but also statistically significant with an EMTR semi-elasticity of about -1.2 and -1.1, respectively. Tax incentives matter less for entities in *Wholesale and retail trade* with a coefficient of -0.17, which is statistically insignificant.

Let us finally focus on a specific type of heterogeneity which we find particularly interesting. It relates to a large literature showing that some firms can avoid taxes by relocating profits to low-tax countries (see the reasoning in *(ii)*).[50] In our basic analysis we include three types of firm entities: stand-alone entities, entities that belong to a domestic firm group, and entities that belong to a multinational firm group. While we exploit this information to

[48]Complete estimation results are available upon request.

[49]Note that since profitability is a key factor for explaining tax-responses, we carry out the remainder of the subgroup analysis using samples that only include firm entities that are mostly profitable

[50]For a recent contribution quantifying profit-shifting activities of multinationals, see Tørsløv et al. (2023).

estimate a coefficient on those that are not part of a firm group above in Table 1.5, the idea is now to examine how firm entities respond to changes in the EMTR given the group-wide (or multinational-firm-wide) minimum statutory corporate tax rate ($Minimum\ tax_j$).[51] For stand-alone entities and national groups, this minimum tax rate equals the statutory rate of the country that they are located in.[52] For an entity that is part of a multinational group, $Minimum\ tax_j$ is calculated by taking the lowest tax rate among all countries that the multinational is operating in according to *Orbis*. The reasoning behind this approach is that multinationals are generally able to shift profits to entities located in low-tax countries to avoid taxes. Therefore, we expect that those entities facing a very high "minimum" tax should be more responsive, compared to others where $Minimum\ tax_j$ is relatively low. The latter suggests that these firms have access to a low-tax country and may shift profits towards related entities in this low-tax country. This, in turn, reduces the cost of capital in the (possibly high-tax) host country (for similar arguments on this mechanism, see Suárez Serrato, 2019). Alternatively, if the entity is itself the low-tax affiliate, then it faces a very low corporate tax rate. Figure 1.3 plots semi-elasticities for various values of $Minimum\ tax_j$.

The pattern we find is highly consistent with the profit-shifting argument. It seems that the negative tax effect only kicks in when the minimum tax rate is 24% or higher. For estimates where the minimum tax is lower, the estimated coefficients are close to zero and statistically insignificant. The increase in tax-responsiveness then increases in the minimum tax (in a not fully monotonic way).[53]

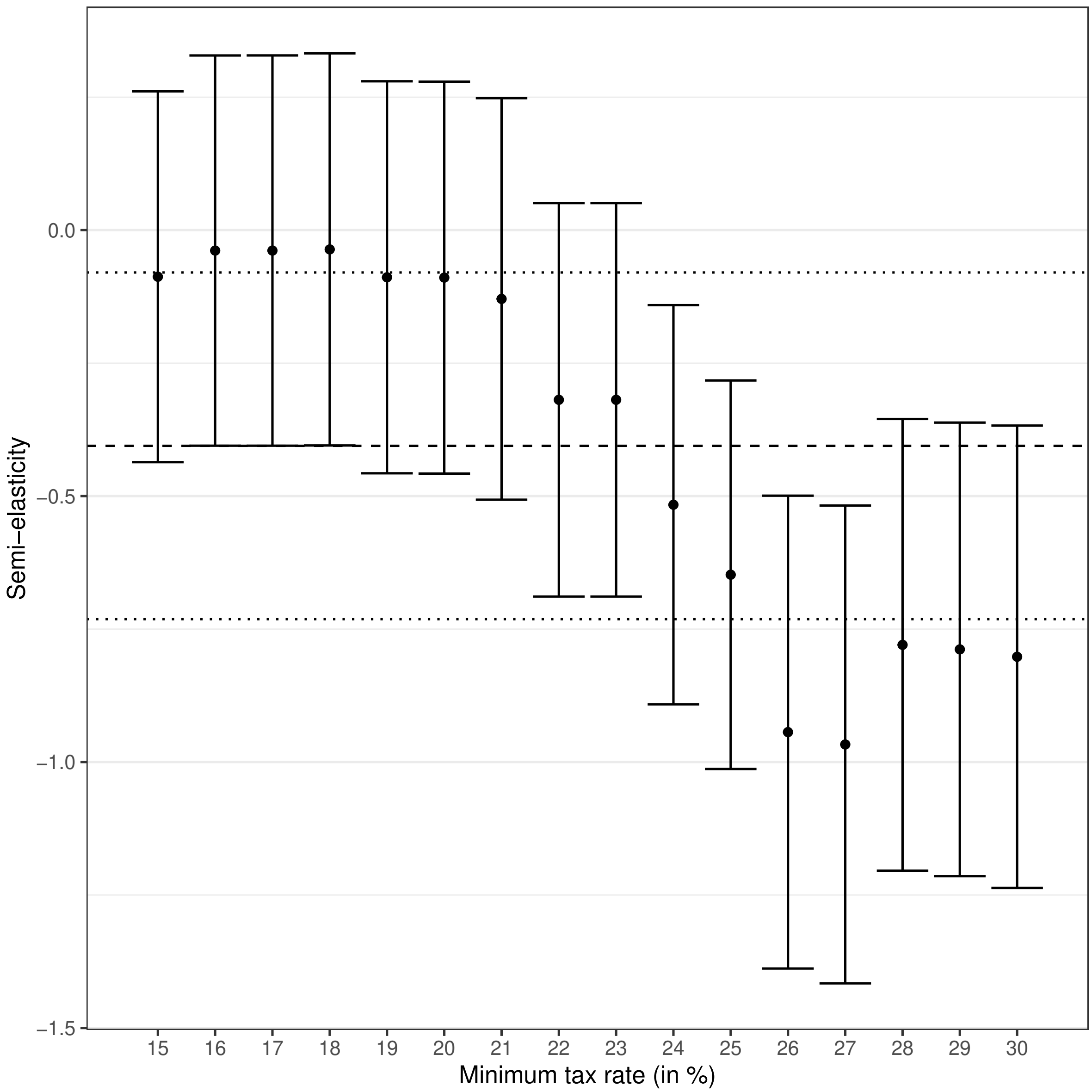
[51]Note that we calculate the minimum for the whole firm group (all entities observed) over all years in our sample.

[52]For firm entities for which we do not have any information on the global ultimate owner, we set the minimum tax rate as if they were stand-alone entities.

[53]Note that we cannot estimate the EMTR responses for groups where the minimum tax is below 15% or above 30% in a sufficiently precise way. Group sizes are too small and the variation over time in EMTRs is limited, introducing too much noise.

Figure 1.3: *TAX-RESPONSIVENESS AND MINIMUM TAX RATES*

The figure presents OLS estimates on $EMTR_{cit}$, and the corresponding 95% confidence intervals. The dependent variable is the logarithm of firm-entity j's tangible fixed assets, $log\, TFAS_{jt}$. The point estimates correspond to subgroups of firm entities that are exposed to a within-firm minimum statutory tax rate that is equal to or higher than the tax rate depicted on the horizontal axis. For the definition of the minimum tax rate, see Section 1.6.5. The confidence intervals are based on robust standard errors (clustered at the country-industry-year level). In terms of control variables and fixed effects, the setup is identical to Table 1.3, column 1. The dashed line gives the semi-elasticity of the benchmark model (Table 1.3, column 1) and the dotted lines the corresponding confidence bounds.



1.7. Conclusions

This paper provides a new approach to calculate country-industry-year-specific forward-looking effective tax rates (FLETRs) for 19 industries, 221 countries and the years 2001 to 2020. Besides statutory tax rate and tax base information, the FLETRs account for typical country-industry-specific financing- and asset structures. These financing- and asset structures are – depending on the data coverage – calculated directly from different data sources or imputed using Predictive Mean Matching. By accounting for the heterogeneity in financing and asset structures, we ensure that our FLETRs adequately capture the variation in the tax incentives that a country's tax code implicitly grants to different industries. We further demonstrate that other commonly used effective tax rate measures suffer from significant measurement error when this country-industry-specific heterogeneity is neglected. Our empirical analysis exploits the substantial variation in FLETRs over time to provide estimates of the tax semi-elasticity of investment. Based on more than 24 million firm-entity observations, our results suggest a tax semi-elasticity of about -0.41, which is at the lower end of previous findings. An interesting additional test focuses on the effect of industry-specific depreciation allowances. Compared to tax cuts, our estimates suggest that firms are very sensitive to changes in depreciation rules. When a government's objective is to stimulate investments, more generous depreciation allowances may thus be the more effective policy instrument.

We further illustrate that different subgroups of firms respond very heterogeneously to tax incentives. For example, when focusing on firm entities operating in the manufacturing sector, we find a substantially bigger semi-elasticity of -1.23. Country-specific economic circumstances as well as profit shifting opportunities also have a significant impact on the tax semi-elasticity. All in all, the estimated semi-elasticities range from values close to zero up to -2.63.

Our study implies that policymakers should be careful when designing tax reforms or when using incentives such as bonus depreciation programs to stimulate corporate investment. The extent to which this leads to more real firm activity depends significantly on the type of business and several other firm- and/or country-specific conditions.

1.8. Appendix

1.8.1. Derivation of the EATR

This section briefly outlines the calculation of the forward-looking effective average tax rate (EATR). For the calculation of the EATR we follow Devereux and Griffith (2003) and Steinmüller et al. (2019). The EATR depicts the effective tax burden of all infra-marginal units invested in a hypothetical investment project. It is the scaled difference between the pre-tax net present value, R^*, and the post-tax net present value, R, of the hypothetical investment that has a given pre-tax rate of return p. This results in a *tax wedge*, reflecting the excess return to investment necessary to compensate for taxation. To obtain the EATR, the tax wedge is divided by the discounted rate of return (using the market interest rate for equity, i, for discounting), yielding

$$EATR_{cit} = \frac{R^* - R}{p/(1+i)} = \frac{\tau\big(p - i\delta\big)}{p}. \tag{1.11}$$

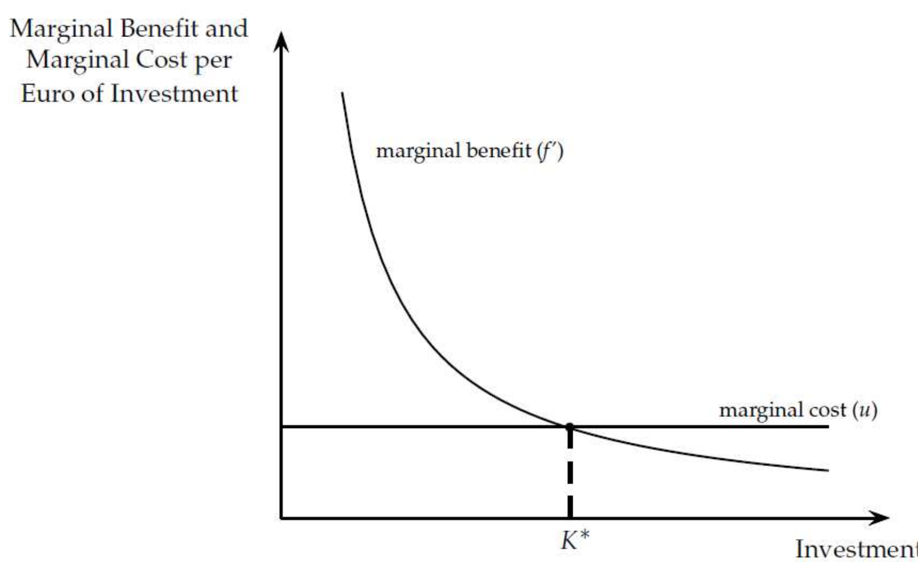
Here, τ represents the statutory corporate tax rate and δ the NPV of depreciation allowances. From (1.11), it is evident that the NPV of depreciation allowances is less relevant for the size of the EATR compared to the EMTR. In fact, the size of the EATR crucially depends on the profitability of the investment as well as the statutory corporate tax rate (see also Devereux and Griffith, 2003). Country-industry-year-specific EATRs can then be calculated using the country-industry-year-specific NPVs of depreciation allowances, δ_{cit}, that are formally depicted in Section 1.2 of the main text:

$$EATR_{cit} = \frac{R^*_{cit} - R_{cit}}{p/(1+i)} = \frac{\tau_{ct}\big(p - i\delta_{cit}\big)}{p}. \tag{1.12}$$

1.8.2. Derivation of the EMTR

This section briefly outlines the calculation of the forward-looking effective marginal tax rate (EMTR). Suppose a firm produces output following the production function $f(K)$ (with properties $f'(K) > 0, f''(K) < 0$) using capital K as the only input. Output is strictly increasing in K, for example investment in machinery, with $\partial f(K)/\partial K > 0$ denoting the marginal product of K. A profit-maximizing firm in a perfectly competitive environment compares marginal benefit of additional investment to marginal cost and increases or decreases K until the two equalize. Let us denote the marginal cost by $u = \sigma + i$, where σ is the economic depreciation rate of K, and i is the cost of equity.[54] We may interpret i as the after-tax return of a risk-free investment and, thus, as opportunity cost.[55] By assuming decreasing returns (a diminishing marginal product) to investment, the profit-maximizing investment K^* is determined by setting marginal benefit equal to marginal cost, i.e., $f'(K) = u$. Thus, in the absence of taxes, optimal investment is given by K^* (see Figure 1.4).

Figure 1.4: *OPTIMAL INVESTMENT*

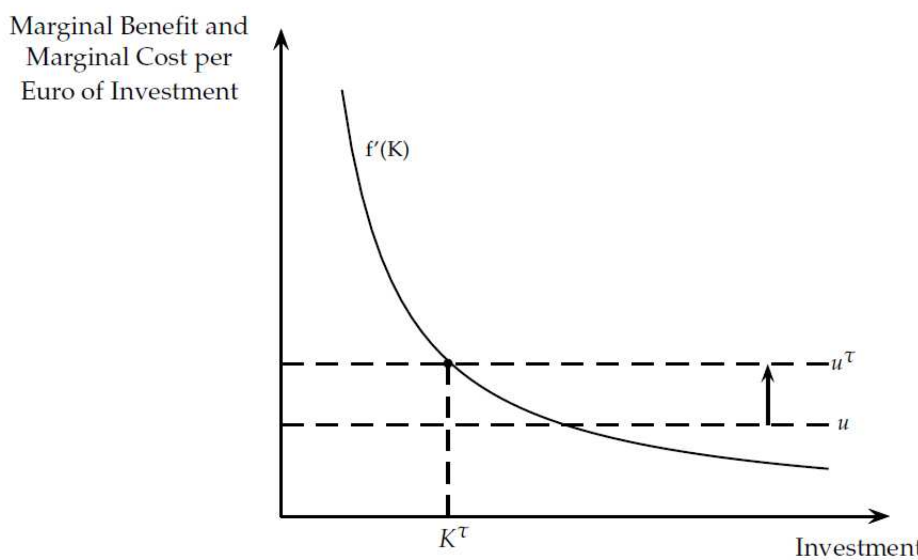


[54] Here, K is equity financed and financing cost are fully taxed (not deductible for tax purposes); prices are kept constant and normalized to one.

[55] We may think of i also as a dividend payout. Note, however, that we are interested in calculating the effective tax burden at the corporate level. Thus, we are abstracting from taxes on dividends.

Introducing a tax τ in this simple model implies that some output is taxed away and the marginal earnings per unit of investment reduce to $f'(K)(1-\tau)$. This suggests a parallel downward shift in the marginal benefit curve and a new equilibrium where investment falls to K^τ, as illustrated in Figure 1.5. Solving for $f'(K)$, we obtain $f'(K) = \frac{1}{(1-\tau)}(\sigma + i) \equiv u^\tau$. Note that the expression on the right-hand side of the equation, u^τ, is the *user cost of capital.* With $\tau \in (0,1)$, the tax increases the required rate of return such that $u^\tau > u$. In order for the new optimality condition to hold, the firm invests less ($K^\tau < K$), leading to an increase of $f'(K)$ by a sufficient amount to just break even. The reduction in K and the concavity of the production function ensure that the pre-tax return with taxation is higher so that the firm is not making a loss.

Figure 1.5: ***OPTIMAL INVESTMENT WITH TAXATION***



We can now account for the fact that governments typically grant tax deductions for the cost of financing and depreciation by introducing *depreciation allowances* into this model. While we only consider the period of the investment, investments generate future returns, and machines or other investments depreciate over time. Accordingly, we need to account for the future stream of depreciation allowances by considering the net present value (NPV) of depreciation allowances, which we denote by δ. Depreciation allowances reduce a firm's tax base, suggesting that for each unit of depreciation allowance subtracted from the tax base, the tax payment equals zero. Thus, there is a tax saving of $\tau \cdot \delta$ per unit of investment. Consequently, the depreciation allowance reduces the user cost of capital:

$$\widehat{u}^{\tau} = \frac{1}{(1-\tau)}(\sigma + i) \cdot (1 - \tau\delta).$$

Note that in a graphical illustration, this would shift the horizontal line of the user cost down. As Figure 1.5 illustrates, a corporate tax τ drives a *wedge* between marginal benefit and marginal cost. The effective marginal tax rate (EMTR) is a measure of the relative size of this tax wedge between user cost of capital with and without taxation. Formally, we thus have

$$EMTR = \frac{\widehat{u}^{\tau} - u}{\widehat{u}^{\tau}} = \frac{\frac{1}{(1-\tau)}(\sigma + i) \cdot (1 - \tau\delta) - (\sigma + i)}{\frac{1}{(1-\tau)}(\sigma + i) \cdot (1 - \tau\delta)} = \frac{(\tau - \tau\delta)}{(1 - \tau\delta)}. \tag{1.13}$$

1.8.3. NACE Rev. 2 (ISIC Rev. 4) Section Descriptions

Table 1.6: ***NACE REV. 2 (ISIC REV. 4) SECTION DESCRIPTIONS***

The table depicts the descriptions of the sections of the *Statistical Classification of Economic Activities in the European Community (NACE) Rev. 2* and the *International Standard Industrial Classification of All Economic Activities (ISIC) Rev. 4* that are used throughout this paper. Note that since NACE Rev. 2 was created based on ISIC Rev. 4, the classification systems are equal at the section level.

Section code	section description
A	Agriculture, forestry and fishing
B	Mining and quarrying
C	Manufacturing
D	Electricity, gas, steam and air conditioning supply
E	Water supply; sewerage, waste management and remediation activities
F	Construction
G	Wholesale and retail trade; repair of motor vehicles and motorcycles
H	Transportation and storage
I	Accommodation and food service activities
J	Information and communication
K	Financial and insurance activities
L	Real estate activities
M	Professional, scientific and technical activities
N	Administrative and support service activities
O	Public administration and defence; compulsory social security
P	Education
Q	Human health and social work activities
R	Arts, entertainment and recreation
S	Other service activities
T	Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use
U	Activities of extraterritorial organizations and bodies

1.8.4. Structure and Preparation of the Eora26 Database

This section provides additional details on the *Eora26* database and how we utilize the data for the purpose of our paper. To get a deeper understanding of the structure of the data, we start out by describing the *Eora Global Supply Chain* database (Lenzen et al., 2012; Lenzen et al., 2013), from which *Eora26* is derived. At the centre of the *Eora Global Supply Chain* database are the yearly multi-region input-output tables (MRIOs). For the countries in the MRIOs, generally either commodities or industries are included, but not both. This results in a mix of different input-output (IO) tables. In detail, three different types of IO tables are distinguished: Industry-by-Industry IO tables, Commodity-by-Commodity IO tables, and Supply-Use tables (SUTs). The latter category includes Commodity-to-Industry as well as Industry-to-Commodity transactions.[56] Furthermore, the industry and commodity classification systems that are used differ strongly between countries. To facilitate between-country analyses, a simplified version of the *Eora* MRIOs is provided, the so-called *Eora26* MRIOs. In this version, all industries and commodities are aggregated to a common 26-sector classification and the SUTs from the full resolution *Eora* MRIOs are converted to symmetric sector-by-sector IO tables using the *Eurostat manual of supply, use and input-output tables* (2008).[57] For our purpose, we translate this 26-sector classification of the *Eora26* database to the NACE Rev. 2 (ISIC Rev. 4) sections that we use throughout the paper. In doing this, we rely on the concordance table provided on the webpage of the *Eora26* database that documents how the different industries and commodity categories from the full resolution *Eora* were transformed to the 26-sector system of *Eora26*.[58] More precisely, we string-search the industry descriptions of the full resolution *Eora* database for the closest matches to the different NACE Rev. 2 (ISIC Rev. 4) section descriptions. Then, we look at how a chosen industry from the full *Eora* was converted to the 26-sector system and reverse this transformation for all countries. The precise assignment is depicted in Table 1.7.

[56] For a graphical illustration of the MRIO layout, see Lenzen et al. (2013, p. 25).

[57] Eurostat (2008). Eurostat manual of supply, use and input-output tables. Office for Official Publications of the European Communities. Eurostat methodologies and working papers. Luxembourg. For more details, see the webpage of Eora26, *https://worldmrio.com/eora26/*.

[58] See *https://worldmrio.com/eora26/*.

Table 1.7: *CONCORDANCE OF EORA26 SECTORS TO NACE REV. 2 (ISIC REV. 4) SECTIONS*

The table depicts the assignment that is used to translate the 26-sector classification of the *Eora26* database to the NACE Rev. 2 (ISIC Rev. 4) sections that we use throughout this paper. The aggregation is based on the concordance table that translates the different industry and commodity categories of the full *Eora* to the 26 sectors used in *Eora26* which can be found on the website of the *Eora26* database (*https://worldmrio.com/eora26/*). Descriptions for the different NACE Rev. 2 (ISIC Rev. 4) sections are provided in Table 1.6.

NACE/ISIC	Eora26 sector(s)
A	0.873377 · Agriculture +
	0.126623 · Fishing
B	Mining and Quarrying
C	0.089343 · Food & Beverages +
	0.181663 · Textiles and Wearing Apparel +
	0.045522 · Wood and Paper +
	0.246543 · Petroleum, Chemical and Non-Metallic Mineral Products +
	0.122740 · Metal Products +
	0.229526 · Electrical and Machinery +
	0.025101 · Transport Equipment +
	0.043395 · Other Manufacturing +
	0.016167 · Recycling
D	Electricity, Gas and Water
E	0.181818 · Electricity, Gas and Water +
	0.818182 · Education, Health and Other Services
F	Construction
G	0.023499 · Maintenance and Repair +
	0.302872 · Wholesale Trade +
	0.673629 · Retail Trade
H	Transport
I	Hotels and Restaurants
J	Post and Telecommunications
K	Finacial Intermediation and Business Activities
L	Finacial Intermediation and Business Activities
M	Finacial Intermediation and Business Activities
N	Finacial Intermediation and Business Activities
O	Public Administration
P	Education, Health and Other Services
Q	Education, Health and Other Services
R	Education, Health and Other Services
S	0.071197 · Education, Health and Other Services +
	0.928803 · Others
T	Private Households
U	Others

1.8.5. Assignment of EUKLEMS & INTANProd Asset Types

Table 1.8: ***ASSIGNMENT OF EUKLEMS & INTANProd RELEASE 2021 ASSET TYPES TO THE ASSET TYPES USED FOR CALCULATIONS OF FLETRs***

The table depicts the assignment of the asset categories from the *EUKLEMS & INTANProd* release 2021 to the asset categories used for the calculations of FLETRs in this paper (excluding the asset type *Inventory*).

Asset type	Assigned EU Klems 2019 asset types
Buildings	N111 Dwellings
	N112 Other buildings and structures
Computer equipment	N11321 Computer hardware
Intangible fixed assets	N1171 Research and development
	N1173 Computer software and databases
Machinery	N11O Other machinery equipment and weapons systems
Office equipment	N11322 Telecommunications equipment
Vehicles	N1131 Transport equipment

1.8.6. Imputation

This section provides additional details on the PMM imputation. In Section 1.4.3, we impute a missing year-specific weight using the observed value corresponding to the data point (the so-called donor) for which the predicted value is closest to the predicted value of the missing data point that we were looking to impute. Alternatively, instead of using just one donor, the mean of the $d > 1$ donors that are closest may be chosen (Van Buuren, 2018). In the extreme case of setting d to the number of available complete cases, one would obtain identical imputed values for all missing data points, that is, the mean over all donor candidates. To not lose variation among the imputed values, typically small d's are chosen.[59] In Figure 1.6, we depict asset weight structures for the section *C Manufacturing* that are imputed as described in Section 1.4.3 but with a varying number of donors d.[60]

It is evident that the imputed asset structures look similar for $d = 1$, 5, 10, and 15. In particular, the reduction in variation between countries when increasing d is small. We therefore conclude that our imputation results are robust to other commonly used choices for the number of donors d.

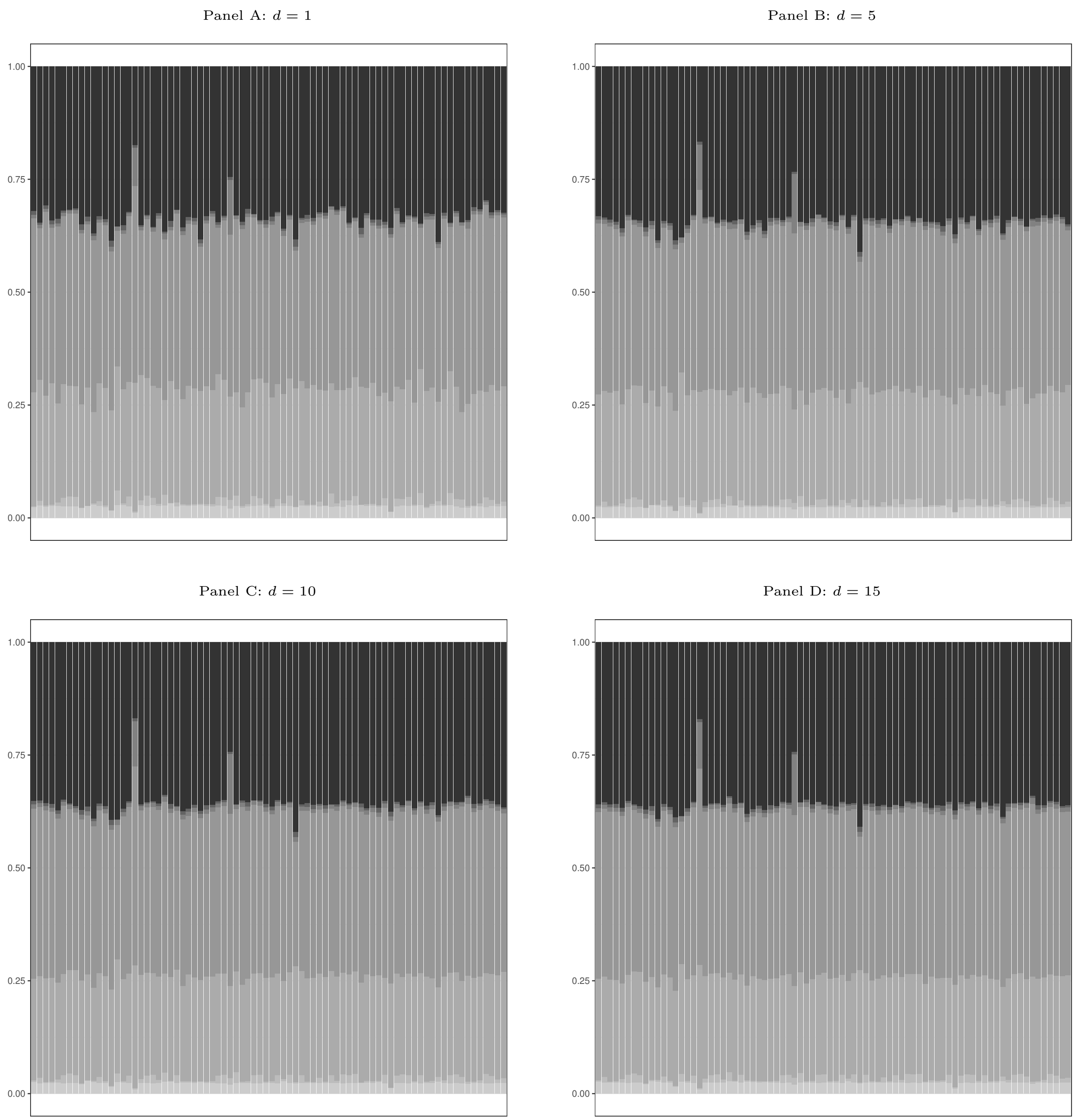
An algorithm that is heavily used in the literature for matching purposes is k-Nearest Neighbor (k-NN) matching (see, e.g., Hastie et al., 2009, ch. 13.3). With k-NN matching, each covariate used for the matching is standardized to have an overall mean of zero and variance of one. A missing data point is then imputed with the mean of the k observed data points for which the Euclidean distance of the covariates to those of the missing is minimal. The key difference between PMM and k-NN matching is that PMM takes into account the importance of each covariate for predicting the dependent variable, whereas k-NN matching assigns each covariate the same weight. For the sake of completeness, we carry out the imputation of asset structures for the sector *C Manufacturing* with k-NN matching using

[59] See Van Buuren (2018) for a thorough literature review on the optimal choice of donors.

[60] Note that the same countries are depicted in the same order as in Panel B of Figure 1.8.

Figure 1.6: ***COUNTRY-SPECIFIC ASSET STRUCTURES OF SECTION C MANUFACTURING IMPUTED USING PMM WITH DIFFERENT NUMBER OF DONORS***

The figure depicts asset structures for the NACE Rev. 2 (ISIC Rev. 4) section *C Manufacturing* by country. The structures of the depicted countries are fully imputed using PMM (see Section 1.4.3). The panels correspond to imputation using a different number of donors d. Each bar corresponds to the asset structure of a different country. The order of the countries is the same in all panels and identical to the one in Panel B of Figure 1.8. The asset types are indicated by the different shadings of the bars. The asset types are – from dark to bright shading – as follows: *Buildings, Computer equipment, Intangible fixed assets, Inventory, Machinery, Office equipment, and Vehicles.*



the same covariates that we used with PMM.[61] The results are depicted in Figure 1.7.

It is evident that, irrespective of k, the imputed asset structures are often identical or extremely similar between countries. Furthermore, it shows that the imputed asset structures seem to be highly dependent on the chosen k, as the amount of variation between countries decreases strongly as k is increased.

Finally, let us also provide some more in-depth plausibility checks of the country-industry-specific asset structures of section *C Manufacturing*. Panel A in Figure 1.8 depicts the asset structures of countries that are fully covered by the primary data sources. Panel B in Figure 1.8 shows the asset structures of countries that were entirely imputed using the PMM procedure.

Comparing these two figures, it can be seen that the imputed results exhibit somewhat less variation. However, as shown in Table 1.1 in the main text, this is not a result that is representative of the imputation of all weights in all industries. In fact, there are several sections where there is more variation among the group of PMM imputed countries than in the one with observed data. One country that stands out in Panel B of Figure 1.8 is Canada that exhibits the lowest share of buildings among the depicted imputed countries. Taking a look at Panel A of Figure 1.8, it can be seen that the imputed asset structure of Canada is similar to the asset structures of other highly developed nations such as Germany, France, Japan, the Netherlands, Sweden, or the USA. Conversely, the imputed less developed countries in Panel B of Figure 1.8 exhibit similarities to the less developed countries that were used for the matching, such as Lithuania or Slovakia.

Another interesting observation that can be made when looking at Panel B of Figure 1.8 is the fact that the different PMM imputed asset structures appear to not be identical. This indicates that the unique weighting of the covariates for the matching of each asset type yielded a mix of different donors matched for the imputation of a single asset structure.

[61]Note, however, that we do not include time dummies, as k-NN matching does not allow for categorical variables. Further note that no logs of the variables are taken.

Figure 1.7: *COUNTRY-SPECIFIC ASSET STRUCTURES OF SECTION C MANUFACTURING IMPUTED USING k-NN WITH DIFFERENT NUMBER OF k*

The figure depicts asset structures for the NACE Rev. 2 (ISIC Rev. 4) section *C Manufacturing* by country. The structures of the depicted countries are fully imputed using k-NN (see Hastie et al., 2009, ch. 13.3). The panels correspond to imputation using a different number of neighbors k. Each bar corresponds to the asset structure of a different country. The order of the countries is the same in all panels and identical to the one in Panel B of Figure 1.8. The asset types are indicated by the different shadings of the bars. The asset types are – from dark to bright shading – as follows: *Buildings, Computer equipment, Intangible fixed assets, Inventory, Machinery, Office equipment, and Vehicles.*

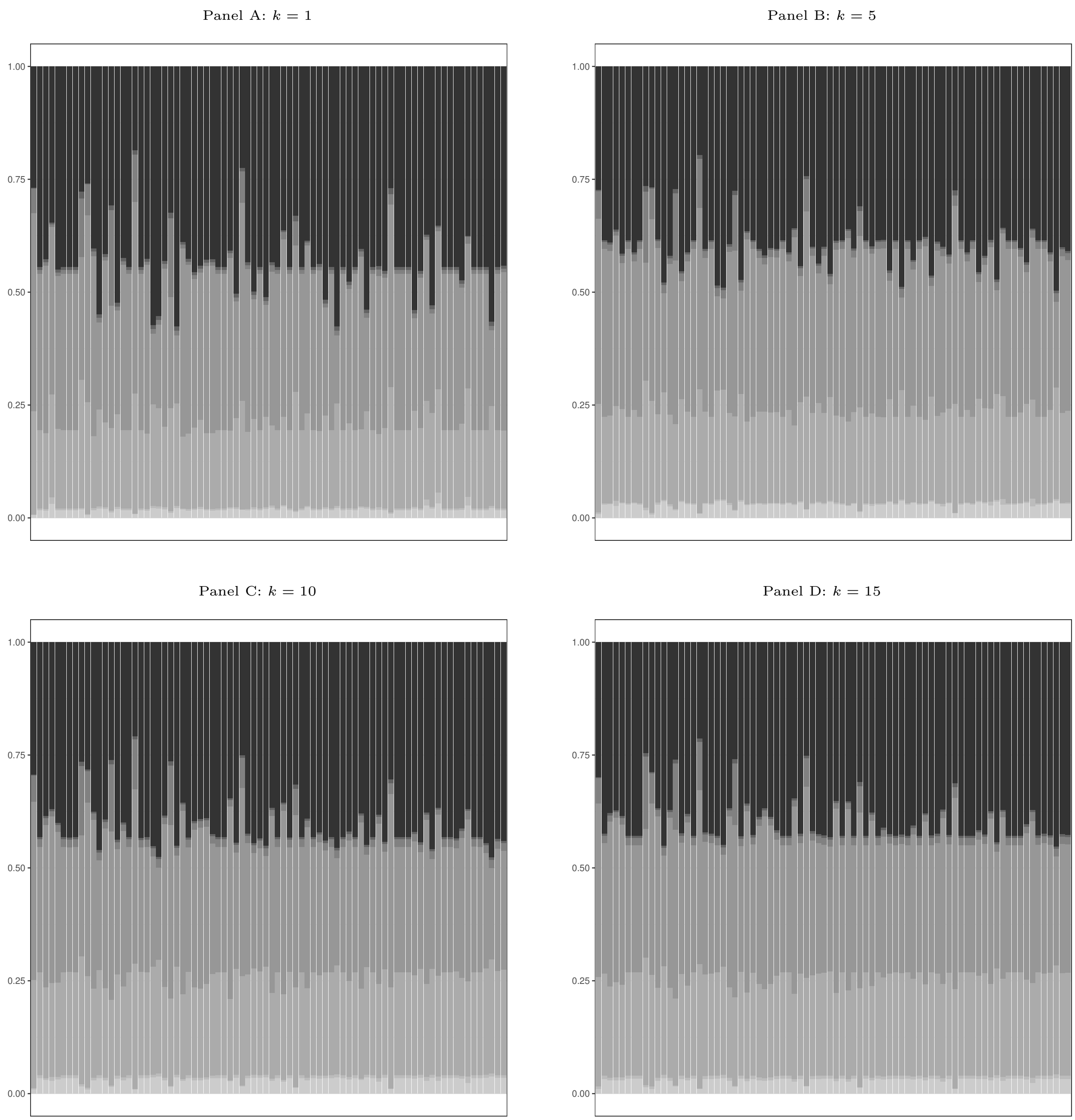
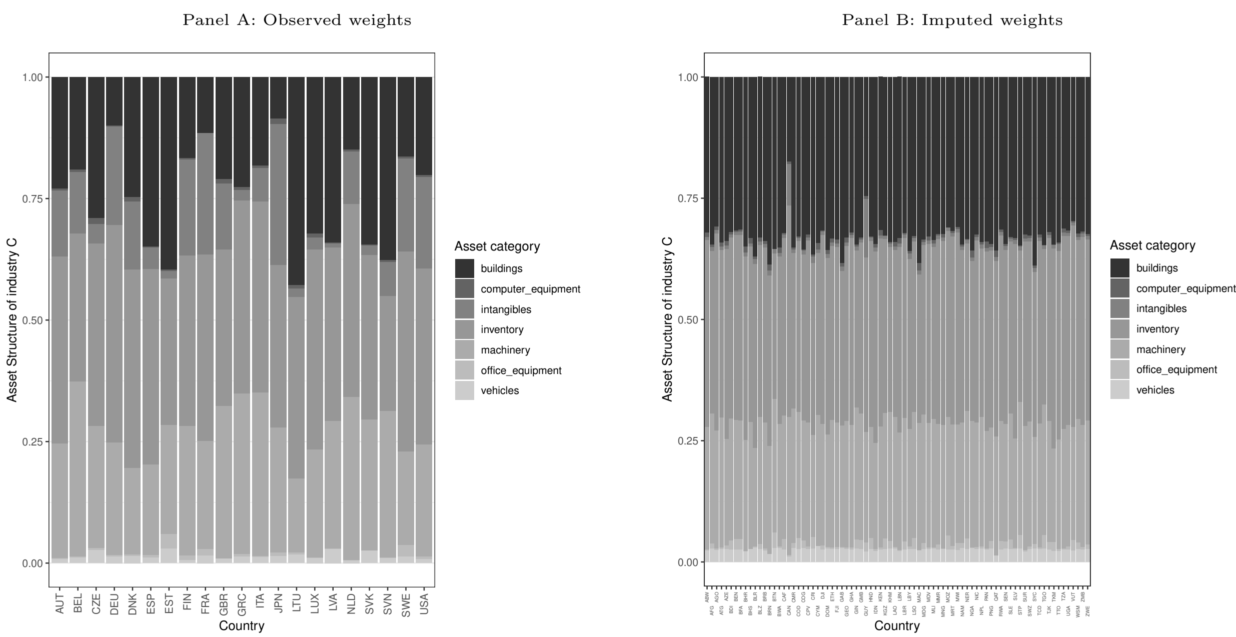


Figure 1.8: *COUNTRY-SPECIFIC ASSET STRUCTURES FROM PRIMARY DATA SOURCES OF SECTION C MANUFACTURING*

The figure depicts asset structures for the NACE Rev. 2 (ISIC Rev. 4) section *C Manufacturing* by country. In Panel A, the structures of the depicted countries are fully derived from the primary data sources *EUKLEMS & INTANProd* release 2021 and *Orbis* (see Section 1.4.2). In Panel B, the structures of the depicted countries are fully imputed using PMM with $d = 1$ donor (see Section 1.4.3).



1.8.7. Descriptive Statistics of Imputation Covariates

Table 1.9: *DESCRIPTIVES ON COVARIATES USED FOR IMPUTATION*

The table depicts descriptive statistics on all the covariates used for the imputation of financing (Panel A) and asset structures (Panel B). The time span that is covered in the sample is 2001 to 2016. Definitions of all variables are provided in Section 1.3.

Panel A: Covariates for financing structure imputation (36,214 observations)		
	Mean	(sd)
τ_{ct}	0.253	(0.094)
ROL_{ct}	0.077	(0.995)
$Corruption_{ct}$	0.094	(1.022)
$log\ DCPS_{ct}$	3.607	(0.980)
$Inflation_{ct}$	5.417	(10.570)
$GDP\ growth_{ct}$	3.638	(4.410)
$log\ GO_{cit}$	15.229	(2.421)
$log\ GI_{cit}$	15.174	(2.440)
$log\ COE_{cit}$	13.519	(2.817)
$log\ NOS_{cit}$	13.055	(3.103)
$log\ NMI_{cit}$	9.335	(4.969)
$log\ net\ TOP_{cit}$	10.217	(3.260)
$log\ COFC_{cit}$	11.954	(3.180)
Panel B: Covariates for asset structure imputation (49,811 observations)		
	Mean	(sd)
$log\ GDP_{ct}$	25.152	(2.006)
$log\ GDP\ p.c._{ct}$	9.288	(1.218)
$log\ COE_{cit}$	13.355	(2.689)
$log\ NOS_{cit}$	12.902	(2.991)
$log\ NMI_{cit}$	9.514	(4.834)
$log\ COFC_{cit}$	11.849	(3.017)
$log\ CO2_{cit}$	6.423	(2.527)

1.8.8. Imputation of Countries without Covariate Data

Table 1.10: ***IMPUTATION OF COUNTRIES WITHOUT COVARIATE DATA***

The table depicts the assignment of countries for which we obtain weights (either directly through data sources or through the imputation algorithm) to countries for which we do not obtain weights. If two more countries are assigned, then the unweighted average of these countries' weights are used for imputation.

Panel A: Financing structure (53 countries with missing weights)

Countries with missing weights	Countries used for imputation
AIA;ANT;BES;CUW;CYM;DMA;GLP;GRD;KNA;LCA;MSR;MTQ;PRI;SXM; TCA;VCT;VGB;VIR	ABW;ATG;BHS;BRB;DOM;JAM;TTO
ASM;COK;FSM;KIR;MHL;MNP;NCL;NIU;NRU;PLW;PYF;SLB; TLS;TON	AUS;FJI;NZL;PNG;VUT;WSM
AND	ESP; FRA
ARG	BOL; BRA; CHL; PRY; URY
BLZ	GTM; MEX
BMU	ABW;ATG;BHS;BRB;DOM;JAM;TTO;USA
COM	MDG; MDV; MUS; SYC
ERI	DJI; ETH; SDN
GGY	FRA; GBR
GIB	ESP
GNB	SEN; GIN
GNQ	GAB; CMR
GRL	CAN; ISL
IMN	GBR
JEY	FRA; GBR
LIE	CHE; AUT
MCO	FRA
PRK	CHN; KOR
SMR	ITA
TKM	AFG; IRN; KAZ
UZB	AFG; KAZ; KGZ; TJK
XKX	SRB
YUG	MNE; SRB

Panel B: Asset structure (56 countries with missing weights)

Countries with missing weights	Countries used for imputation
AIA;ANT;BES;CUW;DMA;GLP;GRD;KNA;LCA;MSR;MTQ;PRI;SXM; TCA;VCT;VGB;VIR	ABW;ATG;BHS;BRB;CYM;DOM;JAM;TTO
ASM;COK;FSM;KIR;MHL;MNP;NCL;NIU;NRU;PLW;PYF;SLB; TLS;TON	AUS;FJI;NZL;PNG;VUT;WSM
AND	ESP; FRA
COM	MDG; MDV; MUS; SYC
ERI	DJI; ETH
GGY	FRA; GBR
GIB	ESP
GNB	SEN; GIN
GNQ	GAB; CMR
GRL	CAN; ISL
IMN	GBR
JEY	FRA; GBR
LIE	CHE; AUT
MCO	FRA
MKD	ALB; BGR; GRC; SRB
MNE	ALB; BIH; HRV; SRB
PRK	CHN; KOR
PSE	EGY; ISR; JOR
SDN	CAF; EGY; ETH; LBY; TCD
SMR	ITA
SSD	CAF; COD; ETH; KEN; UGA
SYR	IRQ; ISR; JOR; LBN; TUR
TWN	CHN; JPN; KOR; PHL
VEN	BRA; COL; GUY
XKX	SRB
YEM	OMN; SAU
YUG	SRB

1.8.9. Country-Industry-Year-specific EATRs

The country-industry-year-specific EATRs are obtained by plugging the country-year-specific statutory tax rate, τ_{ct}, as well as the country-industry-year-specific NPVs of depreciation allowances, δ_{cit}, which are calculated using the financing and asset weights from Section 1.4, into equation (1.12). Note that for the parameterization of the pre-tax rate of return, p, and the market interest rate, i, we follow Steinmüller et al. (2019) and set $p = 0.2$ and $i = 0.05$ across all countries, industries, and years. Similar to Section 1.5 of the main text, we additionally calculate country-year-specific EATRs for the sake of comparison using the same symmetric financing and asset weights from Steinmüller et al. (2019).

In the following, we redo the graphical analysis from Section 1.5 of the main text using the EATR instead of the EMTR as tax measure of interest. The EATRs in Figure 1.9 exhibit a similar downward trend as the EMTRs in Figure 1.1. However, when comparing these two figures, two things stand out. First, the yearly means of the country-year-specific EATR (i.e., the black line) are around 8 percentage points higher each year. Second, the deviation of the yearly means of the country-industry-year-specific EATRs from the respective country-year-specific counterparts is substantially smaller compared to the EMTR figure. Both results can be explained with the fact that the NPV of depreciation allowances plays a relatively small role in determining the magnitude of EATRs compared to the statutory corporate tax rate. This also explains the strong centering of the distribution of $(EATR_{cit} - EATR_{ct})$ right around the zero mark in Figure 1.10. The fact that the country-industry-specific financing and asset structures play a small role is underlined by the high adjusted R^2 of 99.4% that a regression of $EATR_{cit}$ on country-year fixed effects yields.

Figure 1.9: ***DEVELOPMENT OF MEAN COUNTRY-YEAR AND COUNTRY-INDUSTRY-YEAR-SPECIFIC EATRs***

The figure depicts the development of the mean country-year and country-industry-year-specific EATRs calculated in Appendix 1.8.9. The grey dots represent the mean country-industry-year-specific EATRs across countries for each year. The black dots that are connected by black lines represent the mean country-year-specific EATRs across countries for each year. Calculations are based on a sample of 75,126 observations.

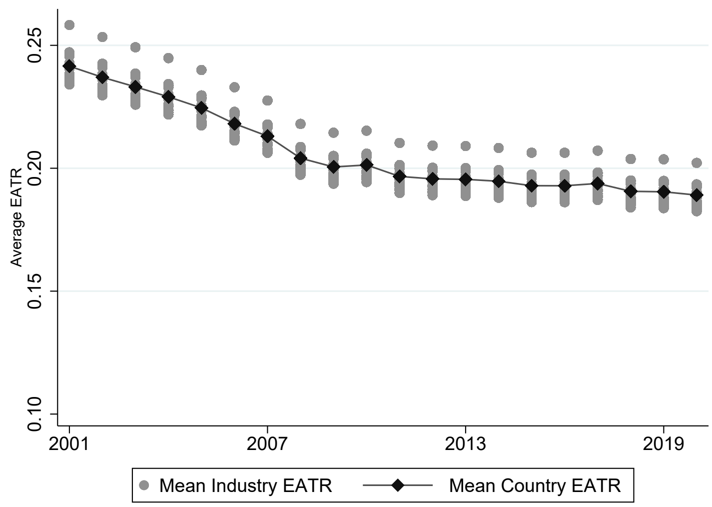
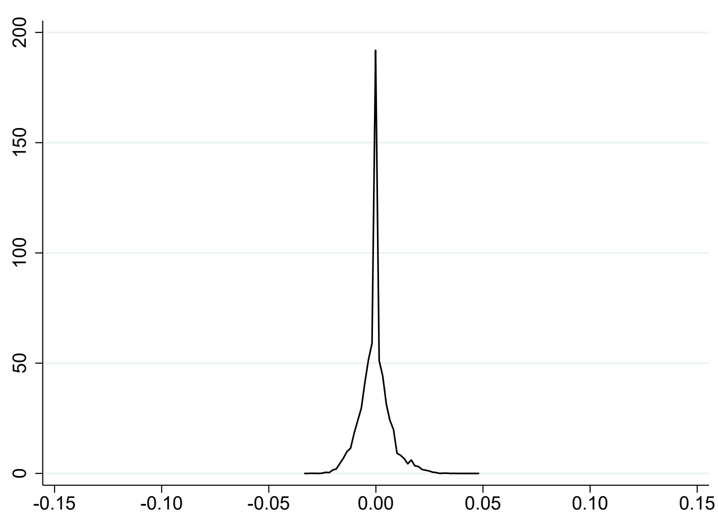


Figure 1.10: ***DISTRIBUTION OF DEVIATIONS FROM THE COUNTRY-YEAR EATRs***

The figure depicts the distribution of the differences between country-industry-year-specific and country-year-specific EATRs calculated in Appendix 1.8.9. The distribution is calculated based on 75,126 observations using a triangle kernel with a bandwidth of 0.0005.



1.8.10. Analysis of Investment Responses using Investment Rates

In the following, as robustness check, we analyze the sensitivity of firm entities' investment with respect to our country-industry-year-specific EMTRs using gross investment rates into fixed assets instead of the logarithm of the asset stock as dependent variable. The setup we use is derived from Liu (2020), who investigates the investment behavior of UK multinationals after the UK's switch from a worldwide to a territorial tax system in 2009. In detail, the gross investment rate into fixed assets ($Gross\ investment\ K_{jt}$) is obtained by adding year t's depreciation and amortization to the net change in the fixed asset stock from the previous to the current year. Then, this term is scaled by the previous year's fixed asset stock.[62] As control variables at the firm-entity level, Liu (2020) uses the one-period lag of the logarithm of sales ($log\ SALES_{jt-1}$) as well as the cash flow rate ($CF\ rate_{jt}$), the one period lag of the sales growth rate ($SALES\ growth_{jt-1}$), and the one period lag of the profit margin ($Profit\ margin_{jt-1}$).[63] To minimize the influence of outliers, following Liu (2020), we winsorize all ratios – including the investment rate – at the top and bottom 1 percentiles. At the country-level, we control for the GDP per capita growth rate, population size, unemployment rate, the *Rule of Law* indicator, as well as a financial institution stability indicator.[64] Note that, as above in the main body of the paper, we exclude certain industries and require a firm entity to appear in the sample at least two times. Following Liu (2020), we estimate a set of models using a variety of control variable and fixed effects combinations. The results of the analysis are depicted in Table 1.11. Note that the use of different variables compared to the analysis in the main text leads to a larger sample size of over 27 million observations. All models yield negative and statistically significant coefficients on $EMTR_{cit}$.

[62]Formally, we get $Net\ investment\ K_{jt} = (K_{jt} - K_{jt-1} + DEPR_{jt})/K_{jt-1}$, with $K_{jt} = TFAS_{jt} + IFAS_{jt}$, denoting the fixed asset stock of firm-entity j in year t, calculated as the sum of tangible and intangible fixed assets. $DEPR_{jt}$ denotes the depreciation and amortization of j's assets in year t. Note that since the *Orbis* database depicts depreciation and amortization jointly using a single variable, we are not able to sensibly compute gross investment rates for the tangible fixed asset stock.

[63]For the definition of the ratio variables, see Section 1.6.2.

[64]See Section 1.3.4 for a detailed description of the variables. The financial institution stability indicator is the *Bank Z-score* from the World Bank's *Global Financial Development* database and estimates the likelihood of country's banking system to default.

In terms of the magnitude of the investment response, we find that – with the exception of columns (1) and (6) – all models yield coefficients that are roughly comparable with the benchmark estimates depicted in Table 1.3.

Table 1.11: *TAX-RESPONSIVENESS USING GROSS INVESTMENT RATES*

The table presents OLS estimates. The dependent variable is the gross investment rate in fixed assets ($Gross\ investment\ K_{jt}$). Robust standard errors are reported in parentheses (clustered at the country-industry-year level). *** denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level. Details on the variables are provided in Appendix 1.8.10.

	(1)	(2)	(3)	(4)	(5)	(6)
$EMTR_{cit}$	-0.653***	-0.560***	-0.550***	-0.399**	-0.382**	-1.000*
	(0.240)	(0.191)	(0.181)	(0.179)	(0.169)	(0.543)
Firm-entity level controls	NO	YES	YES	YES	YES	YES
Country level controls	NO	NO	NO	YES	YES	YES
Entity fixed effects	YES	YES	YES	YES	YES	YES
Year fixed effects	YES	YES	NO	YES	NO	NO
Industry-year fixed effects	NO	NO	YES	NO	YES	NO
Country-year fixed effects	NO	NO	NO	NO	NO	YES
Adjusted R^2	0.068	0.135	0.135	0.136	0.136	0.136
Observations	27,228,920	27,228,920	27,228,920	27,228,920	27,228,920	27,228,788

1.9. Acknowledgements

I am indebted to Georg U. Thunecke and Georg Wamser who are co-authors of Chapter 1. All authors would like to thank Laura Arnemann, Johannes Becker, Giacomo Brusco, Ruby Doeleman, Peter Egger, Gerard Masllorens Fuentes, Dominika Langenmayr, Valeria Merlo, Jacquelyn Pless, Nadine Riedel, Martin Ruf, Frank Stähler, Johannes Voget, and participants at various RSIT workshops and conferences, the Virtual Doctoral Tax Seminar of the Catholic University of Eichstätt-Ingolstadt and University of Paderborn in the winter term 2022/23, the ZEW Public Finance Conference 2023, and the 79th Annual Congress of the International Institute of Public Finance in Utah for helpful comments and suggestions. Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – FOR2783 – is gratefully acknowledged.

1.10. References

Aggarwal, R., & Kyaw, N. A. (2008). Internal capital networks as a source of MNC competitive advantage: Evidence from foreign subsidiary capital structure decisions. *Research in International Business and Finance*, 22(3), 409-439.

Chaney, T., Sraer, D., & Thesmar D. (2012). The collateral channel: How real estate shocks affect corporate investment, *American Economic Review*, 102(6), 2381-2409.

Chirinko, R.S., Fazzari, S.M., & Meyer A. (1999). How responsive is business capital formation to its user cost? An exploration with micro data. *Journal of Public Economics*, 74(1), 53-80.

Cummins, J.G., Hassett, K.A., & Hubbard, R.G. (1994). A reconsideration of investment behavior using tax reforms as natural experiments. R.E. Hall and R.J. Caballero (Eds.), Brookings Papers on Economic Activity, 1994(2), 1-74.

Cummins, J.G., Hassett, K.A., & Hubbard, R.G. (1995). Have tax reforms affected investment? Tax Policy and the Economy 9, NBER, 131-149.

Cummins, J.G., Hassett, K.A., & Hubbard, R.G. (1996). Tax reforms and investment: A cross-country comparison. *Journal of Public Economics*, 62(1-2), 237-273.

Da Rin, M., Sembenelli, A., & Di Giacomo, M. (2010). Corporate taxation and the size of new firms: evidence from Europe. *Journal of the European Economic Association*, 8(2-3), 606-616.

De Mooij, R. A., & Ederveen, S. (2003). Taxation and foreign direct investment: A synthesis of empirical research. *International Tax and Public Finance*, 10(6), 673-693.

De Mooij, R. A., & Ederveen, S. (2008). Corporate tax elasticities: A reader's guide to empirical findings. *Oxford Review of Economic Policy*, 24(4), 680-697.

Devereux, M. P., & Griffith, R. (1998a). The taxation of discrete investment choices. IFS working papers No. W98/16.

Devereux, M. P., & Griffith, R. (1998b). Taxes and the location of production: Evidence from a panel of US multinationals. *Journal of Public Economics*, 68(3), 335-367.

Devereux, M. P., & Griffith, R. (2002). The impact of corporate taxation on the location of capital: A review. *Swedish Economic Policy Review*, 9(1), 79-106.

Devereux, M. P., & Griffith, R. (2003). Evaluating tax policy for location decisions. *International Tax and Public Finance*, 10(2), 107-126.

Edgerton J. (2010). Investment incentives and corporate tax asymmetries. *Journal of Public Economics*, 94(11-12), 936-952.

Egger, P., & Loretz, S. (2010). Homogeneous profit tax effects for heterogeneous firms? *The World Economy*, 33(8), 1023-1041.

Egger, P., Loretz, S., Pfaffermayr, M., & Winner, H. (2009). Firm-specific forward-looking effective tax rates. *International Tax and Public Finance*, 16(6), 850-870.

Egger, P., & Merlo, V. (2011). Statutory corporate tax rates and double-taxation treaties as determinants of multinational firm activity. *FinanzArchiv/Public Finance Analysis*, 67(2), 145-170.

Egger, P., Merlo, V., & Wamser, G. (2014). Unobserved tax avoidance and the tax elasticity of FDI. *Journal of Economic Behavior & Organization*, 108, 1-18.

Fazzari, S.M., Hubbard, R.G., & Petersen, B.C. (1987). Financing constraints and corporate investment. NBER Working Paper 2387.

Feld, L. P., & Heckemeyer, J. H. (2011). FDI and taxation: A meta-study. *Journal of Economic Surveys*, 25(2), 233-272.

Goldbach, S., Møen, J., Schindler, D., Schjelderup, G., & Wamser, G. (2021). The tax-efficient use of debt in multinational corporations. *Journal of Corporate Finance*, 71, 102119.

Goldbach, S., Nagengast, A. J., Steinmüller, E., & Wamser, G. (2019). The effect of investing abroad on investment at home: On the role of technology, tax savings, and internal capital markets. *Journal of International Economics*, 116, 58-73.

Hall, R. E., & Jorgenson, D. W. (1967). Tax policy and investment behavior. *American Economic Review*, 57(3), 391-414.

Hanappi, T. (2018). Corporate effective tax rates: Model description and results from 36 OECD and non-OECD countries. OECD Taxation Working Papers No. 38.

Harhoff D., & Ramb, F. (2001). "Investment and taxation in Germany – Evidence from firm level panel data". In Deutsche Bundesbank (Ed.), *Investing Today for the World of Tomorrow* (pp. 47-83). Berlin, Heidelberg, New York, Springer.

Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (2nd ed.). New York: Springer.

Huizinga, H., Laeven, L., & Nicodème, G. (2008). Capital structure and international debt shifting. *Journal of Financial Economics*, 88(1), 80-118.

Janský, P. (2023). Corporate effective tax rates for research and policy. *Public Finance Review*, 51(2), 171-205.

Kanemoto, K., Moran, D., & Hertwich, E. G. (2016). Mapping the carbon footprint of nations. *Environmental Science & Technology*, 50(19), 10512-10517.

Kanemoto, K., Moran, D., Lenzen, M., & Geschke, A. (2014). International trade undermines national emission reduction targets: New evidence from air pollution. *Global Environmental Change*, 24, 52-59.

Kaufmann, D., Kraay, A., & Mastruzzi, M. (2011). The worldwide governance indicators: Methodology and analytical issues. *Hague Journal on the Rule of Law*, 3(2), 220-246.

Keuschnigg, C., & Ribi, E. (2013). Profit taxes and financing constraints. *International Tax and Public Finance*, 20(5), 808-826.

King, M. A. (1974). Taxation and the cost of capital. *Review of Economic Studies*, 41(1), 21-35.

King, M. A., & Fullerton, D. (1984). Introduction to "The taxation of income from capital: A comparative study of the United States, the United Kingdom, Sweden, and Germany". In M.A. King, D. Fullerton (Eds.), *The taxation of income from capital: A comparative study of the United States, the United Kingdom, Sweden, and Germany* (pp. 1-6). University of Chicago Press.

Lenzen, M., Kanemoto, K., Moran, D., & Geschke, A. (2012). Mapping the structure of the world economy. *Environmental Science & Technology*, 46(15), 8374-8381.

Lenzen, M., Moran, D., Kanemoto, K., & Geschke, A. (2013). Building Eora: A global multi-region input-output database at high country and sector resolution. *Economic Systems Research*, 25(1), 20-49.

Little, R. J. (1988). Missing-data adjustments in large surveys. *Journal of Business & Economic Statistics*, 6(3), 287-296.

Little, R.J., & Rubin, D.B. (2019). *Statistical analysis with missing data* (3rd ed). John Wiley & Sons.

Liu, L. (2020). Where does multinational investment go with territorial taxation? Evidence from the United Kingdom. *American Economic Journal: Economic Policy*, 12(1), 325-358.

MacKie-Mason, J. K. (1990). Do taxes affect corporate financing decisions? *Journal of Finance*, 45(5), 1471-1493.

Moulton, B. R. (1990). An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *Review of Economics and Statistics*, 72(2), 334-338.

OECD (1991). *Taxing profits in a global economy: domestic and international issues.* OECD Publication and Information Center.

Ohrn, E. (2018). The effect of corporate taxation on investment and financial policy: Evidence from the DPAD. *American Economic Journal: Economic Policy*, 10(2), 272-301.

Overesch, M. (2009). The effects of multinationals' profit shifting activities on real investments. *National Tax Journal*, 62(1), 5-23.

Overesch, M., & Rincke, J. (2009). Competition from Low-wage Countries and the Decline of Corporate Tax Rates: Evidence from European Integration. *World Economy*, 32(9), 1348-1364.

Overesch, M., & Wamser, G. (2009). Who cares about corporate taxation? Asymmetric tax effects on outbound FDI. *World Economy*, 32(12), 1657-1684.

Rubin, D. B. (1986). Statistical matching using file concatenation with adjusted weights and multiple imputations. *Journal of Business & Economic Statistics*, 4(1), 87-94.

Sørensen, P. B. (2004). Measuring taxes on capital and labor: An overview of methods and issues. In PB Sørensen (Ed.), *Measuring the tax Burden on capital and labor* (pp. 1-33). Cambridge, MA: MIT Press.

Spengel, C., Heckemeyer, J. H., Bräutigam, R., Nicolay, K., Klar, O., & Stutzenberger, K. (2016a). The effects of tax reforms to address the debt-equity bias on the cost of capital and on effective tax rates. European Commission Taxation Papers No. 65, Brüssel.

Spengel, C., Heckemeyer, J. H., Nusser, H., Klar, O., & Streif, F. (2016b). The impact of tax planning on forward-looking effective tax rates. European Commission Taxation Papers No. 64, Brüssel.

Spengel, C., Heckemeyer, J. H., & Streif, F. (2016c). The effect of inflation and interest rates on forward-looking effective tax rates. European Commission Taxation Papers No. 63, Brüssel.

Spengel, C., Schmidt, F., Heckemeyer, J., Nicolay, K., Bartholmeß, A., Ludwig, C., Steinbrenner, D., Buchmann, P., Bührle, T., Dutt, V., Fischer, L., Spix, J., Stage, B., Weck, S., Wickel, S. (2021). Effective tax levels using the Devereux/Griffith methodology. Final Report 2021. Project for the EU Commission TAXUD/2020/DE/308.

Steinmüller, E., Thunecke, G. U., & Wamser, G. (2019). Corporate income taxes around the world: A survey on forward-looking tax measures and two applications. *International Tax and Public Finance*, 26(2), 418-456.

Stöwhase, S. (2005). Tax-rate differentials and sector-specific foreign direct investment: Empirical evidence from the EU. *FinanzArchiv/Public Finance Analysis*, 61(4), 535-558.

Suárez Serrato, J.C. (2019). Unintended consequences of eliminating tax havens. Unpublished manuscript.

Tørsløv, T., Wier, L., & Zucman, G. (2023). The missing profits of nations. *Review of Economic Studies*, 90(3), 1499-1534.

Van Buuren, S. (2018). *Flexible imputation of missing data* (2nd ed.). CRC Press.

Wamser, G., Merlo, V., Ruf, M., Stähler, F., Strohmaier, K., Eklund, J., Hansen, J., Laudage, S., Mc Auliffe, S., Paraknewitz, J., & Thunecke, G. U. (2023). A new database on international tax institutions. Forthcoming in RSIT Working Papers.

Zwick, E., & Mahon, J. (2017). Tax policy and heterogeneous investment behavior. *American Economic Review*, 107(1), 217-248.

2. Effective Corporate Income Taxation and Corruption

Abstract[§] –We show that effective corporate income taxes are lower in EU NUTS 2 regions where citizens perceive corruption to be comparatively more prevalent. We develop a new approach for calculating region-industry-year-specific empirical effective income tax rates (EEITRs) using firm-entity-level income statement data. Controlling for proxies for deductions that could legally be claimed (e.g., depreciation allowances, deduction of interest payments, potential for loss carryforwards, preferential treatment of patent revenues) and additional controls (e.g., regional GDP), as well as country-industry-year fixed effects, our benchmark model suggests that a one standard deviation increase in corruption leads to a statistically significant decrease in EEITRs of approximately 0.4 percentage points. From an economic point of view, this effect is sizeable given that several countries in our sample exhibit between regions differences in corruption of more than one standard deviation. Our findings suggest that high corruption regions exhibit higher levels of tax evasion via overstated deductions.

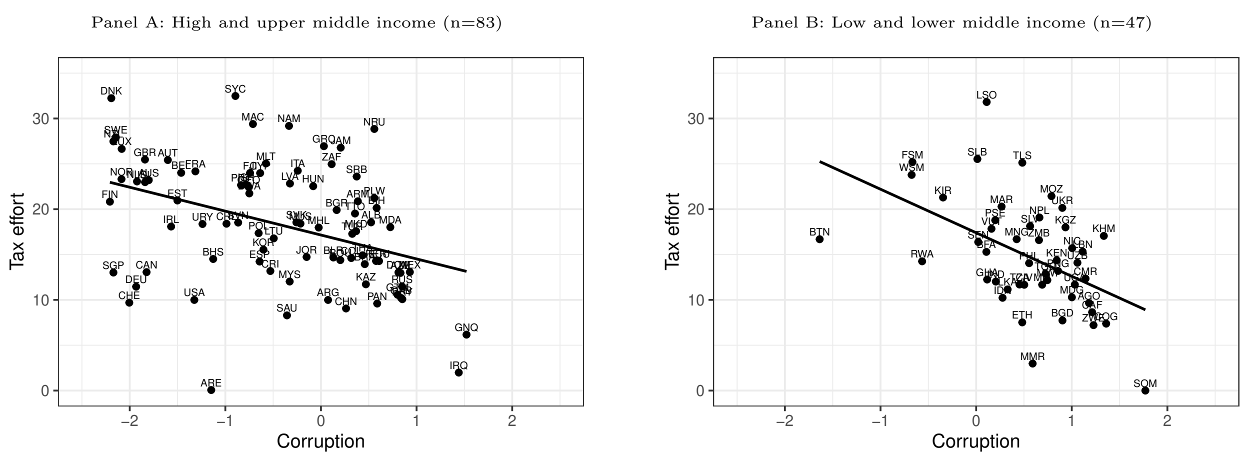
[§]This paper is based on joint work with Peter Egger, Valeria Merlo, and Georg Wamser.

2.1. Introduction

The adverse effects of corruption on countries' ability to raise domestic revenue, the so-called fiscal capacity (Kaldor, 1963), have been extensively studied and documented in the context of the developing world.[1] Interestingly, however, a negative correlation between tax effort, i.e., tax revenue as percentage of GDP, and corruption can also be found for developed countries, as can be seen in Figure 2.1. An important aspect to keep in mind with such correlations is that the interpretation of tax effort is generally ambiguous, as a comparatively low tax effort may be the result of either a generous tax code or tax evasion, or a combination of both. Therefore, a meaningful analysis of the relationship between taxation and corruption requires controlling for all relevant tax regulations.

Figure 2.1: *TAX EFFORT AND CORRUPTION BY COUNTRY FOR DIFFERENT INCOME LEVELS IN 2018*

The figure depicts scatter plots of the country level variables tax effort (tax revenue as % of GDP) and a corruption measure for different income levels of the World Bank's classification system. The labels above to the dots depict the ISO 3 codes of the respective countries. The data refers to the year 2018, which is the last year for which the corruption measure is available. The line is fit by OLS (models include a constant). The tax revenue (% of GDP) variable is taken from the World Bank's *World Development Indicator Database*. Tax revenue refers to compulsory transfers to the central government for public purposes, excluding penalties, fines, and most social security contributions. The corruption measure is the so-called *Control of Corruption* measure from the World Bank's *Worldwide Governance Indicators* database and measures "the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as 'capture' of the state by elites and private interests" (Kaufmann et al., 2011, p. 223). The corruption variable varies on an interval -2.5 to 2.5. Originally, it is constructed such that a higher value corresponds to less corruption (Kaufmann et al., 2011). The graphs depict the original measure multiplied by -1, such that a higher value indicates more corruption. Note that we exclude Croatia as it exhibits a tax effort of over 150%.



In this paper, motivated by the findings in Figure 2.1, we analyze this relationship in the

[1] See, e.g., Besley and Persson (2013, 2014), Bird et al. (2008), Ghura (1998).

context of developed countries. More precisely, our analysis focuses on member countries of the EU. To manage complexity, we focus solely on corporate income taxation. Our research design exploits the fact that the tax code applies to all firms located in a given country equally, which allows us to base our identification on the variation between different regions of the same country. As a simple aggregation of income statement data shows, the regional variation in firm entities' total tax to profit ratios is strong, even when comparing aggregates corresponding to the same industry and year, see Figure 2.2.[2] Noteworthy regional patterns in the figure that hold for all depicted industries are, e.g., that the total tax to profits ratio is substantially higher in the northern part of Italy compared to its south or that the ratio is higher in the Madrid region compared to Madrid's surrounding regions.[3]

One particular empirical challenge that we face lies in the fact that large-scale databases that provide financial accounting information, such as Bureau van Dijk's *Orbis*, only provide a single aggregated tax liability variable that includes not only the corporate income tax, but also other taxes, e.g., the carbon taxes or the property taxes. Therefore, simply analyzing the total tax over profits ratio, which is depicted in Figure 2.2, would lead to measurement error. To solve this issue, we propose a novel approach for calculating region-industry-year-specific empirical effective income tax rates (EEITRs) using *Orbis* that exploits the fact that the unobserved income tax payments contained in the total tax liability variable are the only taxes that are directly derived from the respective firm entities' profits.

For our analysis of the EEITRs, we proxy the income tax deductions that firm entities may claim by combining information on firm entities from income statements, balance sheets, and patent records with information from countries' tax codes. Note that this is crucial even when comparing EEITRs within the same jurisdiction, as deductions – despite being calculated on the same legal basis – can vary in magnitude depending on individual firm entities' characteristics such as asset structures, financing compositions, or R&D activities.

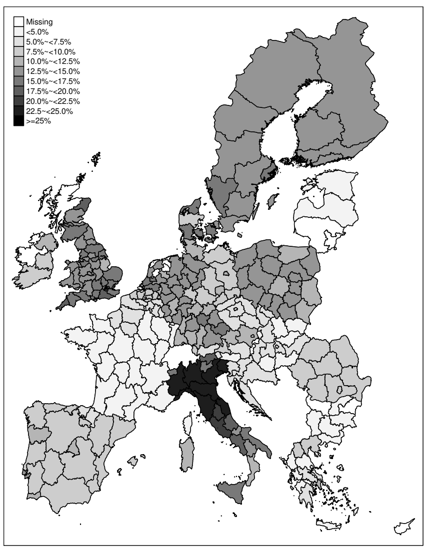
[2]More detailed descriptive statistics on the range of the regional aggregated tax to profit ratio are provided in Appendix 2.6.1.

[3]Note that the strong regional variation in Germany is mainly due to regional trade taxes that are added to the country-wide statutory tax rate. More detail on this is provided below.

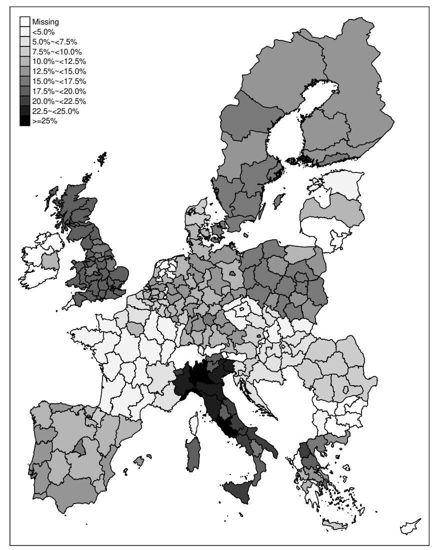
Figure 2.2: ***MEDIAN OF THE RATIO TAX LIABILITY OVER EARNINGS BEFORE INTEREST, TAXES, DEPRECIATION, AND AMORTIZATION IN DIFFERENT NUTS 2 REGIONS IN 2013***

The figure depicts the median of the firm-entity-specific ratio tax liability over Earnings Before Interest, Taxes, Depreciation, and Amortization (EBITDA) for different NUTS 2 regions (version 2016) of the EU 28. The different panels depict the ratios corresponding to firm entities operating in different industries (NACE Rev. 2 sections). All plots correspond to data for the year 2013. Firm entities belonging to MNEs are excluded. Only observations with strictly positive EBITDA are used for the calculation. Observations of the depicted ratio in the top and bottom one percentile were excluded from the sample. A minimum of 25 firm entity observations per region and industry combination was required. The source of the data is *Orbis*. Maps plotted with the *tmap* package for *R* (Tennekes, 2018).

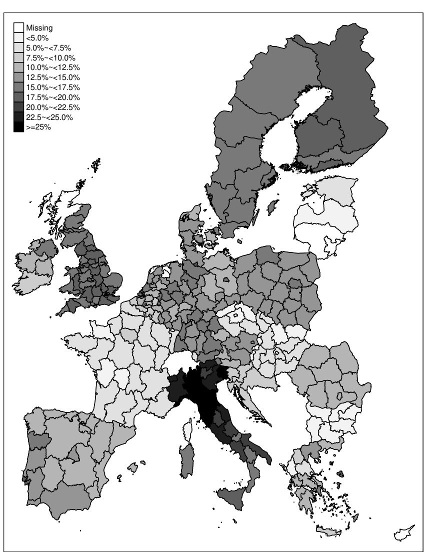
Panel A: Manufacturing



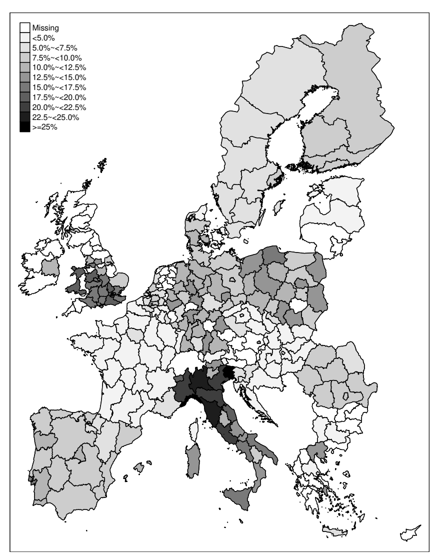
Panel B: Construction



Panel C: Wholesale and retail trade



Panel D: Transportation and storage



Furthermore, loss carryforwards and loss carrybacks in combination with regional economic shocks may play a role in explaining regional differences in effective income taxation. Controlling for these deduction proxies as well as additional controls (e.g., regional GDP) and country-industry-year fixed effects, we analyse if variation in a regional corruption measure contributes to explaining the variation in the EEITRs. Our results show that EEITRs are lower in NUTS 2 regions where citizens perceive corruption to be more prevalent. More precisely, we find that a one standard deviation increase in corruption is associated with a statistically significant decrease in the EEITR of approximately 0.4 percentage points. This effect is economically substantial, given that several countries in our sample exhibit between regions differences in corruption of more than one standard deviation. Additional results show that EEITRs are slightly higher in regions where survey results suggest that the tax morale is higher.[4]

Since we control for the legal ways to decrease the EEITR, our results imply that the lower EEITRs in high corruption regions are likely the result of tax evasion, i.e., the illegal and intentional actions taken by firms to reduce their legally due tax obligations. More precisely, our approach suggests the tax evasion is carried out via overstated deductions. The reason for this supposition is that we base our EEITRs on profits obtained from the income statements of firm entities. Since this information is publicly available, we do not expect it to deviate from the figures reported to the tax authority. In this regard, our paper differs from the existing body of empirical literature on corporate tax evasion, which focuses on the underreporting of profit or sales figures to tax authorities, but not on how the reported figures are transformed into the final tax base via deductions. For instance, the study related closest to ours by Alm et al. (2016) uses firm level survey data from the World Bank that measures the degree of tax evasion as the share of sales reported for tax purposes.[5] Using

[4]The latter finding is in line with the previous literature that shows that tax morale is a determinant for tax evasion, see, e.g., Richardson (2006), Torgler (2007), or Torgler et al. (2008).

[5]Note that this survey question was not asking the surveyed firm directly about its own sales ratio reported for taxes, but instead asked for an estimate of this ratio for a“typical” firm in the same area of business. See Alm et al. (2019) for a brief discussion of resulting potential issues regarding this surrogate.

instrumental variable methods, they find that higher corruption, measured by tax related bribes that were paid, leads to lower reported sales ratios. Using similar data but focusing on a limited number of transition countries, Uslaner (2010) also finds that the decision to pay taxes is negatively affected by corruption.[6] Other studies that use the same survey data from the World Bank are Alm et al. (2019), who show that financial constraints are a determinant of tax evasion, and Beck et al. (2014), who find that there is less tax evasion in countries with better credit information–sharing systems and higher branch penetration. Best et al. (2015) evaluate tax evasion behavior by firms in the context of the Pakistani tax system that – depending on the expected tax liability – taxes either profits or turnover. DeBacker et al. (2015) analyse confidential audit data from the Internal Revenue Service and find that owners from countries with more pronounced corruption norms tend to evade more taxes in the United States. Carrillo et al. (2017) evaluate the effectiveness of combating tax evasion using third-party information to verify tax reports. They evaluate an Ecuadorian policy intervention in which firms were notified about revenue discrepancies and find that most firms did not react and some adjusted their reporting to match the discrepancy amount. Doerr and Necker (2021) conduct a field experiment in which they compare offers for home improvement services on online markets for the cases where an invoice is requested or not. They find that in particular in markets that allow to sell anonymously, the willingness to evade taxes is given. In the context of corporate taxation, many studies evaluate the effectiveness of taxpayer audits to combat tax evasion. A recent example is Bergolo et al. (2023), who find that letters sent out by the tax authority in Uruguay that announce audits significantly affected tax compliance by firms regarding the value-added tax. Concerning the value-added tax, another study worth noting is Pomeranz (2015), who use randomized experiments conducted in Chile to analyse the role that third-party information plays for enforcement. Another paper that evaluates the effects of audits is Lediga et al. (2020) who focus on spillover effects from tax audits in South Africa and find that the tax liability of

[6]Theoretical contributions analysing the relationship between corruption and taxation include, e.g., Brueckner (2000), Flatters and MacLeod (1995), and Litina and Palivos (2016).

unaudited firms in the same local network as audited firms increases.[7]

The remainder of this paper is structured as follows. Section 2.2 presents the institutional setting. Section 2.3 describes the estimation strategy as well as the data used to carry out the analysis. The results are discussed in Section 2.4. Finally, Section 2.5 concludes.

2.2. Institutional Setting

Suppose firm entity j's Earnings Before Interest, Taxes, Depreciation, and Amortization (EBITDA) in period t amount to π_{jt}. Filing its tax return, j reduces π_{jt} to the final income tax base π_{jt}^{tax}, which is then taxed at the respective country c's statutory income tax rate τ_{ct} in case $\pi_{jt}^{tax} > 0$ and not at all in case $\pi_{jt}^{tax} \leq 0$. The final income tax liability, $ITAX_{jt}$, is hence given by

$$ITAX_{jt} = \begin{cases} \tau_{ct}\pi_{jt}^{tax} & \text{if } \pi_{jt}^{tax} > 0, \\ 0 & \text{if } \pi_{jt}^{tax} \leq 0. \end{cases} \tag{2.1}$$

The instruments that a firm entity may use to adjust π_{jt} to the final tax base π_{jt}^{tax} are defined in the respective country's tax code. Generally, countries grant depreciation allowances for certain assets and allow for the deduction of interest payments on debt. Furthermore, some countries grant preferential tax treatment for firm activities associated with R&D or innovation (so-called "patent boxes"). Other tools to adjust the current year's tax liabilities are loss carryforwards, i.e., applying losses from previous periods to the current period's income, and loss carrybacks, i.e., applying current losses to a previous year's tax return for an immediate refund of previously paid taxes.[8] In the context of MNEs, there may be additional charges when income is repatriated to foreign parent companies and when

[7] Other papers investigating the effect of tax audits include, e.g., Advani et al. (2021), DeBacker et al. (2018), and Kleven et al. (2011). Furthermore, Xu et al. (2019) investigate how regional political corruption levels impact auditor behavior in the United States.

[8] Note that in the context of loss carrybacks, the tax liability TAX_{jt} may be negative in case $\pi_{jt}^{tax} \leq 0$ – a special case that is not depicted in (2.1).

the country where the parent company is located seeks to tax worldwide income.

It is important to note that due to confidentiality, it is generally not possible to observe a firm entity's income tax return and the composition of the deductions that are claimed. However, it is possible to combine observable firm entity information with tax code regulations to proxy the deductions that could *legally* be claimed. In this paper, we control for such proxies when analysing regional empirical effective income tax rates (EEITRs), i.e., empirical measures that state the aggregated relationship between the income tax liability and the EBITDA of firm entities located in a given region. In particular, we are interested in whether regional corruption can contribute to explaining variation in EEITRs after controlling for the proxies for legal deductions. Based on results from the previous literature (Alm et al., 2016; Uslaner, 2010), we expect EEITRs to be lower in regions where corruption is more prevalent, as such environments have the potential to facilitate tax evasion. For instance, entities located in such high corruption regions may be more likely to successfully collude with or bribe officials. In our framework, we may think of tax evasion as illegally overstating deductions, which decreases the tax base and therefore the EEITR. It is important to note that we are not able to evaluate if the EBITDA that we use to calculate the EEITRs is itself correctly reported. In fact, it is a well-documented tax evasion strategy to underreport sales or profits to the tax authorities (see, e.g., Alm et al., 2016, 2019; Beck et al., 2014; Doerr and Necker, 2021; Uslaner, 2010). The results by Alm et al. (2016) and Uslaner (2010) suggest that tax evasion via underreporting sales figures is more pronounced in high corruption locations, which suggests that our results should be interpreted as lower bounds.

2.3. Empirical Approach

2.3.1. Estimation Strategy

For the analysis of the relationship between EEITRs and corruption we run OLS regressions at the NUTS 2 region, NACE Rev. 2 section,[9] and year level. The distinction between industries is crucial, as different industries have been shown to use fundamentally different typical financing and asset structures. This is tax-relevant since interest payments on debt are tax-deductible and depreciation allowances differ between asset types (see, e.g., Fabling et al., 2014; Mc Auliffe et al., 2022; Steinmüller et al., 2019). Furthermore, different industries in the same country may be exposed to heterogeneous shocks in a given year. This matters for the magnitude of EEITRs, as consequently the subsequent potential for adjustments of the tax base using loss carryforwards and loss carrybacks differs. Using industry-specific EEITRs as well as tax base determinants that account for industry heterogeneity ensures that the results are not contaminated by differences in industry structures between different regions. Formally, the equation that we estimate states as follows:

$$EEITR_{rit} = \beta Corruption_{rt} + \boldsymbol{\psi X_{rit}} + \boldsymbol{\zeta X_{rt}} + c_{cit} + \varepsilon_{rit}. \tag{2.2}$$

The indices c, r, i, and t denote country, region, industry, and year, respectively. $EEITR_{rit}$ is the region-industry-year-specific EEITR. β is the coefficient on our region-year-specific corruption measure, $Corruption_{rt}$. We further control for a set of region-industry-year-specific variables, contained in $\boldsymbol{X_{rit}}$, as well as for region-year-specific variables, contained in $\boldsymbol{X_{rt}}$. The corresponding parameter estimates are collected in the vectors $\boldsymbol{\psi}$ and $\boldsymbol{\zeta}$, respectively. These sets of controls include different determinants of the tax base and general economic measures. We further include country-industry-year fixed effects, denoted by c_{cit}, to control for level differences that are due to factors that equally impact all regional EEITRs corre-

[9]Note that we shall henceforth use the term "industry" for the sake of simplicity to denote NACE Rev. 2 sections.

sponding to the same country, industry, and year.[10] Finally, the error component is denoted by ε_{rit}.

2.3.2. Data and Sample

To carry out the analysis we use data from a number of different sources. In the following, the data preparation, variable construction, as well as the resulting sample are described. Note that we generally exclude *Orbis* observations corresponding to firm entities that are part of an MNE, as we are not able to accurately observe tax-relevant profit-shifting in the data (see discussion above). That is, all results presented are based on stand-alone firms and national groups.[11]

Empirical Effective Income Tax Rate ($EEITR_{rit}$): In the literature, empirical tax rates are often calculated from income statement data as total taxes paid relative to a pre-tax profit measure, as also done in the introduction of this paper.[12] While this ratio holds valuable information, it has to be noted that the total tax liability item from income statements that is commonly used (and is often the only measure available, e.g., when using *Orbis*) may also contain taxes other than the corporate income tax, e.g., carbon taxes or property taxes. These "non-income taxes" are, however, not of interest for our analysis, as we do not observe their tax bases (e.g., carbon emissions or property values) and can therefore make no statement about whether the correct amount was paid or not. Instead, we are interested in constructing an EEITR that depicts the relationship between income tax payments and profits only. Using *Orbis*, we propose a new approach for obtaining EEITRs that exploits the

[10] Note that several years of various key variables used in the analysis, including our corruption measure, are imputed using information from observed years (see Section 2.3.2). Furthermore, the *Orbis* that we use to calculate various variables is highly unbalanced, with the general tendency that more firms are included each year. Therefore, a panel analysis with, e.g., region or region-industry fixed effects that exploits variation across time is not feasible in a sensible way.

[11] Note that we identify MNEs in *Orbis* using the information on the Global Ultimate Owner (GUO). We define an MNE as a corporate group with at least two firm entities that have the same GUO and are located in different countries. In Appendix 2.6.3, we provide our results table using data that also includes MNEs. The results including the MNEs are highly similar to the ones where they are excluded. This may be due to the fact that the number of MNEs in our sample is small.

[12] A recent and extensive overview of such measures is provided in Janský (2023).

fact that the unobserved income tax payments contained in the total tax liability variable are the only tax payments that directly vary with and depend on profits. In detail, our approach defines the EEITR as the marginal effect of a one unit increase in the EBITDA on the tax liability for firm entities that report a strictly positive EBITDA in the given year.[13] The estimation of these marginal effects is carried out using the following regression, which is separately run for every region, industry, and year combination:

$$\frac{TAX_{jt}}{EMPL_{jt}} = \beta_1 \mathbb{1}(EBITDA_{jt} > 0) \cdot \frac{EBITDA_{jt}}{EMPL_{jt}} + \beta_2 \mathbb{1}(EBITDA_{jt} \leq 0) + \varepsilon_{jt}. \qquad (2.3)$$

The indices j and t denote firm entity and year, respectively. TAX_{jt}, $EMPL_{jt}$, and $EBITDA_{jt}$ denote the *Orbis* variables tax liablity, number of employees, and EBITDA, respectively. $\mathbb{1}(EBITDA_{jt} > 0)$ is an indicator function equal to one if the EBITDA of j in year t is strictly positive and zero if not. $\mathbb{1}(EBITDA_{jt} \leq 0)$ is equal to one if the EBITDA of j in t is non-positive and zero if not. β_1 and β_2 are the coefficients we are looking to estimate, with β_1 being the EEITR corresponding to the region, industry, and year of the respective estimation sample. Finally, ε_{jt} denotes the error component.

The interaction term $\mathbb{1}(EBITDA_{jt} > 0) \cdot (EBITDA_{jt}/EMPL_{jt})$ in (2.3) is endogenous due to both simultaneity and correlation with omitted variables. The simultaneity issue arises due to the fact that the precise magnitude of the EBITDA is jointly determined with the resulting (expected) tax burden by firm entities' accounting divisions. Potential variables that are omitted that may be correlated with $\mathbb{1}(EBITDA_{jt} > 0) \cdot (EBITDA_{jt}/EMPL_{jt})$ include taxes other than the income tax, such as carbon or property taxes, as well as any tax-relevant

[13] Note that both the EBITDA as well as the tax liability figures from *Orbis* stem from the entities' income statements and must not coincide with the profit reported to the tax authority and the true tax liability that they are obliged to pay, respectively. However, since the EBITDA is based on (billed) real economic transactions, we do not expect a systematic bias. This argument could not be made for, e.g., the EBIT or other profit measures from the income statement that already account for depreciation, since the amounts of depreciation on the income statement and on the tax return may differ due to differences in tax and financial accounting (see, e.g., Graham et al., 2012). Concerning the total tax liability variable form *Orbis*, Arulampalam et al. (2012) argue that it is a good approximation for the true tax obligation, especially when only focusing on national firms, as we do in this paper.

deductions, including interest payments, depreciation, or losses from previous periods.[14] To obtain unbiased estimates of β_1, we construct an instrument for every firm entity observation j in the given year t, which we define as the mean of the ratio EBITDA/EMPL over all firm entities, excluding j itself, that operate in the same 3-digit industry and the same country as j in year t and report a strictly positive EBITDA. We argue that the exclusion restriction, which requires there to be no direct impact of the instrument on $TAX_{jt}/EMPL_{jt}$, is satisfied, as for the determination of j's tax liability only j's own tax base is of relevance.

To mitigate the influence of outliers we drop the top and bottom one percentile of the firm-entity-specific ratios TAX/EMPL, EBITDA/EMPL, as well as TAX/EBITDA. Furthermore, to ensure meaningful estimates are obtained, we require a minimum number of 50 firm entity observations per region, industry, and year combination. The time span we use for the estimation is 2009 to 2018, since these are the years of our *Orbis* sample with the most complete coverage. Using 2SLS estimation, we obtain 22,389 EEITRs for 18 industries and 261 NUTS 2 regions that span across 25 EU countries.[15] For 99.97% of the regressions, we find a strong first stage[16] and 99.12% of the EEITRs are in the "plausible range" $(0; \tau_{ct}]$, with τ_{ct} denoting the statutory income tax rate in country c in year t. In Figure 2.3, the EEITR is depicted for the same industries and for the same year for which Figure 2.2 depicted the median of the ratio total tax liability to EBITDA. It can be seen that the regional patterns

[14] Note that we treat the indicator $\mathbb{1}(EBITDA_{jt} \leq 0)$ as exogenous, as it is not possible to construct a strong instrument for this variable. One reason for this lies in the fact that the total number of observations per region, industry, and year bin for which the EBITDA is not positive is often very small. We are aware that exogeneity is a strong assumption which may potentially introduce a bias to the estimates. However, arguments can be found for why this issue may not be as severe. For instance, regarding simultaneity, one may argue that – unlike the exact amount of the EBITDA – the fact whether or not a firm entity is loss making is not always at the discretion of the accounting unit.

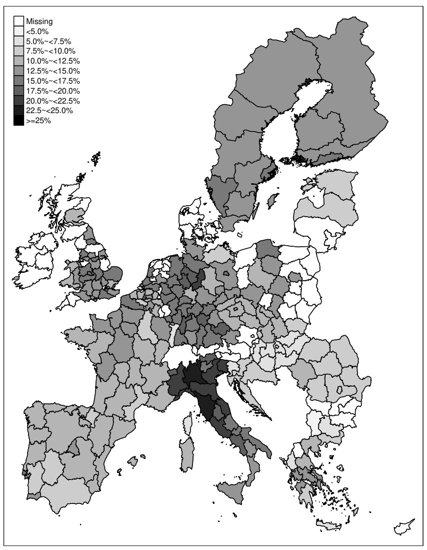
[15] The three NACE Rev. 2 sections for which we do not obtain any EEITRs due to data coverage reasons are *O Public administration and defence; compulsory social security*; *T Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use*; as well as *U Activities of extraterritorial organizations and bodies*. The EU countries without any coverage are Cyprus, Lithuania, and Malta.

[16] We consider the first stage of a regression strong, if the coefficient estimate on the instrument is different from zero at a significance level of 10 percent. The tests are based on robust standard errors.

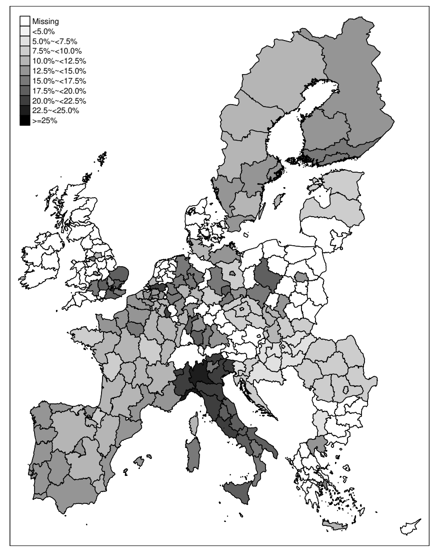
Figure 2.3: *EMPIRICAL EFFECTIVE INCOME TAX RATES IN DIFFERENT NUTS 2 REGIONS IN 2013*

The figure depicts EEITRs for different NUTS 2 regions (version 2016) of the EU 28. The different panels depict the EEITRs corresponding to firm entities operating in different industries (NACE Rev. 2 sections). All plots correspond to data for the year 2013. Firm entities belonging to MNEs are excluded. The calculation of the EEITRs is detailed in Section 2.3.2. Maps plotted with the *tmap* package for *R* (Tennekes, 2018).

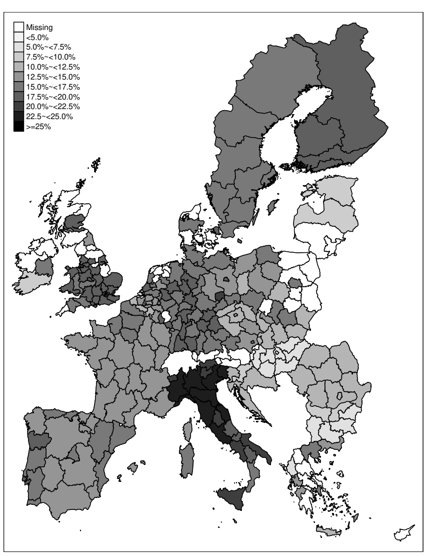
Panel A: Manufacturing



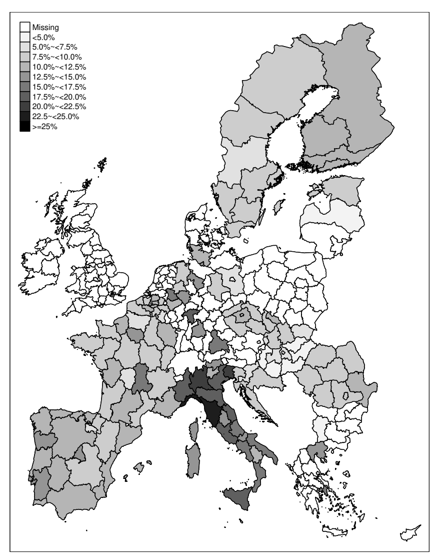
Panel B: Construction



Panel C: Wholesale and retail trade



Panel D: Transportation and storage



are very similar in the two figures.[17] This observation is supported by the high unconditional correlation of the two measures which amounts to 0.84. This suggests that the income tax payments are the main driver of the variation in firm entities' total tax burden.

In a last step, we analyse the region-industry-year-specific EEITRs using the analysis of variance (ANOVA) approach. The ANOVA setup allows us to quantify how much of the variance in the EEITRs is attributable to the country level (i.e., the national tax codes) and the industry level.[18] The simple ANOVA model that we use has the form

$$EEITR_{rit} = \alpha + \mu_c + \lambda_i + \theta_t + \eta_{rit}. \tag{2.4}$$

$EEITR_{rit}$ is the region-industry-year-specific EEITR. α denotes the constant. Country-, industry-, and year-specific sets of dummy variables are contained in μ_c, λ_i, and θ_t, respectively. η_{rit} is the remainder component which is not attributable to either countries, industries, or years. We denote the total variance in $EEITR_{rit}$ as SS_{EEITR} and the partial sums of squares of the country, industry, and year effects as SS_μ, SS_λ, and SS_θ, respectively. The model's residual sum of squares is SS_η. Note that it holds that $SS_{EEITR} = SS_\mu + SS_\lambda + SS_\theta + SS_\eta$. The results of the ANOVA in Table 2.1 suggest that the sums of squares of the country effects, SS_μ, contribute to SS_{EEITR} in a major way ($SS_\mu/SS_{EEITR} = 42.57\%$). These findings suggest that the national tax codes are the main contributors to the variation in the EEITR. The industry effects play a smaller – nonetheless sizeable – role with $SS_\lambda/SS_{EEITR} = 21.43\%$. It is important to note that the R^2 of the model, which is given by $(SS_\mu + SS_\lambda + SS_\theta)/SS_{EEITR}$, amounts to 64.8% only. This suggests that idiosyncratic effects at the regional level also play an important role for explaining the variance in our EEITRs, which corroborates our approach of identifying the effect of corruption on EEITRs via within country, industry, and year variation.

[17] Note that the EEITR exhibits more regions with missing values since the minimum number of observations is set higher for the EEITR calculation compared to the calculation of the median of the ratio total tax liability to EBITDA.

[18] Note that the setup of the ANOVA broadly follows Egger et al. (2009).

Table 2.1: *ANOVA OF REGION-INDUSTRY-YEAR-SPECIFIC EEITRs*

The table depicts analysis of variance (ANOVA) results of the region-industry-year-specific Empirical Effectictive Income Tax Rates ($EEITR_{rit}$) that are calculated in Section 2.3.2. The ANOVA is based on 22,389 observations.

	Partial sum of squares	Degrees of freedom	F-statistic	p-value
Country effects	23.039	24	1103.79	0.000
NACE Rev. 2 section effects	11.596	17	784.35	0.000
Year effects	0.059	9	7.57	0.000
Model	35.700	50	820.99	0.000
Residual	19.427	22,338		
Total	55.127	22,388		
R^2	0.648			

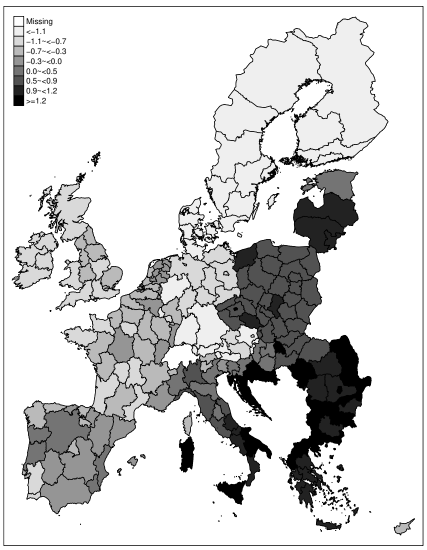
Corruption ($Corruption_{rt}$): The regional corruption measure that we use is taken from the *European Quality of Governance Index* (EQI). The corruption measure of the EQI aims to capture citizens' perceptions and experiences with corruption and is based on a set of survey questions that could be answered using a numeric scale (for details, see Charron et al., 2022). It is important to note that the survey questions on which the corruption measure is based do not specifically ask about tax evasion behavior of firms located in the given region but instead focus on corruption in the context of the local public school system, the public health care system, and the police force (Charron et al., 2022). Therefore, and also since the tax payments of local firms are generally not public knowledge, we argue that the corruption measure is exogenous in our regression setup. For our purpose, we use all four previous waves of the EQI for the years 2010 (Charron et al., 2014), 2013 (Charron et al., 2015), 2017 (Charron et al., 2019), and 2021 (Charron et al., 2022). The data is provided in a balanced panel spanning across 220 regions of all EU 28 countries.[19] For Belgium, Germany, and the UK the corruption measure is provided at the NUTS 1 level. For all other countries the measure is provided at the NUTS 2 level. A graphical depiction of the survey measures is provided in Figure 2.4. Note that the measure is standardized to have a mean equal to zero and a standard deviation of one across the sample. Further note that we adjust the measure

[19]Note that due to Brexit, the UK is not covered in the 2021 wave.

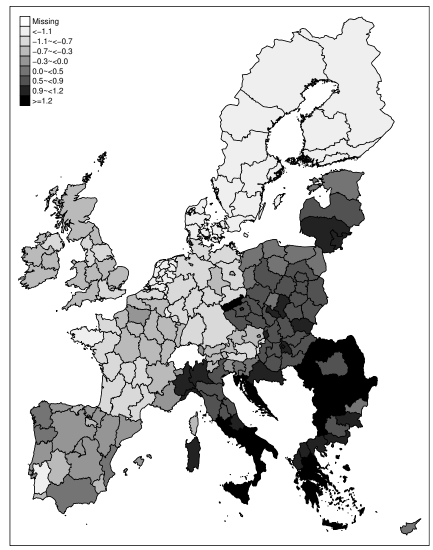
Figure 2.4: *REGIONAL CORRUPTION SURVEY MEASURE FOR EU COUNTRIES*

The figure depicts the corruption measure of the *European Quality of Governance Index* (EQI). The different panels depict different waves of the survey. For Belgium, Germany, and the UK the measure is depicted at the NUTS 1 level. For all other 25 EU countries the measure is depicted at the NUTS 2 level. There is no data available for the UK in 2021. The measure is standardized to have a mean equal to zero and a standard deviation of one. Further note that the raw measure provided by the EQI database is multiplied by -1, such that a higher value corresponds to a higher perceived corruption level. Maps plotted with the *tmap* package for *R* (Tennekes, 2018).

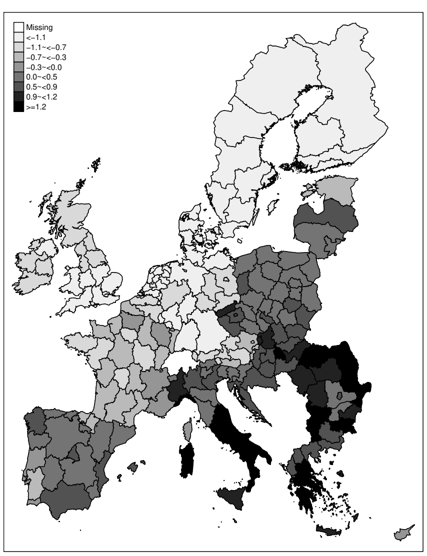
Panel A: 2010



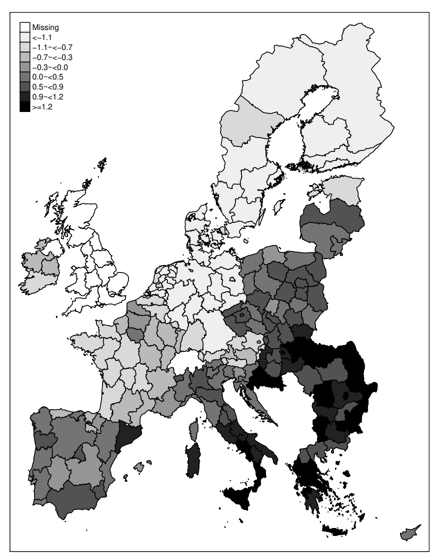
Panel B: 2013



Panel C: 2017



Panel D: 2021



such that a higher value corresponds to a higher perceived corruption level. At first glance, it is very apparent that the southern and eastern EU countries have the highest perceived levels of corruption. The Scandinavian countries as well as Germany and Austria, on the other hand, are among the countries with the lowest corruption levels. The ordering of the countries seems fairly time consistent between the waves, with the exception of the Baltic states that exhibit decreasing levels of corruption over time. Another interesting observation that we exploit in our empirical strategy is the fact that there is a lot of within country and year variation. The country with the most pronounced within country differences is Italy with its moderate corruption levels in the north and very high corruption levels in the south. In 2010, the difference between the Italian NUTS 2 regions with the highest corruption value (Apulia) and the lowest corruption value (South Tyrol) amounts to 2.775 standard deviations. Other countries that have within country and year differences of more than one standard deviation in one or more years are Bulgaria, France, Portugal, Romania, and Spain. To be able to carry out our analysis of the EEITRs with more years than the ones in which the surveys were conducted, we linearly interpolate missing years between the survey years. While we do think that this is a legitimate imputation approach given the moderate variation of individual regions between waves, we do also carry out our analysis using only the observed years.[20]

Statutory income tax rate (τ_{ct} or in the case of Germany τ_{rt}): Since our regression setup controls for country-industry-year fixed effects, country-year-specific variables such as the statutory tax rate, denoted by τ_{ct}, generally drop out. However, in the case of Germany, the statutory income tax rate varies across regions. More precisely, in addition to the country-wide statutory tax rate, each German municipality (*Gemeinde*) sets their trade tax (*Gewerbesteuer*) which is added to the country-wide rate to obtain the final, municipality-specific income tax rate. For our purpose, we aggregate the municipality-specific tax rates to the NUTS 2 level by taking population-weighted averages across all municipalities located

[20]We find that the results with and without the imputed years are highly similar (see Section 2.4).

in a given NUTS 2 region. We denote the resulting NUTS 2 level statutory tax rates for Germany by τ_{rt}. The data on the region-specific tax rates for Germany is obtained from the Research School of International Taxation's (RSIT) *International Tax Institutions* (ITI) database (Wamser et al., 2023). The municipality level population data is obtained from the Federal Statistical Office of Germany.

Net present value of depreciation allowances (δ_{rit}): Our EEITRs are based on the EBITDA, i.e., a raw profit measure that does not take asset depreciation and interest deductions into account. As reasoned above, the choice of this profit variable is deliberate, as in particular depreciation figures may differ between the available financial accounting data and the unobserved tax returns due to differences in the respective accounting rules and practices (Graham et al., 2012). Since consequently our EEITRs capture depreciation and interest deductions, we control for net present values (NPV) of depreciation allowances per unit of investment to avoid omitted variable bias in the analysis. This measure accounts for asset depreciation as well as interest deductions that could legally be claimed and is constructed by combining country level tax code information with region-industry-specific asset and financing structures. Note that using time-constant rather than time-varying asset and financing structures ensures that the NPVs of depreciation allowances are exogenous, as they do not capture dynamic tax-planning behavior of firm entities (Mc Auliffe et al., 2022). For the construction of the measure we distinguish two financing modes, financing via retained earnings and debt financing, and seven different asset categories: Buildings, Machinery, Office equipment, Computer equipment, Intangible fixed assets, Vehicles, and Inventory. Formally, the NPVs of depreciation allowances can be stated as follows:

$$\delta_{rit} = ES_{ri} \sum_{a=1}^{7} w_{ari} \cdot A^{E}_{act} + DS_{ri} \sum_{a=1}^{7} w_{ari} \cdot A^{D}_{act}. \tag{2.5}$$

A^{E}_{act} and A^{D}_{act} denote the NPV of depreciation allowances for an investment in asset type a in country c in year t that is purely financed through retained earnings or debt financing, respectively. They are obtained from the RSIT's *ITI* database (Wamser et al., 2023). It is important to note that these asset- and financing-mode-specific NPVs of depreciation allowances are determined purely by the national tax codes and are therefore identical for all firm entities located in a given country c in year t. The region-industry-specific variation in the δ_{rit}'s stems from the typical asset and financing structures, i.e., from taking into account that firm entities that are located in different regions and operate in different industries differ in terms of their typical asset and financing compositions. Formally, in (2.5), this heterogeneity enters the equation via the region-industry-specific shares of the different asset types that the typical asset structure is comprised of, w_{ari}, as well as the via region-industry-specific shares of financing through retained earnings, ES_{ri}, and debt, DS_{ri}.[21] For the calculation of DS_{ri} we first calculate region-industry-year-specific long-term debt ratios for each year of the sample period (2009 to 2018) using *Orbis*.[22] These year-specific debt ratios are then aggregated to the time-constant DS_{ri}'s by taking unweighted averages over all years, similar to Mc Auliffe et al. (2022). ES_{ri} is then obtained by subtracting DS_{ri} from unity. The calculation of the region-industry-specific asset structures is undertaken in two steps. First, we obtain region-industry-specific shares for the asset types Inventory, Intangible fixed assets, as well as for the whole tangible fixed asset stock from *Orbis* using a similar aggregation approach as for the financing structures.[23] Since *Orbis* does not provide more detailed information on the composition of the tangible fixed asset stock, we use the country-industry-specific weights proposed by Mc Auliffe et al. (2022) to further divide the tangible fixed asset share into Buildings, Machinery, Office equipment, Computer equipment, and Vehicles.

Forward-looking Effective Marginal Tax Rate ($FL\ EMTR_{rit}$): As an alternative to con-

[21] Note that for each region-industry combination it holds that $\sum_{a=1}^{7} w_{ari} = 1$ and $ES_{ri} + DS_{ri} = 1$.

[22] More precisely, the long-term debt ratio is defined as the ratio of the *Orbis* variables non-current liabilities over total assets.

[23] In more detail, the inventory share, the intangible fixed asset share, and the tangible fixed asset shares are obtained by dividing the *Orbis* variables stocks of current assets (i.e., inventories), intangible fixed assets, and tangible fixed assets by the sum of these three variables, respectively. The approach for obtaining the asset structure from *Orbis* was first introduced by Egger et al. (2009) as well as Egger and Loretz (2010).

trolling for the statutory tax rate as well as the NPV of depreciation allowances separately, we also construct forward-looking effective marginal tax rates ($FL\ EMTR_{rit}$) that combine these two measures. Forward-looking EMTRs quantify the income tax burden a firm would face on a hypothetical marginal investment that just breaks even.[24] We calculate the EMTRs using the simple representation proposed by Mc Auliffe et al. (2022):

$$FL\ EMTR_{rit} = \frac{(\tau_{ct} - \tau_{ct}\delta_{rit})}{(1 - \tau_{ct}\delta_{rit})}. \tag{2.6}$$

δ_{rit} denotes the region-industry-year-specific NPV of depreciation allowances that is discussed above and τ_{ct} denotes the statutory tax rate. Note that for the calculation of the EMTRs for Germany we use the NUTS 2 specific statutory tax rates, τ_{rt} (see above).

Patent box regime ($IP\ box_{ct}$): Over the past two decades, more and more countries introduced so-called "patent boxes", i.e., special tax regimes that aim at incentivizing R&D by taxing patent revenues at preferential rates. In our analysis, we use a dummy ($IP\ box_{ct}$) that is equal to unity if country c has a patent box in place in year t and zero if not. We obtain the data on the patent boxes from the RSIT's *ITI* database (Wamser et al., 2023).

Patent density ($Patent\ density_{rit}$): Next, we construct a region-industry-year-specific variable that captures patent activity. In detail, we define a patent density measure which we define as the share of firm-entities in *Orbis* that hold at least one patent. The source of the patent data is Bureau van Dijk's *Orbis Intellectual Property* database.

Share of firm entities with strictly negative EBITDA in $t-1$ in firm entities that have strictly positive EBITDA in t ($Share\ loss\ in\ t-1_{rit}$): Due to the confidentiality of tax accounting, it is not possible to determine if and to which degree loss carrybacks and loss carryforwards are used by firm entities. Since our EEITRs are based on observations with

[24]The concept of the EMTR is formally developed in the seminal contributions by Devereux and Griffith (1998), Hall and Jorgenson (1967), King (1974), King and Fullerton (1984), and OECD (1991). Note that forward-looking tax measures do not aim to proxy the actual income tax burden that an individual firm entity faces, but rather state the incentive of the tax code to invest in a simplified setup where the tax code is applied as intended and where the absence of tax base adjustments outside of depreciation is assumed.

strictly positive EBITDA, however, we can expect loss carrybacks to play a negligible role.[25] It cannot be said with certainty that losses in $t-1$ are used against profits in t via a loss carryforwards to reduce the tax base in t, as, for instance, the losses in $t-1$ could have been used against profits in $t-2$ or earlier periods using loss carrybacks. However, Rechbauer and Rünger (2023) demonstrate that the earnings in $t-1$ serve as reliable indicator for the existence of loss carryforwards in t. As proxy for the extent of the use of loss carryforwards we calculate the region-industry-year-specific share of firm entities that report strictly negative EBITDA in $t-1$ in firm entities that have a strictly positive EBITDA figure in t.

GDP ($log\ GDP_{rt}$) and GDP per capita ($log\ GDP\ p.c._{rt}$): As additional controls, we use the logarithm of real GDP as well as real GDP per capita as measures for the economic development of the NUTS 2 regions. Both variables are obtained from the *ARDECO* online database.

Tax morale ($Tax\ morale_r$): As last variable, we prepare a regional tax morale measure that we derive from from the joint European Values Survey/ World Values Survey 2017-2022 dataset (EVS/WVS, 2022). More precisely, we prepare the question item that asks whether cheating on tax if you have the chance is justifiable.[26] Surveyed individuals could answer this question on a numeric scale from one ("Never justifiable") to ten ("Always justifiable"). The survey results are provided at the individual respondents level and also include information about the NUTS 2 region or in the case of Germany about the NUTS 1 region where the interview was conducted. We aggregate the individual level data to the regional level by taking unweighted averages across all respondents located in the respective region, excluding the answer options "no answer" and "don't know".[27] Note that while

[25]It has to be noted that it is possible that firm entities that report a strictly positive EBITDA file a loss carryback in the same period. That is in the case when tax-relevant deductions are higher than the EBITDA, which results in a negative income tax base. However, those cases seem to be an exception, as almost all firm entities that report strictly positive EBITDA figures also report positive tax liabilities in the data.

[26]Note that constructing a proxy for tax morale from survey responses concerning the justifiability of evading taxes is an approach commonly used in the literature (see, e.g., Torgler, 2007; Torgler et al., 2008).

[27]Note that due to lack of data, we cannot calculate averages that are weighted by socio-demographic characteristics.

the survey was conducted over a time span of several years, the resulting measure is time-constant, as regions in different countries were generally only surveyed once. Further note that we set our aggregated region-specific tax morale measure missing for regions with less than 50 respondents to only include meaningful values into our analysis. In a last step, we standardize the variable such that in the final regression sample it has a mean equal to zero and a standard deviation of one and multiply it with minus one such that a higher value corresponds to a higher tax morale (i.e., higher agreement that cheating on taxes is never justifiable).

Table 2.2: *DESCRIPTIVE STATISTICS*

The table depicts descriptive statistics on all the variables used for the empirical analysis. Panel A reports summary statistics. Panel B depicts Pearson correlation coefficients for key variables. Definitions of the variables are provided in Section 2.3.2.

Panel A: Summary statistics

	Observations	Mean	(sd)
$EEITR_{rit}$	19,095	0.128	(0.050)
τ_{ct} (τ_{rt} for DEU)	19,095	0.261	(0.072)
δ_{rit}	19,095	0.557	(0.134)
$FL\ EMTR_{rit}$	19,095	0.134	(0.051)
$Share\ loss\ in\ t-1_{rit}$	19,095	0.123	(0.052)
$IP\ box_{ct}$	19,095	0.503	(0.500)
$Patent\ density_{rit}$	19,095	0.008	(0.023)
$Patent\ density_{rit} \times IP\ box_{ct}$	19,095	0.002	(0.007)
$log\ GDP_{rt}$	19,095	24.351	(0.964)
$log\ GDP\ p.c._{rt}$	19,095	9.994	(0.566)
$Corruption_{rt}$	19,095	0.050	(0.956)
$Tax\ morale_{r}$	15,476	0.000	(1.000)

Panel B: Correlation matrix (19,095 observations)

	$EEITR_{rit}$	τ_{ct}	δ_{rit}	$FL\ EMTR_{rit}$	$Corruption_{rt}$
$EEITR_{rit}$	1.000				
τ_{ct} (τ_{rt} for DEU)	0.381	1.000			
δ_{rit}	-0.035	0.196	1.000		
$FL\ EMTR_{rit}$	0.342	0.737	-0.469	1.000	
$Corruption_{rt}$	-0.117	-0.317	-0.003	-0.240	1.000

Finally, descriptive statistics on the estimation sample are provided in Table 2.2. The correlation matrix in Panel B suggests a negative relationship between the EEITRs and the corruption measure, as expected. Furthermore, the signs of the correlation coefficients between the EEITRs and the statutory tax measures are in line with basic intuition: a higher NPV of depreciation allowances results in more deductions from the tax base and therefore lower effective tax payments, which is indicated by the negative sign of the coefficient. On the other hand, a higher statutory tax rate or a higher forward-looking EMTR mean that a higher share of the profits is taxed away, which is in line with the positive sign of the respective coefficients.

2.4. Results

Our estimation results are provided in Table 2.3. Note that since our dependent variable, the EEITR, is a regression coefficient itself, we bootstrap the standard errors. The specification

Table 2.3: *ESTIMATION RESULTS*

The table presents OLS estimates. The dependent variable is the region-industry-year specific EEITR ($EEITR_{rit}$). Bootstrapped standard errors (based on 10,000 bootstrap replications) are reported in parentheses. *** denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level. Specification (2) only uses the years 2010, 2013, and 2017, i.e., the years of the sample period for which the corruption survey measure is observed. Specification (3) is identical to (2) but additionally uses linearly interpolated years, which results in all years 2010 to 2018 being included in the sample. Specifications (4) and (5) exclude observations corresponding to Germany and Italy, respectively. Specification (6) excludes the NACE Rev. 2 sections A, B, K, P, and Q (for section descriptions, see Appendix 2.6.2). All specifications control for country-industry-year fixed effects (FEs). Definitions and descriptive statistics on the explanatory variables are provided in Section 2.3.2.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
τ_{ct} (τ_{rt} for DEU)	0.406***	0.393***	0.408***		0.417***	0.425***		0.386***
	(0.053)	(0.100)	(0.057)		(0.056)	(0.059)		(0.052)
δ_{rit}	-0.069***	-0.074***	-0.068***	-0.047***	-0.108***	-0.075***		-0.083***
	(0.008)	(0.014)	(0.008)	(0.008)	(0.009)	(0.009)		(0.008)
$FL\ EMTR_{rit}$							0.300***	
							(0.030)	
$Share\ loss\ in\ t-1_{rit}$	-0.030***	-0.022*	-0.024***	-0.028***	-0.012	-0.025***	-0.022***	-0.024***
	(0.007)	(0.012)	(0.008)	(0.007)	(0.008)	(0.008)	(0.007)	(0.008)
$Patent\ density_{rit}$	0.042***	0.035**	0.031***	0.023***	0.002	0.041***	0.029***	0.020**
	(0.009)	(0.015)	(0.009)	(0.008)	(0.007)	(0.010)	(0.009)	(0.008)
$Patent\ density_{rit}\ \times$	-0.086**	-0.139*	-0.130***	-0.116***	-0.150***	-0.165***	-0.131***	-0.014
$IP\ box_{ct}$	(0.039)	(0.081)	(0.044)	(0.044)	(0.053)	(0.050)	(0.043)	(0.053)
$log\ GDP_{rt}$	0.008***	0.009***	0.009***	0.009***	0.008***	0.010***	0.009***	0.007***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$log\ GDP\ p.c._{rt}$	0.007***	0.006***	0.005***	0.005***	0.005***	0.006***	0.005***	0.010***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$Corruption_{rt}$		-0.004***	-0.003***	-0.003***	-0.004***	-0.004***	-0.003***	-0.006***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$Tax\ morale_r$								0.001***
								(0.000)
Country-industry-year FEs	YES	YES	YES	YES	YES	YES	YES	YES
Adjusted R^2	0.864	0.863	0.866	0.874	0.840	0.852	0.866	0.886
Observations	21,056	6,379	19,095	17,936	16,213	14,888	19,095	15,476

in column (1) uses all years 2009 to 2018. The coefficient on the statutory tax rate is positive and suggests that a one percentage point increase in the statutory tax rate is associated with an increase in the EEITR of 0.41 percentage points. Note that due to the country-industry-year fixed effects, this coefficient is solely determined by the regional variation in the German statutory tax rate due to the regional trade tax. The NPV of depreciation allowances exhibits a negative sign, which is explained by the fact that more depreciation reduces the tax base and therefore yields lower EEITRs. Note that since the statutory depreciation rules are determined at the country level, the regional variation in the depreciation allowances is solely due to regional variation in financing and asset structures. The sign of the coefficient on the share of firm entities that have a strictly negative EBITDA in $t-1$ out of the entities that have a strictly positive EBITDA in t is significant negative. This suggests that at least some of these firm entities use the losses of the previous period to reduce the tax base in the current period via loss carryforwards. The results further imply that EEITRs are higher in regions with higher patent densities. However, for regions in countries that have low tax regimes for patent revenue in place, a higher patent density is associated with lower EEITRs (compare magnitude of coefficients on $Patent\ density_{rit}$ and $Patent\ density_{rit} \times IP\ box_{ct}$). This finding suggests that firm entities make use of the patent box regimes when filing their tax returns, which results in a more favorable taxation of revenues associated with patents and therefore lower EEITRs. We further find that EEITRs are higher in regions with higher GDP and higher GDP per capita. Columns (2) and (3) introduce the corruption measure, with column (2) only using years in which the EQI survey was conducted (i.e., 2010, 2013, and 2017) and column (3) also using linearly interpolated years, which results in a broader sample that spans across all years 2010 to 2018. Both models yield highly similar results with statistically significant coefficients on the corruption measure, however, the model in column (2) yields a (in absolute terms) slightly larger coefficient than the one in column (3) (-0.004 vs. -0.003). The coefficient estimates on the corruption measure suggest that a one standard deviation increase in the measure is associated with – depending on the specification – a 0.4 or 0.3 percentage point decrease in the EEITR. This is a sizeable effect, considering that there are several EU countries that exhibit within country and year differences in corruption of more than one standard deviation. Since we control for the legal ways to decrease the EEITR, the likely explanation for lower EEITRs in high corruption regions is tax evasion via overstated deductions. It has to be noted that the effect of corruption on tax evasion has been established in the empirical literature before (see, e.g., Uslaner, 2010; Alm et al., 2016). However, our approach fundamentally differs from the previous contributions in the sense that we do not measure tax evasion behavior by how much profits or sales are underreported, but instead take the reported EBITDA on which we base our EEITRs as given and estimate the effect of corruption on these EEITRs (see discussion above). The subsequent specifications in columns (4) to (7) present a number of robustness checks. Since the statutory tax rate in Germany is aggregated from the municipality level, at which the trade tax is levied, to the NUTS 2 level, at which the analysis is carried out, there may be a certain degree of measurement error. For this purpose, we redo specification (3) without any observations corresponding to Germany. The results, depicted in column (4), are similar to those in (3), with the coefficient on the corruption measure being unchanged.[28] As shown above, Italy is the country with the highest within country variation in the corruption measure and is also among the countries with the highest within country and industry variation in the EEITRs. To check if the coefficient on the corruption measure is driven by Italy, we rerun specification (3) excluding Italy. Interestingly, the coefficient on corruption is even larger in magnitude when excluding Italy (-0.004; see column (5)). This rules out that the corruption results are exclusively driven by Italy. Specification (6) excludes certain industries that are known to often be subject to differential tax treatment, such as the agricultural or the financing sector.[29] The results in column (6) show that the exclusion of these industries yields a (in absolute terms) slightly larger coefficient of -0.004 on the corruption measure

[28]Note that since all other countries have country-year-specific statutory tax rates, the coefficient on the statutory tax rate is not identified due to the fixed effects.

[29]In detail, the excluded industries are the NACE Rev. 2 sections A, B, K, P, and Q (for section descriptions, see Appendix 2.6.2). Note that it is common in tax related empirical analyses to exclude these industries, see, e.g., Liu (2020), Mc Auliffe et al. (2022), or Steinmüller et al. (2019).

compared to the base specification in (3) that includes all industries (coefficient: -0.003). The model presented in column (7) applies forward-looking, region-industry-year-specific EMTRs that combine both the statutory tax rate and the NPV of depreciation allowances instead of controlling for these two measures individually. The negative sign allows for a similar interpretation as of the sign of the statutory tax rate in that a higher EMTR corresponds to a higher share of the EBITDA being taxed away, which results in higher EEITRs. Compared to (3), the magnitude of the coefficient on corruption remains unchanged. Finally, column (8) adds a second survey measure, namely the tax morale index. The sign on the tax morale measure is positive and significant, which suggests that EEITRs are on average higher in regions where the citizens agree more with the statement that cheating on taxes is never justifiable, which is in line with the previous literature (Richardson, 2006; Torgler, 2007; Torgler et al., 2008). However, the coefficient of 0.001, which suggests that a one standard deviation increase in the measure is associated with a 0.1 percentage point increase in the EEITRs, is quite small. It has to be noted that the measure is time-constant and that the survey from which it is taken was conducted at the end of our sample period, making it potentially imprecise for the earlier sample years. Additionally, the aggregation from the individual respondent level to the NUTS level was done without taking socio-demographic characteristics into account due to lack of data, which likely further reduces the accuracy of the measure. Interestingly, the coefficient on the corruption measure is twice as large in specification (8) compared to specification (3), which outside of not controlling for tax morale uses the same variables (-0.006 vs. -0.003). However, it has to be noted that due to data availability of the tax morale measure, specification (8) is estimated using a much smaller sample.

In the last step of our analysis, we illustrate the size of the effect of corruption on EEITRs by carrying out a simple back-of-the-envelope calculation. For this purpose, we assume that the relationship between corruption and the EEITRs is causal. Using this assumption, we calculate how much higher country-wide EEITRs of EU countries would

Table 2.4: *OBSERVED VERSUS HYPOTHETICAL COUNTRY LEVEL EEITRs*

The table depicts country-year-specific observed EEITRs ($EEITR_{ct}$), country-year-specific hypothetical EEITRs ($EEITR_{ct}^{\text{hypoth}}$), as well as the change from the observed to the hypothetical EEITR in percent (Δ in %) for the years 2010, 2013, and 2017. The $EEITR_{ct}$ are obtained by taking weighted averages across all region-industry-year-specific EEITRs ($EEITR_{rit}$; see Section 2.3.2). The weights are calculated as the share of the region-industry-year-specific sums of strictly positive tax liabilities (obtained from *Orbis*) in the total country-year-specific sum, considering only values corresponding to region-industry-year combinations for which we obtain EEITRs to ensure the weights add up to unity. Note that we drop tax liability values in the top and bottom percentile to mitigate the influence of outliers. The $EEITR_{ct}^{\text{hypoth}}$'s are weighted averages of region-industry-year-specific hypothetical EEITRs ($EEITR_{rit}^{\text{hypoth}}$) that are calculated using the same aforementioned weighting. The $EEITR_{rit}^{\text{hypoth}}$'s reflect the hypothetical scenario in which the respective regional corruption levels are adjusted from the observed level to the average across all Scandinavian EU regions (i.e., regions of the countries Denmark, Finland, and Sweden). For details of the calculation, see Section 2.4.

	2010			2013			2017		
	$EEITR_{ct}$	$EEITR_{ct}^{\text{hypoth}}$	Δ in %	$EEITR_{ct}$	$EEITR_{ct}^{\text{hypoth}}$	Δ in %	$EEITR_{ct}$	$EEITR_{ct}^{\text{hypoth}}$	Δ in %
AUT	n.a.	n.a.	n.a.	0.151	0.155	3.14	0.145	0.149	2.59
BEL	0.141	0.146	3.58	0.140	0.143	2.37	0.163	0.166	2.06
BGR	0.060	0.075	24.92	0.059	0.073	24.41	n.a.	n.a.	n.a.
CZE	0.106	0.117	10.62	0.105	0.115	9.14	0.109	0.118	8.19
DEU	0.155	0.158	1.95	0.157	0.160	1.82	0.179	0.181	1.24
ESP	0.127	0.133	4.84	0.136	0.142	4.27	0.159	0.168	5.35
EST	0.102	0.109	7.38	0.094	0.100	6.99	0.095	0.100	5.13
FRA	0.140	0.144	3.40	0.135	0.138	2.79	0.148	0.153	3.13
GBR	0.189	0.194	2.44	0.167	0.170	2.14	0.152	0.154	1.47
GRC	0.140	0.151	7.95	0.147	0.159	8.24	0.180	0.192	6.42
HRV	0.108	0.120	11.18	0.101	0.112	10.79	0.090	0.100	11.00
HUN	0.078	0.089	13.83	0.064	0.073	15.05	0.061	0.072	18.09
IRL	0.111	0.114	2.68	0.133	0.137	2.73	0.121	0.124	2.55
ITA	0.218	0.228	4.52	0.215	0.225	4.63	0.195	0.205	5.34
LUX	0.151	0.153	1.09	0.187	0.187	0.37	0.175	0.176	0.46
LVA	0.070	0.082	16.02	0.082	0.091	11.09	0.082	0.091	10.82
NLD	0.142	0.148	4.06	0.145	0.146	0.65	0.190	0.192	1.03
POL	0.141	0.151	7.26	0.140	0.149	6.37	0.130	0.138	6.13
PRT	0.120	0.126	5.66	0.152	0.158	3.88	0.140	0.146	4.61
ROU	0.126	0.140	11.19	0.099	0.112	12.73	0.065	0.076	16.60
SVK	0.092	0.103	11.14	0.110	0.120	8.98	0.127	0.137	8.18
SVN	0.109	0.116	6.72	0.092	0.098	7.26	0.110	0.117	6.75

be if all of the regional corruption levels decreased to the mean corruption level of the Scandinavian countries Denmark, Finland, and Sweden of the respective year. In a first step, we calculate hypothetical EEITRs ($EEITR_{rit}^{\text{hypoth}}$) that consist of the sum of the observed EEITR ($EEITR_{rit}$) plus the difference between the respective observed regional corruption level ($Corruption_{rt}$) and the mean Scandinavian corruption level ($Corruption_{t}^{\text{Scandinavia}}$), multiplied by the marginal effect of a one unit decrease in corruption on the EEITR (i.e., 0.004; see Table 2.3, Column (2)). Formally, we get

$$EEITR_{rit}^{\text{hypoth}} = EEITR_{rit} + (Corruption_{rt} - Corruption_{t}^{\text{Scandinavia}}) \cdot 0.004. \tag{2.7}$$

Next, we aggregate both the observed region-industry-year-specific EEITRs as well as the hypothetical counterparts to the respective country-year-levels by taking weighted averages. The weights that we use proxy the contribution of the different region-sector-combinations to the countries' overall corporate income tax revenue in the respective year and are derived from *Orbis* data.[30] Table 2.4 provides a juxtaposition of the aggregated country-level observed EEITRs and the hypothetical counterparts for the years 2010, 2013, and 2017. The increase from the observed to the hypothetical EEITRs is substantial for high-corruption-countries like Bulgaria, the Czech Republic, Greece, Croatia, Hungary, Italy, Latvia, Poland, Romania, or Slovakia, with differences of one percentage point or more. This suggests substantial increases in the corporate income tax revenues collected by these countries in the hypothetical scenario compared to the observed one.

2.5. Conclusions

This paper provides evidence on tax evasion via overstated deductions, a tax evasion channel that was so far unexplored in the empirical literature on tax evasion. We show that this tax evasion strategy is more extensively used in regions of the EU with higher levels of corruption. Our analysis suggests that policymakers seeking to combat tax evasion and increase corporate income tax revenue should focus on tackling high corruption environments within the respective country. Methodologically, this paper proposes a novel approach for calculating EEITRs in scenarios where only an aggregated tax liability variable is available that contains all types of taxes that the respective firm entities paid in a given year.

[30] In detail, we construct region-industry-year-specific sums of the firm-entity-level variable total tax liability, considering only observations with strictly positive values. The weights are then obtained as the share of these region-industry-year-specific sums in the total country-year-specific sum, with the latter only comprising values corresponding to region-industry-year combinations for which we obtain EEITRs to ensure the weights add up to unity. Note that we drop tax liability values in the top and bottom percentile to mitigate the influence of outliers.

2.6. Appendix

2.6.1. Range of NUTS 2-specific Ratio Tax Liability over EBITDA by Country and Industry in 2013

Table 2.5: ***RANGE OF THE NUTS 2-SPECIFIC RATIO TAX LIABILITY OVER EARNINGS BEFORE INTEREST, TAXES, DEPRECIATION, AND AMORTIZATION BY COUNTRY AND INDUSTRY IN 2013***

The table depicts descriptive statistics on the firm-entity-specific ratio tax liability over Earnings Before Interest, Taxes, Depreciation, and Amortization (EBITDA) for different NUTS 2 regions (version 2016) of EU 28 countries. Panel A states the percentage point difference between the NUTS 2 regions with the highest and the lowest median of this ratio for a given country and industry (NACE Rev. 2 section). Panel B states the respective percentage point difference for the mean of NUTS 2-specific ratios. All data corresponds to the year 2013. Firm entities belonging to MNEs are excluded. Only observations with strictly positive EBITDA are used for the calculation. Observations of the depicted ratio in the top and bottom one percentile were excluded from the sample. A minimum of 25 firm entity observations per region and industry combination was required. Country-industry combinations with less than two observed ratios are set missing. Note that Cyprus, Estonia, Luxembourg, Latvia, and Malta are excluded, as these countries only have one NUTS 2 region, i.e., the whole country. Furthermore, Lithuania is excluded due to poor data coverage. The sections O, T, and U are not depicted due data coverage. For descriptions of the sections, see Appendix 2.6.2. The source of the data is *Orbis*.

Panel A: Difference between maximum and minimum of NUTS 2-specific median of tax liability over EBITDA ratio																		
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	P	Q	R	S
AUT	n.a.	n.a.	9.1	n.a.	n.a.	8.0	10.1	11.4	n.a.	n.a.	n.a.	4.1	0.0	n.a.	n.a.	n.a.	n.a.	n.a.
BEL	4.4	n.a.	4.7	0.2	5.1	6.6	3.7	6.8	3.3	13.1	4.4	2.1	3.8	10.5	8.0	2.5	10.4	5.4
BGR	6.0	n.a.	1.3	n.a.	n.a.	n.a.	0.1	0.0	n.a.	n.a.	n.a.	0.0	0.7	n.a.	n.a.	n.a.	n.a.	n.a.
CZE	3.6	n.a.	1.2	2.4	5.9	3.4	3.4	3.9	0.0	4.4	7.1	2.6	4.2	6.1	9.5	2.5	8.1	5.3
DEU	6.5	n.a.	11.5	9.3	10.5	7.5	7.6	10.1	12.3	16.8	17.0	5.5	8.8	10.5	1.1	2.1	16.9	9.7
DNK	n.a.	n.a.	9.0	n.a.	n.a.	7.2	4.3	12.1	3.1	2.7	8.2	2.8	3.7	5.0	n.a.	1.4	n.a.	n.a.
ESP	6.6	15.6	4.3	7.7	8.1	6.4	5.7	5.2	6.7	12.8	7.6	6.2	6.9	8.3	10.7	6.7	9.9	6.6
FIN	8.1	0.4	7.5	1.8	7.7	2.1	2.4	2.3	8.8	3.4	3.5	4.2	1.5	2.5	1.0	1.2	2.3	1.3
FRA	2.3	13.3	6.8	3.5	12.7	10.7	5.5	6.2	4.3	10.3	9.7	4.2	13.6	8.9	9.1	10.8	4.8	1.0
GBR	8.3	n.a.	8.0	5.4	n.a.	2.5	3.6	10.8	14.7	2.5	19.6	9.5	1.8	5.0	2.8	10.6	14.8	7.3
GRC	n.a.	n.a.	10.3	1.3	n.a.	12.5	18.6	15.0	0.4	1.2	n.a.	8.9	1.5	12.9	n.a.	5.3	n.a.	n.a.
HRV	0.3	0.4	0.5	4.0	0.4	0.2	0.3	0.6	1.0	1.2	4.1	1.0	0.7	2.3	0.7	0.9	0.4	2.9
HUN	1.4	1.3	0.9	2.1	1.3	0.7	1.0	1.3	1.7	1.5	0.8	0.3	0.6	1.0	1.8	0.5	1.4	1.1
IRL	n.a.	n.a.	0.3	n.a.	n.a.	n.a.	2.8	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	0.8	n.a.	n.a.	n.a.	n.a.
ITA	7.3	9.8	13.9	11.4	10.3	8.7	10.8	19.9	10.9	15.9	10.3	6.7	13.6	17.1	19.7	22.9	8.9	12.6
NLD	n.a.	n.a.	5.6	n.a.	n.a.	3.5	5.4	3.3	n.a.	3.4	4.6	n.a.	15.3	2.8	n.a.	n.a.	n.a.	n.a.
POL	19.8	2.5	3.0	5.6	6.7	4.6	3.4	6.2	3.8	7.5	2.7	7.2	3.4	2.6	8.9	11.1	4.9	n.a.
PRT	3.7	1.6	9.1	4.5	4.3	7.0	7.6	10.3	5.9	13.9	10.5	7.6	12.5	12.1	4.9	8.9	10.6	8.9
ROU	2.8	5.7	2.0	6.5	2.0	1.5	1.1	1.9	2.2	2.0	3.2	1.1	1.8	1.3	2.7	1.8	2.1	3.0
SVK	0.7	n.a.	2.9	2.1	2.3	1.7	1.1	1.1	0.1	4.6	2.7	1.2	1.8	1.1	5.6	1.8	1.6	1.1
SVN	0.9	n.a.	0.6	1.0	1.2	0.9	0.1	0.7	1.2	0.4	1.4	0.2	0.3	0.5	0.2	1.3	1.0	2.3
SWE	6.0	5.7	2.5	2.3	5.9	3.4	1.7	4.0	1.4	2.4	1.9	1.9	1.3	5.0	3.7	2.3	7.2	2.3
Panel B: Difference between maximum and minimum of NUTS 2-specific mean of tax liability over EBITDA ratio																		
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	P	Q	R	S
AUT	n.a.	n.a.	8.4	n.a.	n.a.	8.5	6.5	6.6	n.a.	n.a.	n.a.	0.5	6.8	n.a.	n.a.	n.a.	n.a.	n.a.
BEL	9.0	n.a.	3.8	2.4	3.1	5.4	3.9	4.8	3.8	10.7	3.1	2.3	3.0	8.7	5.8	1.9	7.6	4.6
BGR	6.5	n.a.	2.9	n.a.	n.a.	n.a.	1.6	0.2	n.a.	n.a.	n.a.	1.2	0.2	n.a.	n.a.	n.a.	n.a.	n.a.
CZE	1.6	n.a.	0.6	2.4	5.0	1.4	1.6	1.3	1.0	2.0	2.3	1.8	1.8	3.2	7.4	2.0	10.0	2.7
DEU	2.5	n.a.	6.6	15.6	9.9	6.2	5.1	7.8	8.5	9.9	11.5	10.8	11.7	11.0	2.6	13.4	19.2	8.3
DNK	n.a.	n.a.	6.4	n.a.	n.a.	7.6	2.7	6.0	3.0	4.6	2.5	5.5	3.5	15.5	n.a.	1.2	n.a.	n.a.
ESP	10.3	11.5	5.5	12.2	8.8	6.3	4.9	6.0	11.3	14.9	7.8	5.8	5.9	10.7	12.4	6.5	11.0	7.0
FIN	4.2	0.8	6.7	0.7	3.8	1.6	1.0	0.9	5.7	2.1	3.3	1.2	2.5	3.0	1.3	1.2	1.1	1.5
FRA	10.1	9.8	6.9	8.2	15.3	7.5	7.1	5.6	9.6	7.0	9.3	5.7	8.1	8.6	6.3	9.6	10.0	6.8
GBR	6.5	n.a.	9.4	3.6	n.a.	6.7	3.4	10.3	12.2	8.8	11.8	11.9	5.3	8.9	7.0	9.3	8.3	7.7
GRC	n.a.	n.a.	7.0	0.4	n.a.	9.7	8.6	4.4	3.5	1.4	n.a.	0.2	2.9	5.2	n.a.	4.1	n.a.	n.a.
HRV	0.4	0.4	0.3	2.1	0.3	0.1	0.0	0.1	0.9	1.2	1.6	0.4	0.2	0.8	1.0	0.2	0.8	1.0
HUN	2.1	1.9	1.6	8.3	2.4	0.9	1.2	2.3	1.6	2.4	1.5	0.6	1.1	1.1	2.2	1.3	1.6	2.2
IRL	n.a.	n.a.	0.6	n.a.	n.a.	n.a.	2.9	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	0.8	n.a.	n.a.	n.a.	n.a.
ITA	8.9	10.5	11.8	12.5	10.4	8.8	9.7	17.7	9.0	13.3	12.5	5.7	11.3	13.1	13.6	19.4	10.3	11.7
NLD	n.a.	n.a.	9.1	n.a.	n.a.	6.8	3.3	3.9	n.a.	5.9	8.0	n.a.	5.0	2.8	n.a.	n.a.	n.a.	n.a.
POL	21.5	3.7	3.9	8.5	6.5	8.2	3.5	7.1	5.0	3.4	6.5	8.7	4.6	5.5	5.6	7.3	2.3	n.a.
PRT	6.6	1.0	9.5	2.9	5.8	7.2	8.3	7.2	7.3	11.4	13.2	7.8	8.9	10.3	3.6	7.9	9.3	7.4
ROU	1.6	4.4	1.9	5.6	2.8	1.7	2.6	1.8	4.6	2.3	2.9	2.1	0.9	2.0	6.7	2.3	3.8	6.7
SVK	0.7	n.a.	1.6	2.9	2.0	1.0	0.6	1.7	0.3	1.8	3.9	1.4	1.0	1.3	3.3	1.0	2.6	1.8
SVN	1.7	n.a.	1.0	0.9	0.4	0.6	0.1	0.6	0.5	0.1	3.8	0.3	0.5	0.1	0.7	2.5	0.6	1.2
SWE	5.3	2.5	2.8	1.9	5.6	3.6	1.9	3.1	2.3	3.6	2.2	1.3	1.6	5.6	4.1	2.6	6.5	3.9

2.6.2. NACE Rev. 2 (ISIC Rev. 4) Section Descriptions

Table 2.6: *NACE REV. 2 (ISIC REV. 4) SECTION DESCRIPTIONS*

The table depicts the descriptions of the sections of the *Statistical classification of economic activities in the European Community (NACE) Rev. 2* and the *International Standard Industrial Classification of All Economic Activities (ISIC) Rev. 4* that are used throughout this paper. Note that since NACE Rev. 2 was created based on ISIC Rev. 4, the classification systems are equal at the section level.

Section code	section description
A	Agriculture, forestry and fishing
B	Mining and quarrying
C	Manufacturing
D	Electricity, gas, steam and air conditioning supply
E	Water supply; sewerage, waste management and remediation activities
F	Construction
G	Wholesale and retail trade; repair of motor vehicles and motorcycles
H	Transportation and storage
I	Accommodation and food service activities
J	Information and communication
K	Financial and insurance activities
L	Real estate activities
M	Professional, scientific and technical activities
N	Administrative and support service activities
O	Public administration and defence; compulsory social security
P	Education
Q	Human health and social work activities
R	Arts, entertainment and recreation
S	Other service activities
T	Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use
U	Activities of extraterritorial organizations and bodies

2.6.3. Estimation Results Including MNEs

Table 2.7: *ESTIMATION RESULTS INCLUDING MNEs*

The table presents OLS estimates similar to the ones depicted in Table 2.3, however, MNEs were not excluded in the computation of the variables. The dependent variable is the region-industry-year specific EEITR ($EEITR_{rit}$). Bootstrapped standard errors (based on 10,000 bootstrap replications) are reported in parentheses. *** denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level. Specification (2) only uses the years 2010, 2013, and 2017, i.e., the years of the sample period for which the corruption survey measure is observed. Specification (3) is identical to (2) but additionally uses linearly interpolated years, which results in all years 2010 to 2018 being included in the sample. Specifications (4) and (5) exclude observations corresponding to Germany and Italy, respectively. Specification (6) excludes the NACE Rev. 2 sections A, B, K, P, and Q (for section descriptions, see Appendix 2.6.2). All specifications control for country-industry-year fixed effects (FEs). Definitions of the explanatory variables are provided in Section 2.3.2.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
τ_{ct} (τ_{rt} for DEU)	0.386***	0.368***	0.381***		0.386***	0.374***		0.391***
	(0.049)	(0.088)	(0.052)		(0.051)	(0.052)		(0.053)
δ_{rit}	-0.071***	-0.071***	-0.073***	-0.052***	-0.114***	-0.075***		-0.076***
	(0.008)	(0.014)	(0.008)	(0.008)	(0.009)	(0.009)		(0.008)
$FL\ EMTR_{rit}$							0.320***	
							(0.030)	
$Share\ loss\ in\ t-1_{rit}$	-0.032***	-0.030**	-0.028***	-0.030***	-0.018**	-0.032***	-0.027***	-0.024***
	(0.007)	(0.012)	(0.007)	(0.007)	(0.008)	(0.009)	(0.007)	(0.008)
$Patent\ density_{rit}$	0.237***	0.235***	0.213***	0.213***	0.073**	0.237***	0.204***	0.229***
	(0.024)	(0.043)	(0.026)	(0.029)	(0.029)	(0.028)	(0.025)	(0.026)
$Patent\ density_{rit}\ \times$	-0.219***	-0.255***	-0.233***	-0.226***	-0.150***	-0.216***	-0.227***	-0.145***
$IP\ box_{ct}$	(0.040)	(0.083)	(0.043)	(0.045)	(0.053)	(0.047)	(0.042)	(0.047)
$log\ GDP_{rt}$	0.008***	0.009***	0.009***	0.009***	0.008***	0.010***	0.009***	0.006***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$log\ GDP\ p.c._{rt}$	0.007***	0.005***	0.006***	0.005***	0.006***	0.006***	0.006***	0.009***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$Corruption_{rt}$		-0.003***	-0.002***	-0.002***	-0.005***	-0.003***	-0.002***	-0.005***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$Tax\ morale_{r}$								0.001***
								(0.000)
Country-industry-year FEs	YES	YES	YES	YES	YES	YES	YES	YES
Adjusted R^2	0.855	0.854	0.858	0.869	0.829	0.843	0.858	0.887
Observations	22,043	6,699	20,022	18,569	17,117	15,692	20,022	15,501

2.7. Acknowledgements

I am indebted to Peter Egger, Valeria Merlo, and Georg Wamser who are co-authors of Chapter 2. All authors would like to thank Giacomo Brusco, Camila Cisneros, Sean Fahle, Anna Gumpert, Jaqueline Hansen, Michael Knaus, Wilhelm Kohler, Martin Ruf, and Andrea Tulli for helpful comments and suggestions. Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – FOR2783 – is gratefully acknowledged.

2.8. References

Advani, A., Elming, W., & Shaw, J. (2021). The Dynamic Effects of Tax Audits. *Review of Economics and Statistics*, 105(3), 545–561.

Alm, J., Martinez-Vazquez, J., & McClellan, C. (2016). Corruption and firm tax evasion. *Journal of Economic Behavior & Organization*, 124, 146-163.

Alm, J., Liu, Y., & Zhang, K. (2019). Financial constraints and firm tax evasion. *International Tax and Public Finance*, 26, 71-102.

Arulampalam, W., Devereux, M. P., & Maffini, G. (2012). The direct incidence of corporate income tax on wages. *European Economic Review*, 56(6), 1038-1054.

Beck, T., Lin, C., & Ma, Y. (2014). Why do firms evade taxes? The role of information sharing and financial sector outreach. *The Journal of Finance*, 69(2), 763-817.

Bergolo, M., Ceni, R., Cruces, G., Giaccobasso, M., & Perez-Truglia, R. (2023). Tax Audits as Scarecrows: Evidence from a Large-Scale Field Experiment. *American Economic Journal: Economic Policy*, 15(1), 110-153.

Besley, T., & Persson, T. (2013). Taxation and Development. In A.J. Auerbach, R. Chetty, M. Feldstein, & E. Saez (Eds.), *Handbook of public economics* (Vol. 5, pp. 51-110). Elsevier.

Besley, T., & Persson, T. (2014). Why Do Developing Countries Tax So Little?. *Journal of Economic Perspectives*, 28(4), 99-120.

Best, M. C., Brockmeyer, A., Kleven, H. J., Spinnewijn, J., & Waseem, M. (2015). Production versus revenue efficiency with limited tax capacity: theory and evidence from Pakistan. *Journal of Political Economy*, 123(6), 1311-1355.

Bird, R. M., Martinez-Vazquez, J., & Torgler, B. (2008). Tax effort in developing countries and high income countries: The impact of corruption, voice and accountability. *Economic Analysis and Policy*, 38(1), 55-71.

Brueckner, J. K. (2000). Fiscal decentralization in developing countries: The effects of local corruption and tax evasion. *Annals of Economics and Finance*, 1(1), 1-18.

Carrillo, P., Pomeranz, D., & Singhal, M. (2017). Dodging the taxman: Firm misreporting and limits to tax enforcement. *American Economic Journal: Applied Economics*, 9(2), 144-164.

Charron, N., Dijkstra, L., & Lapuente, V. (2014). Regional governance matters: Quality of government within European Union member states. *Regional studies*, 48(1), 68-90.

Charron, N., Dijkstra, L., & Lapuente, V. (2015). Mapping the regional divide in Europe: A measure for assessing quality of government in 206 European regions. *Social indicators research*, 122, 315-346.

Charron, N., Lapuente, V., & Annoni, P. (2019). Measuring quality of government in EU regions across space and time. *Papers in Regional Science*, 98(5), 1925-1953.

Charron, N., Lapuente, V., Bauhr, M., & Annoni, P. (2022). Change and Continuity in Quality of Government: Trends in subnational quality of government in EU member states. *Investigaciones Regionales-Journal of Regional Research*, 2022(53), 5-23.

DeBacker, J., Heim, B. T., & Tran, A. (2015). Importing corruption culture from overseas: Evidence from corporate tax evasion in the United States. *Journal of Financial Economics*, 117(1), 122-138.

DeBacker, J., Heim, B. T., Tran, A., & Yuskavage, A. (2018). Once bitten, twice shy? The lasting impact of enforcement on tax compliance. *The Journal of Law and Economics*, 61(1), 1-35.

Devereux, M. P., & Griffith, R. (1998). The taxation of discrete investment choices. IFS working papers No. W98/16.

Doerr, A., & Necker, S. (2021). Collaborative tax evasion in the provision of services to consumers: A field experiment. *American Economic Journal: Economic Policy*, 13(4), 185-216.

Egger, P., & Loretz, S. (2010). Homogeneous profit tax effects for heterogeneous firms? *The World Economy*, 33(8), 1023-1041.

Egger, P., Loretz, S., Pfaffermayr, M., & Winner, H. (2009). Firm-specific forward-looking effective tax rates. *International Tax and Public Finance*, 16(6), 850-870.

EVS/WVS (2022). European Values Study and World Values Survey: Joint EVS/WVS 2017-2022 Dataset (Joint EVS/WVS). GESIS Data Archive, Cologne. ZA7505. Dataset Version 3.0.0, doi:10.4232/1.13899.

Fabling, R., Gemmell, N., Kneller, R., & Sanderson, L. (2014). Estimating firm-level effective marginal tax rates and the user cost of capital in New Zealand. Motu working paper 13-14. Motu Economic and Public Policy Research.

Flatters, F., & MacLeod, W. B. (1995). Administrative corruption and taxation. *International Tax and Public Finance*, 2, 397-417.

Graham, J. R., Raedy, J. S., & Shackelford, D. A. (2012). Research in accounting for income taxes. *Journal of Accounting and Economics*, 53(1-2), 412-434.

Ghura, M. D. (1998). Tax revenue in Sub-Saharan Africa: Effects of economic policies and corruption. IMF Working Paper WP/98/135.

Hall, R. E., & Jorgenson, D. W. (1967). Tax policy and investment behavior. *American Economic Review*, 57(3), 391-414.

Janský, P. (2023). Corporate effective tax rates for research and policy. *Public Finance Review*, 51(2), 171-205.

Kaldor, N. (1963). Taxation for economic development. *The Journal of Modern African Studies*, 1(1), 7-23.

Kaufmann, D., Kraay, A., & Mastruzzi, M. (2011). The worldwide governance indicators: Methodology and analytical issues. *Hague Journal on the Rule of Law*, 3(2), 220-246.

King, M. A. (1974). Taxation and the cost of capital. *Review of Economic Studies*, 41(1), 21-35.

King, M. A., & Fullerton, D. (1984). Introduction to "The taxation of income from capital: A comparative study of the United States, the United Kingdom, Sweden, and Germany". In M.A. King, D. Fullerton (Eds.), *The taxation of income from capital: A comparative study of the United States, the United Kingdom, Sweden, and Germany* (pp. 1-6). University of Chicago Press.

Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., & Saez, E. (2011). Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica*, 79(3), 651-692.

Lediga, C., Riedel, N., & Strohmaier, K. (2020). Tax Enforcement Spillovers – Evidence from Business Audits in South Africa. Available at SSRN 3616817.

Litina, A., & Palivos, T. (2016). Corruption, tax evasion and social values. *Journal of Economic Behavior & Organization*, 124, 164-177.

Liu, L. (2020). Where does multinational investment go with territorial taxation? Evidence from the United Kingdom. *American Economic Journal: Economic Policy*, 12(1), 325-358.

Mc Auliffe, S., Thunecke, G. U., & Wamser, G. (2022). The Tax-Elasticity of Tangible Fixed Assets: Evidence from Novel Corporate Tax Data. RSIT Working Paper 11/2022.

OECD (1991). *Taxing profits in a global economy: domestic and international issues.* OECD Publication and Information Center.

Pomeranz, D. (2015). No taxation without information: Deterrence and self-enforcement in the value added tax. *American Economic Review*, 105(8), 2539-2569.

Rechbauer, M., & Rünger, S. (2023). Measuring the impact of tax loss carryforwards on capital structure choice: the accuracy of earnings-based proxies. *Review of Managerial Science*, 1-28.

Richardson, G. (2006). Determinants of tax evasion: A cross-country investigation. *Journal of International Accounting, Auditing and Taxation*, 15(2), 150-169.

Steinmüller, E., Thunecke, G. U., & Wamser, G. (2019). Corporate income taxes around the world: A survey on forward-looking tax measures and two applications. *International Tax and Public Finance*, 26(2), 418-456.

Tennekes, M. (2018). tmap: Thematic Maps in R. *Journal of Statistical Software*, 84(6), 1-39.

Torgler, B. (2007). *Tax compliance and tax morale: A theoretical and empirical analysis.* Edward Elgar Publishing.

Torgler, B., Demir, I. C., Macintyre, A., & Schaffner, M. (2008). Causes and consequences of tax morale: An empirical investigation. *Economic Analysis and Policy*, 38(2), 313-339.

Uslaner, E. M. (2010). Tax evasion, corruption, and the social contract in transition. In J. Alm, J. Martinez-Vazquez, & B. Torgler (Eds.), *Developing alternative frameworks for explaining tax compliance* (pp. 174–190). Routledge International Studies in Money and Banking

Wamser, G., Merlo, V., Ruf, M., Stähler, F., Strohmaier, K., Eklund, J., Hansen, J., Laudage, S., Mc Auliffe, S., Paraknewitz, J., & Thunecke, G. U. (2023). A new database on international tax institutions. Forthcoming in RSIT Working Papers.

Xu, H., Dao, M., & Petkevich, A. (2019). Political corruption and auditor behavior: Evidence from US firms. *European Accounting Review*, 28(3), 513-540.

3. A Welfare Analysis of the Negative Income Tax with Nonlinear Labor Supply Estimation

Abstract–This paper conducts a social welfare analysis of the negative income tax (NIT), explicitly allowing for nonlinear labor supply responses with respect to the take-back rate, i.e., the rate at which NIT transfer payments are reduced per additional dollar of income. We derive a theoretical model which yields a notion of social welfare as function of the take-back rate that we calibrate using data from two NIT experiments that were conducted in the US in the 1970s. We find (i) that both the theoretical model and the empirical estimates suggest that the labor supply with respect to the take-back rate is nonlinear; (ii) that the social welfare optimizing take-back rates strongly differ between models calibrated with nonlinear versus linear labor supply functions; and (iii) that the welfare optimizing take-back rate lies between 65% and 69% for most tested parameterizations. However, due to poor data quality, the validity of the empirical findings is limited.

3.1. Introduction

Standard social welfare analysis often uses a single parameter, typically an elasticity, to quantify key behavioral responses to policy changes. This "sufficient statistics approach" (Chetty, 2009), however, may produce results that strongly differ from approaches that explicitly allow for variation in elasticities across different policy levels. A striking example is provided by Kasy (2018), who analyzes optimal coinsurance rates for health insurance using data from the RAND health insurance experiment. Allowing for arbitrary variation in health care expenditure elasticities across different coinsurance rates, he finds a welfare maximizing coinsurance rate of 18%, whereas the sufficient statistics approach yields an optimal rate of 50%.

In this paper, we use a similar approach to that of Kasy (2018) to analyze the negative income tax (NIT), a transfer scheme that aims at reducing poverty. The NIT provides families without any income with a transfer equal to the guaranteed income level G. The NIT transfer, however, linearly declines in family income at the take-back rate t up until a break-even point where the transfer becomes zero (see, e.g., Saez, 2002). In our analysis, we focus on the role of t to answer the research question: which level of t is social welfare maximizing? To this end, we first theoretically derive a notion of social welfare that takes into account the families' allocation decision regarding their disposable time which can either be used for work (which earns labor income that is used for consumption, but also reduces the NIT transfer) or leisure. Our notion of social welfare describes the trade-off between the policy maker's two objectives that both depend on the magnitude of the take-back rate t: (i) maximizing private utility, which is achieved by reducing t, i.e., all other things equal, increasing the NIT transfer, and (ii) minimizing government spending for the NIT, which is achieved by increasing t, assuming all other things equal. The key behavioral relationship in our model is the one between t and the labor supply, as a marginal increase in t induces a decrease in labor supply, which in turn diminishes the mechanical savings effect regarding the transfers. Unlike previous theoretical analyses of the NIT (see, e.g., Saez, 2002), we do not rely on the aforementioned sufficient statistics approach that summarizes key behavioral relationships using a single elasticity parameter. Instead, we derive simple expressions of the labor supply and social welfare that are functions of t. In a next step, we empirically estimate the labor supply function using data from two NIT experiments that were conducted in the US in the 1970s. In these experiments, treated families were assigned to different NIT plans, i.e., different combinations of G and t. Controlling for G and a number of family-specific controls, we find that families assigned to higher take-back rates supply less labor in terms of hours worked, as predicted by our theoretical model. Furthermore, also in line with the theory model, the empirical labor supply function is concave for most of the observed range of t. However, considering the full available range of t, this result does not hold. Finally, we plug the estimated labor supply function into our notion of social welfare. For most of the tested parameterizations, our findings suggest that the welfare optimizing take-back rate is quite large, lying between approximately 65% and 69%. However, due to poor data quality of the NIT experiment data, the validity of these empirical findings is limited. Furthermore, we find that the social welfare optimizing take-back rates strongly differ depending on whether we allow for nonlinearities in the labor supply estimation or not.

Besides the aforementioned contribution to the theoretical literature regarding the welfare analysis of the NIT, we add to two strands of the empirical literature. First, we contribute to the large body of studies estimating labor supply responses to government-run transfer programs. Regarding the NIT experiments, a comprehensive list of studies is provided by Widerquist (2005), who summarizes previous labor supply response estimates as being varying in size.[1] Most of the previous estimations using the NIT experiment data regress labor supply on changes in the net wage rate that are induced by the introduction of the

[1]The lack of an agreed acceptable level of work-disincentive also had effects on the policy debate surrounding the NIT in the US (Widerquist, 2005). In the end, the NIT was not introduced. Instead, the US opted for the Earned Income Tax Credit regime, which, up to a certain threshold, matches each dollar of earned income with a certain transfer, but pays nothing in the case of zero income (Saez, 2002). As pointed out by Saez (2002), however, many transfer programs in European countries work like the NIT.

NIT, often distinguishing different subgroups, such as husbands, wives, or single female heads (Robins, 1985). In contrast, our estimation approach of the labor supply response focuses on the direct effects of different levels of the take-back rate on aggregated family labor supply. The evaluation of labor supply responses also plays a large role in the context of randomized control trials (RCTs) that provide targeted transfers. A recent example is Verho et al. (2022) who study the work-disincentive effect of participants in an RCT conducted in Finland that replaced minimum unemployment benefits with an unconditional income of the same size. They find that the days in employment did not statistically change during the first year. Banerjee et al. (2017) provide a comprehensive re-evaluation of data from seven RCTs of cash transfers in six developing countries. They do not find systematic evidence of a work-discouragement effect. Second, we add to the growing literature in public finance that estimates behavioral effects for different levels of policy variables rather than relying on a single aggregating estimate. Besides the aforementioned study by Kasy (2018), a recent example can be found in Fuest et al. (2022) who study the profit shifting behavior of multinational firms. They show that profit elasticities depend nonlinearly on the magnitude of countries' tax rates and argue that taking these nonlinearities into account is key for obtaining accurate profit shifting estimates.

The remainder of the paper is structured as follows. Section 3.2 theoretically derives our notion of social welfare. Section 3.3 describes the estimation strategy as well as the data. The results of our empirical labor supply and social welfare estimations are presented in Section 3.4. Finally, Section 3.5 concludes.

3.2. Theory

3.2.1. Household Optimization Problem

We start out by describing the decision problem of the family that is subject to the NIT. The theory builds on standard intensive labor supply models in the context of income taxation as discussed in, e.g., Hausman (1985) or Keuschnigg and Wamser (2024).[2] The key decision that the family has to make is the one concerning the allocation of the available time $T > 0$ between hours worked $L \in (0, T)$, that are compensated at the wage rate $w > 0$,[3] and hours used for leisure F. Hence, leisure is given by $F = T - L$. The opportunity cost of leisure is the labor income that could have been earned instead. The total income of the family states as follows:

$$Y = wL + I + S + P = wL + I + S + \max\{G - t(wL + I) - S,\ 0\}. \tag{3.1}$$

The labor income is given by wL. $I \geq 0$ denotes unearned income, such as interest, dividends, or capital gains. $S \geq 0$ gives the total public assistance that the family receives, including, e.g., Aid to Families with Dependent Children (AFDC). We treat both I and S as exogeneously given. Finally, $P = \max\{G - t(wL + I) - S,\ 0\}$ denotes the NIT payment, with $G > 0$ denoting the guaranteed income level and $t \in (0, 1)$ denoting the take-back rate. Note that while labor income and unearned income are taxed at t, the welfare income S is taxed at 100%, which means that the NIT payment effectively replaced the multitude of different welfare programs during the US experiments (Mathematica Policy Research, Inc. [MPR], 1980). For the remainder of this section, we assume that the family receives some strictly positive NIT payment P in the optimum, i.e., $G > t(wL + I) + S$.[4] Furthermore, we set unearned income to zero, i.e., $I = 0$. This assumption is plausible in the context of the US NIT experiments of the 1970s that we use for our empirical analysis below, as the samples consist of poor families with no or negligibly small unearned income.[5] Finally, we introduce the "keep-rate" k, which we define as $1 - t$. The use of the keep-rate rather than the take-back rate is solely due to practical reasons regarding the formulation of the social

[2]Note that the model is kept simple, as the goal is to obtain a notion of social welfare function that can be directly estimated and that allows for nonlinearities in the labor supply. For a more rigorous theoretical analysis of the NIT, see Aboudi et al. (2014) or Saez (2002).

[3]We assume that families are price takers and take the wage rate as given.

[4]We further assume that also after marginal changes in exogenous model parameters, in particular changes in t, the family still receives some strictly positive NIT transfer.

[5]See descriptive statistics below in Section 3.3.2 or Widerquist (2005).

welfare function. The budget constraint in (3.1) then simplifies to

$$Y = G + kwL. \tag{3.2}$$

It is important to note that since our model is static and therefore does not allow for savings, consumption equals total income Y. Given that the NIT is a policy instrument that is designed to target low-income households and given that the transfers from the NIT experiment typically do not raise families' disposable incomes much above the poverty line (see, e.g., discussion of the NIT experiments that were conducted in the US in the 1970s below), this assumption seems plausible. Finally, we assume that the family's preferences for consumption Y and leisure F are captured by the Cobb–Douglas utility function

$$U(Y, F) = U(Y, T - L) = Y^{\alpha}(T - L)^{1-\alpha}. \tag{3.3}$$

α and $1 - \alpha$ are the utility elasticities of consumption and leisure, respectively, which we take as given, with $\alpha \in (0, 1)$.[6] We proceed to formulate the utility maximization problem of the household (with λ denoting the Lagrange multiplier and $V(\cdot)$ denoting the indirect utility function):

$$V(k, w, \alpha, T, G) = \max_{Y,L} \left[Y^{\alpha}(T - L)^{1-\alpha} + \lambda(G + kwL - Y)\right]. \tag{3.4}$$

Solving the first-order conditions corresponding to (3.4), we obtain the Marshallian labor supply:

$$L^{*} = L(k, w, \alpha, T, G) = \alpha T - \frac{G(1-\alpha)}{kw}. \tag{3.5}$$

[6]Note that $\alpha \in (0, 1)$ implies homogeneity of degree one, i.e., multiplying both Y and F by the same factor $a > 0$ leads to an increase in utility by the same factor. Formally: $U(aY, aF) = aU(Y, F)$ (see, e.g., Mas-Colell et al., 1995).

For our analysis, we are particularly interested in the response of L^* with respect to marginal changes in the keep-rate k. The Marshallian labor supply in (3.5) implies that L^* is strictly increasing in k, as

$$\frac{\partial L^*}{\partial k} = \frac{G(1-\alpha)}{k^2 w} > 0. \tag{3.6}$$

Note that this result implies that the substitution effect is larger than the income effect. The substitution effect states that less leisure is consumed as a result of an increase in the after-tax wage rate kw, i.e., the opportunity cost of leisure. Consequently, the labor supply $L = T - F$ increases. The income effect, on the other hand, suggests that the labor supply decreases as kw increases, as the household can maintain its original consumption level Y with a labor supply level that is lower than the initial one.[7]

A key result of our theoretical model is that the labor supply response is *nonlinear*. More precisely, the second derivative of (3.5) with respect to the keep-rate k implies concavity:

$$\frac{\partial^2 L^*}{\partial k^2} = \frac{-2G(1-\alpha)}{k^3 w} < 0. \tag{3.7}$$

This means that the labor supply response induced by a marginal increase in k is smaller when k is comparatively high already. Consequently, this finding suggests that characterizing the complete labor supply function with a single parameter – as often done in simple sufficient statistics welfare formulas – is not feasible in our setup.

The family's consumption in the optimum is given by:

$$Y^* = Y(k, w, \alpha, T, G) = G + kwL^* = \alpha(kwT + G). \tag{3.8}$$

It can easily be seen that the first derivative of Y^* with respect to k is also strictly positive: $\partial Y^*/\partial k = \alpha wT > 0$. This is due to (i) the fact that the family increases its labor supply

[7]For an in-depth discussion of the substitution and income effects, see Keuschnigg and Wamser (2024).

in response to an increase in k (see (3.6)) and thereby increases its labor income; and (ii) a mechanical increase in the NIT transfer due to an increase in k.

Finally, the indirect utility function is obtained by plugging (3.5) and (3.8) into (3.4):[8]

$$\begin{aligned} V^* = V(k, w, \alpha, T, G) &= (G + kwL^*)^\alpha (T - L^*)^{1-\alpha} \\ &= [\alpha(kwT + G)]^\alpha \left[(1-\alpha)\left(T + \frac{G}{kw}\right)\right]^{(1-\alpha)}. \end{aligned} \tag{3.9}$$

The derivation of V^* with respect to k states as follows:

$$\frac{\partial V^*}{\partial k} = \frac{(1-\alpha)}{k} Y^{*\alpha}(T - L^*)^{-\alpha} L^*. \tag{3.10}$$

Given the assumptions regarding ranges of the parameters and choice variables, it follows that $\partial V^*/\partial k > 0$. A proof of this result as well as a detailed derivation of (3.10) are provided in Appendix 3.6.1. Invoking the Envelope Theorem, it can be shown that changes in the choice variables (i.e., L, Y, and λ) as response to a marginal change in k have no effect on V^* in the optimum. Instead, the derivative of V^* with respect to k equals its direct derivative (see, e.g., Mas-Colell et al., 1995 or Keuschnigg and Wamser, 2024). A brief demonstration of the Envelope Theorem is provided in Appendix 3.6.2.

3.2.2. NIT Payment

We now turn to the government side. Using the family's optimal labor supply choice L^* from above, we can compute the NIT payment that the government makes to the family:[9]

$$P^* = G - (1-k)wL^* - S. \tag{3.11}$$

[8]Note that the term $\lambda^*(G + kwL^* - Y^*)$ in (3.4) becomes zero, as the budget constraint is satisfied with equality in the optimum.

[9]Keep in mind that we assume that $P > 0$ always holds.

A marginal change in k has two effects on the magnitude of P^*. (i) A mechanical effect that can simply be computed by holding the family's labor supply fixed at L^*. It amounts to wL^*. Note that this effect is always strictly positive, i.e., it increases the NIT payment, as both $w > 0$ and $L^* > 0$. (ii) A behavioral effect, which results from the family adjusting its labor supply in response to the change in k. In detail, this change amounts to $-(1-k)w(\partial L^*/\partial k)$. As $k \in (0, 1)$, $w > 0$, and $\partial L^*/\partial k > 0$ (see (3.6) above), this behavioral effect is strictly negative, i.e., reduces the NIT payment. Adding up the mechanical and the behavioral effects, we obtain the partial derivative of (3.11) with respect to k:

$$\frac{\partial P^*}{\partial k} = wL^* - (1-k)w\frac{\partial L^*}{\partial k} = w\left(\alpha T - \frac{G(1-\alpha)}{k^2 w}\right). \tag{3.12}$$

The last equality is obtained by inserting the value function for L^* (see (3.5)) as well as its derivative with respect to k (see (3.6)). The labor supply response of the household at k plays an important role for the magnitude of (3.12), with a strong response, i.e., a steeply upward sloping labor supply curve, being beneficial for the government. Note that our model does not suggest any particular sign for (3.12); in theory, payments could decrease as result of an increase in k. This is the case when the behavioral effect outweighs the mechanical effect. In our empirical application below, however, this special case does not play a role.

3.2.3. Social Welfare Function

Finally, we define social welfare as a function of k that the policy maker seeks to maximize. Similar to Kasy (2018), we define social welfare as the difference between household utility (see (3.9)) and the transfer payment (see (3.11)), both of which are value functions of the take-back rate k and all other exogenous parameters. For the sake of notational simplicity we shall henceforth denote the value function of the labor supply depicted in (3.5) with

$L^* = L(k)$. Formally, the social welfare function states as follows:

$$\begin{aligned} SW(k) &= V^* - P^* \\ &= (G + kwL(k))^{\alpha}(T - L(k))^{1-\alpha} - [G - (1-k)wL(k) - S]. \end{aligned} \tag{3.13}$$

Note that while both our notion of social welfare as well as the one proposed by Kasy (2018) account for nonlinear behavioral responses, there are two key aspects in which they differ. First, we explicitly model the household's trade-off between consumption and leisure under the assumption of Cobb-Douglas preferences. In Kasy (2018), the individual chooses a level of health care expenditure while being confronted with a given coinsurance rate. However, the trade-off the individual faces in this setup, i.e., staying/becoming healthy versus reducing out-of-pocket costs and thereby increasing, e.g., other consumption, is not theoretically modeled. Instead, the only assumption that is made by Kasy (2018) is that maximized private utility changes linearly with respect to the coinsurance rate. In our setup, the indirect utility function is a nonlinear function of the policy parameter of interest k, see (3.10).[10] Second, the social welfare function used by Kasy (2018) assumes that the policy maker sets a marginal value of an additional dollar transferred to the sick relative to the cost of an additional dollar of expenditure for the health insurance provider, which is assumed to be larger than one. Without this parameter, the social welfare function in Kasy (2018) essentially collapses. In contrast, our setup does not necessitate invoking such a parameter.

As mentioned above, the policy maker sets the keep-rate such that social welfare is maximized. We denote the maximizing level of the keep-rate as k^*. The first-order condition is

$$SW'(k^*) = \left.\frac{\partial V^*}{\partial k}\right|_{k=k^*} - \left.\frac{\partial P^*}{\partial k}\right|_{k=k^*} = 0. \tag{3.14}$$

[10]Note that if one would assume quasilinear preferences for our problem with consumption Y entering utility linearly, the indirect utility function would also change linearly with respect to k. However, we believe that assuming Cobb-Douglas preferences with diminishing marginal utility with respect to both consumption and leisure is more adequate.

Using the expressions for the first derivatives of V^* and P^* with respect to k from above (see (3.10) and (3.12), respectively), we can rewrite (3.14) as a function of the Marshallian labor supply $L(k)$ and its first derivative with respect to k, $L'(k)$:

$$SW'(k^*) = \frac{(1-\alpha)}{k^*} \left(G + k^* w L(k^*)\right)^{\alpha} \left(T - L(k^*)\right)^{-\alpha} L(k^*) - wL(k^*) + (1-k^*)wL'(k^*) = 0. \tag{3.15}$$

Note that the first derivative of the social welfare function is conceptually similar to the notion of excess burden as described in, e.g., Keuschnigg and Wamser (2024) in the context of income taxation. This becomes apparent when we think about a marginal decrease in k, which is identical to an increase in the take-back rate t:[11] An increase in t reduces utility (see (3.10)), which reduces overall social welfare (see (3.13)). However, the increase in t mechanically lowers the NIT payment, which is beneficial for overall welfare, as the government has more money at its disposal for other purposes that increase welfare. In this sense, from a social welfare perspective, a lower NIT payment is similar to an increase in income tax revenue. The degree to which the lower NIT payments make up for the decrease in private utility hinges on the behavioral response of the household, or, more precisely, the magnitude of the substitution effect at t. In case the household is strongly decreasing its labor supply in response to a marginal increase in t, i.e., a reduction in the after-tax wage rate $(1-t)w$, the mechanical decrease in the NIT payment is largely canceled out. Our notion of the excess burden states how much of the loss in utility cannot be offset by the decrease in the NIT payment. They key innovation compared to other standard formulas of social welfare and excess burden is that we do not rely on a single elasticity parameter characterizing the curvature of the labor supply function. Instead, we account for the response intensity at each level of k. In the remainder of the paper, we use data from the US NIT experiments to estimate $L(k)$, explicitly allowing for nonlinearities. Then, we use the estimate of $L(k)$ to estimate (3.13) for different levels of k to determine the welfare optimizing keep-rate k^*.

[11] Remember that $k = 1 - t$.

3.3. Empirical Approach

3.3.1. Estimation Strategy

The key ingredient for the estimation of the social welfare function for different levels of the take-back rate k is the labor supply $L(k)$, see (3.13). With all exogenous model parameters at hand, one could simply calculate $L(k)$ using (3.5). However, we are interested in an empirical estimate of the labor supply, $\hat{L}(k)$, that accounts for behavioral responses that cannot be captured by our simple model. In doing so, we accept potential contradictions between $\hat{L}(k)$ the purely model-based $L(k)$. In the following, we describe the estimation strategy for obtaining $\hat{L}(k)$ using data from two NIT experiments that were conducted in the US in the 1970s. In these experiments, families were assigned to different NIT plans consisting of combinations of take-back rates ($k = 0.30$, $k = 0.40$, $k = 0.50$, or $k = 0.60$) and guaranteed income levels G, which amounted to either 75%, 95%, 100%, 120%, or 140% of the respective family's poverty line (see MPR, 1980; Robins, 1985; Widerquist, 2005). In addition to the families in the NIT plans, the experiments observed a number of families in the control group with $k = G = 0$. We predict the marginal effects of the different take-back rates on the labor supply of the families using the following linear estimation equation:[12]

$$\begin{aligned} \textit{Labor supply}_{iq} &= \beta_{0.3}\mathbb{1}(k_{iq} = 0.3) + \beta_{0.4}\mathbb{1}(k_{iq} = 0.4) + \beta_{0.5}\mathbb{1}(k_{iq} = 0.5) + \beta_{0.6}\mathbb{1}(k_{iq} = 0.6) \\ &\quad + \kappa G_{iq} + \gamma \textit{No treatment}_{iq} + \boldsymbol{\phi X_i} + \boldsymbol{\psi X_{iq}} + \boldsymbol{\zeta X_{iq-1}} + \theta_y + \varepsilon_{iq}. \end{aligned} \tag{3.16}$$

[12]Note that our labor supply estimation differs from most previous studies using the same NIT experiment data in that these studies typically conduct the estimation at the individual level, often times focusing on and comparing the labor supply responses of different groups such as husbands, wives, or single female heads of households. See Robins (1985) or Widerquist (2005) for overviews of such studies. Note that in our context, estimating the labor supply at the individual level would complicate the analysis, as, e.g., the labor supply of the husband is a function of his wife's labor supply and vice versa. Controlling for the respective spouse's labor supply, however, would lead to endogeneity issues. Furthermore, since the NIT payments are administered at the family level, either way some sort of (non-trivial) aggregation from the individual response to the family level response would still be necessary in such a setup.

The indices i and q denote family and experimental quarter, respectively.[13] The dependent variable $\textit{Labor supply}_{iq}$ gives the sum of hours worked by all members of family i on all regular jobs in the given quarter q. Following Kasy (2018), we estimate the marginal effects of different take-back rates using dummy variables, denoted by $\mathbb{1}(k_{iq} = k)$. The corresponding OLS coefficients are denoted by the β_k's. The marginal effect of the guaranteed income level G_{iq} is given by κ. Note that for the families in the control group, G_{iq} is equal to zero in all quarters. $\textit{No treatment}_{iq}$ is an indicator variable that is equal to unity if a family was in the control group, i.e., was never eligible for any NIT payment. To ensure that the estimation of the β_k's and κ is not contaminated by families that are treated with an NIT plan, however, do not receive NIT payments as their income is too high,[14] we assign families that did not receive any NIT payments in the previous quarter $q - 1$ to the control group in q. The underlying rationale of this assignment of certain observations to the control group is that the magnitudes of the keep-rate and the guaranteed income level are only relevant when a strictly positive NIT payment is expected. Not making this adjustment would lead to a systematic bias in the estimation of the β_k's and κ, as one would expect more generous plans (i.e., plans with high k and high G) to have a higher share of families receiving payments than less generous ones (i.e., plans with low k and low G). We further control for a set of time-constant variables, contained in $\boldsymbol{X_i}$, as well as sets of contemporary and lagged variables, contained in $\boldsymbol{X_{iq}}$ and $\boldsymbol{X_{iq-1}}$, respectively. The corresponding coefficients for the variables in $\boldsymbol{X_i}$, $\boldsymbol{X_{iq}}$, and $\boldsymbol{X_{iq-1}}$ are collected in the vectors $\boldsymbol{\phi}$, $\boldsymbol{\psi}$, and $\boldsymbol{\zeta}$, respectively. These additional variables control for factors that potentially influence the labor supply other than the NIT variables and include, e.g., the pre-experimental average wage rate across the different working family members, the pre-experimental quarterly total labor income, the number of adults, the number of minors, the gender of the family head, or the received

[13] The experimental quarters denote the months since the enrollment of the family, with the first quarter comprising the first, second, and third month after the enrollment month (Mathematica Policy Research, Inc. [MPR] and Social & Scientific Systems, Inc. [SSS], 1980). Therefore, the experimental quarters generally do not align with the calendar quarters, i.e., January to March, April to June, etc.

[14] For details on the calculation of the NIT payments, see Section 3.2.1.

welfare income. In particular controlling for pre-experimental income as well as the family size and composition is crucial, as the assignment of the different NIT plans was not random but instead was based on the Conlisk-Watts assignment model (Conlisk and Watts, 1979). This method aims to reduce the expected costs of the experiment, i.e., the sum of NIT payments, and therefore favors families with high pre-tax income and small families in the assignment of generous plans, i.e., plans with high k and high G.[15] As pointed out by Keeley and Robins (1978), controlling for the Conlisk-Watts assignment variables is the only way to correct for the bias from the non-random assignment, even though the inclusion of additional control variables may reduce the reliability of the estimates (Keeley, 1981). In addition to the family-specific control variables we control for year fixed effects, which are denoted by θ_y, with y indicating the calendar year. Note that since each unique NIT plan was tested only at one of the two sites,[16] we pool over the two locations, i.e., do not include site fixed effects. Further note that since a family's NIT plan did not change over the course of the experiment, we do not include family fixed effects. However, we cluster our standard errors at the family level to account for non-independence between the different quarters. Finally, ε_{iq} denotes the error component. Similar to Kasy (2018), we de-mean all covariates except for the $\mathbb{1}(k_{iq} = k)$'s and the $No\ treatment_{iq}$ dummy. Since our model does not include a constant, this allows for the interpretation of the β_k's in (3.16) as the average labor supply corresponding to the given coinsurance rate for a hypothetical family that has control variables that are equal to the respective sample means.

Regarding the labor supply estimation, there is an important aspect that deserves some special attention, namely the distinction between intensive and extensive margin responses (Heckman, 1993). Unlike Saez (2002), our theoretical model – for the sake of simplicity – focuses exclusively on intensive margin responses and rules our non-participation in the labor market by assumption. Given that we consider the aggregated labor supply of a family rather

[15] Note that family size is relevant for the magnitude of NIT payments, as G is calculated by multiplying a constant factor with the family-specific poverty line, with the latter being higher for large families. For more details on the calculation of the poverty line, see Section 3.3.2.

[16] More details on this are provided below in Section 3.3.2.

than of an individual and given that our observations correspond to rather long three month time periods, this assumption does not seem completely implausible. In fact, only 16.78% of family-quarter observations in the final sample report zero hours worked. Nevertheless, we want to clarify that the empirical estimates of the labor supply are the result of a combination of responses along both margins.

In a second step, to be able to compute and evaluate the social welfare function not only at the observed k's but at all k's in the range $[0.30, 0.60]$, we apply a cubic spline monotonic interpolation approach (Dougherty et al., 1989; Forsythe et al., 1977; Hyman, 1983). Unlike conventional cubic splines, this method ensures that the slope of the labor supply curve is monotonically increasing. Of course, this requires that the β_k estimates that are passed to the spline are monotonically increasing as well, i.e., $\beta_{0.3} \leq \beta_{0.4} \leq \beta_{0.5} \leq \beta_{0.6}$. As can be seen in Section 3.4.1, this is the case in our analysis. Another feature of the spline is that the interpolated line passes right through observed data points, i.e., the β_k's. The main advantage of cubic spline monotonic interpolation lies in its simplicity, in particular in comparison to machine learning based algorithms that are usually more computationally intensive, require tedious (and often times arbitrary) tuning of hyperparameters (see, e.g., Hastie et al., 2009), and usually do not work with categorical variables (Potdar et al., 2017). On the downside, compared to some machine learning algorithms such as the Gaussian Process Priors used by Kasy (2018) that allow for the computation of confidence bounds, the interpolated line obtained with cubic splines does not allow to make statements about statistical uncertainty.

In a third and final step, the estimated labor supply curve is plugged into the social welfare function (3.13). Other exogenous parameters that are needed for the computation of social welfare as function of k, i.e., G, T, w, or S, are calibrated using the data from the NIT experiments. The utility elasticities of consumption and leisure, α and $1-\alpha$, are varied in the empirical analysis to simulate different family preferences.

3.3.2. Data, Sample, and Parameterization

As mentioned above, the data used for the empirical analysis stem from the NIT experiments that were conducted in different areas of the United States between 1968 and 1980.[17] The implementation of the experiments was the result of political debate surrounding the NIT in the context of the "war on poverty" that President Lyndon Johnson's called for in his state of the union address in 1964. Advocates of the NIT, including most prominently the Office of Economic Opportunity, which was established by the US Congress to administer the war on poverty, were confronted with the criticism that NIT programs could promote idleness. While a reduction of the labor supply in response to the NIT is in line with economic theory (see Section 3.2), reliable empirical estimates regarding the magnitude of the work-disincentive effect were not available at the time. As a result, the NIT experiments were implemented to obtain the necessary information to settle the debate (Hum and Simpson, 1993).

For our analysis, we use the "Cross-Site Analysis File", which includes records for the New Jersey, the Gary (Indiana), and the Seattle/Denver Income Maintenance Experiments (SIME/DIME) and was provided by the Data and Information Services Center (DISC) at the University of Wisconsin-Madison.[18] The Cross-Site Analysis File provides relevant variables for the different experiments in a common format using the same concepts and definitions for the construction of the variables (MPR and SSS, 1980). In detail, it provides information on each individual enrolled in the experiment as head of a family for the 12 experimental quarters as well as for the four quarters preceding the start of the NIT treatment.[19] Note that since

[17]Note that additionally to the NIT experiments conducted in the US, there was also an experiment conducted in Manitoba, Canada between 1975 and 1978, the so-called "Manitoba Basic Annual Income Experiment" or "Mincome" (Widerquist, 2005). However, due to data availability, we only consider the US experiments.

[18]Note that the fourth US experiment, the Rural Income-Maintenance Experiment (RIME) which was conducted in Iowa and North Carolina from 1970 to 1972 (Widerquist, 2005) is not included in the Cross-Site Analysis File and therefore is not part of our analysis.

[19]Note that due to infrequency of interviews and attrition, not all families are covered in each quarter. Regarding the survey method of interviews, there is evidence of systematic misreporting of labor supply (Greenberg et al., 1981; Greenberg and Halsey, 1983). Since the periodic interviews are the only data available to us, we cannot correct for this bias, however, we want the reader to be aware that the results presented in this paper may possibly be biased due to misreporting.

we also control for pre-experimental variables, we exclude experimental observations where the number of family heads does not coincide with the one of the quarter before the start of the treatment, as this suggests that either a marriage or a divorce happened (MPR and SSS, 1980) and the pre-experimental variables therefore lose validity as controls.[20] We convert all monetary variables to 1971 dollars using consumer price based annual inflation rates that we obtain from the World Bank's *World Development Indicators* database. Since several variables that we need for our analysis are not provided for the New Jersey experiment, we only use data from the Gary and the Seattle/Denver experiments.[21]

The Gary experiment was carried out between 1971 and 1974 and initially comprised a total of 1,799 households. It included only black families, with the majority of them being single-headed (54%). The requirements for participation were that the head of the family was between 18 and 58 years old and that the family income was below 240% of the poverty line in the year of enrollment. The NIT plans that were tested in Gary were combinations of the guaranteed income levels G of either 75% or 100% and keep-rates k of 0.40 or 0.60 (Robins, 1985; Widerquist, 2005). Note that the assignment of treatment in all experiments was based on information obtained from a couple of interviews that were conducted before enrollment (MPR, 1980). Compared to the Gary experiment, the Seattle/Denver experiments, for which we observe the time span 1971 to 1975,[22] exhibited a much larger sample size of initially 4,800 households. The households in the Seattle/Denver experiments were primarily black (43%), followed by an almost equally big share of white households (39%), and a comparatively small share of Latino households (18%). The share of single-headed families amounted to

[20]Interestingly, the families in the NIT experiments exhibited a substantial number of divorces, which in itself is a research subject. Widerquist (2005) provides an extensive list of papers regarding this topic.

[21]The variables that are missing for New Jersey include, e.g., the quarterly experimental payments or family social security income (MPR and SSS, 1980).

[22]Note that for 71% of the households, the treatment was planned – and communicated – to last three years, which is also the treatment duration of the Gary experiment. For the other households in of the sample, the treatment duration was longer, either five years (25%) or 20 years (4%). However, the experiments were cancelled in 1980, such that the maximum treatment duration only amounted to 9 years. For our analysis, we only use the first three years of treatment, irrespective of the communicated total duration. We are aware that the labor supply response may differ depending on the communicated duration of treatment, however, due to a lack of variation in treatment duration across sites and therefore NIT plans, we do not control for this in our analysis.

39%. For eligibility, the households' income had to be below 325% of the respective poverty line in the year of enrollment. The plans that were tested in the Denver/Seattle experiments were combinations of the guaranteed income levels G of either 95%, 120%, or 140% and keep-rates k of 0.30 or 0.50 (Robins, 1985; Widerquist, 2005).[23]

The variables from the Cross-Site Analysis File that we use for our labor supply estimation are the following. The dependent variable, labor supply in hours in the given quarter ($Labor\ supply_{iq}$), is calculated as the sum of a family's heads' hours worked on all regular jobs. In an alternative specification of our model, we use the labor income ($Labor\ income_{iq}$) rather than the labor supply as dependent variable. Labor income is defined as the total of gross wages earned by a family's heads on all regular jobs in the respective quarter. As mentioned above, we code indicator variables for each of the four keep-rates k used in the experiment, which are denoted by $\mathbb{1}(k_{iq} = k)$. We also control for the guaranteed income level G_{iq} which states the NIT transfer in the absence of any family income. Note that since the Cross-Site Analysis File only provides G as the share of the respective family's poverty line, we use poverty thresholds tables for the year 1971 which we obtain from the US Census Bureau to calculate the actual thresholds using the relevant information on the families.[24] The indicator $No\ treatment_{iq}$ is equal to unity if a family is in the control group, in which case all $\mathbb{1}(k_{iq} = k)$'s as well as G_{iq} are equal to zero. As discussed above, besides the actual control group of the experiment, we also assign families that are assigned to an NIT plan but did not receive any transfers in the previous quarter to the control group of the current quarter. As final NIT related variable, we control for the NIT payments that the family received in the previous quarter ($NIT\ payment_{iq-1}$). Next, we control for family size, distinguishing the number of adults ($Persons\ 18\ or\ older_{iq}$) as well

[23]Note that additionally flexible keep-rates that depended on the household's income were tested in Seattle/Denver. However, we exclude households treated with such flexible rates, as our setup is using indicator variables to control for different levels of the keep-rate, which does not allow us to sensibly control for continuous variation in k (see Section 3.3.1). Furthermore, note that flexible keep-rates on the right-hand side of our labor supply estimation equation would be endogenous due to simultaneity, as they depend on income and are therefore correlated with the labor supply.

[24]Note that we converted all monetary variables to 1971 dollars before.

as minors ($Persons\ 17\ or\ younger_{iq}$). We further control for the gender of the head of the family ($Female\ head_{iq}$),[25] as well as the family type. Regarding the latter, we distinguish two-headed families ($Two\ heads_{iq}$), one-headed families with at least one dependent under 18 years old ($One\ head,\ dependents_{iq}$), and one-headed families without dependents ($One\ head,\ dependents_{iq}$). Since the NIT plans were not assigned randomly, but dependent on family size and pre-experimental income (Conlisk and Watts, 1979; Widerquist, 2005), we also control for the average quarterly family labor income across the four pre-experimental quarters ($Labor\ income\ pre-exp_i$). We furthermore control for a family's pre-experimental wage rate ($Wage\ rate\ pre-exp_i$), which we calculate as the weighted mean of all hourly wage rates of a family's heads, with the weights reflecting the share of total family hours worked by the given individual. Note that using wage figures during the treatment quarters would be endogenous, as family members of treated families might be reluctant to accept low paying jobs, effectively making the wage rate a treatment outcome itself.[26] The idea behind controlling for pre-experimental wage is to account for the families' value in the labor market. Finally, we control for families' income sources outside of labor income, namely the previous period's non-labor income ($Non-labor\ income_{iq-1}$), which includes interest and rent income (i.e., the I in the theoretical model in Section 3.2), total welfare income ($Welfare\ income_{iq-1}$), and social security income ($Social\ security\ income_{iq-1}$). The final sample includes a total of 30,307 family-quarter observations. Descriptive statistics on all variables are depicted in Table 3.1.

Finally, we estimate the other parameters needed for the welfare analysis. For the wage rate w, we simply use the sample mean of the pre-experimental wage rate ($Wage\ rate\ pre-exp_i$), depicted in Table 3.1. We also parameterize the guaranteed income level G in the theoretical model using the sample mean of G_{iq}, however, we exclude observations where $G_{iq} = 0$ from our calculations, yielding $G = 1142.76$ dollars. The total public assistance that

[25] Note that the head of the family is considered female if there is no male head.

[26] Of course, the wages in the pre-experimental quarters may be contaminated by anticipation effects, though we deem this a comparatively smaller issue.

Table 3.1: *DESCRIPTIVE STATISTICS*

The table depicts descriptive statistics on all the variables used for the empirical labor supply estimation. Definitions of the variables are provided in Section 3.3.2.

	Observations	Mean	(sd)
Panel A: Dependent variables			
$Labor\ supply_{iq}$	30,307	472.373	303.755
$Labor\ income_{iq}$	30,307	1500.315	1122.858
Panel B: Experimental treatment variables			
$\mathbb{1}(k_{iq} = 0.3)$	30,307	0.166	0.372
$\mathbb{1}(k_{iq} = 0.4)$	30,307	0.083	0.276
$\mathbb{1}(k_{iq} = 0.5)$	30,307	0.164	0.371
$\mathbb{1}(k_{iq} = 0.6)$	30,307	0.083	0.276
G_{iq}	30,307	567.255	631.516
$No\ treatment_{iq}$	30,307	0.504	0.500
$NIT\ payment_{iq-1}$	30,307	186.684	310.510
Panel C: Family control variables			
$Persons\ 18\ or\ older_{iq}$	30,307	2.112	0.895
$Persons\ 17\ or\ younger_{iq}$	30,307	2.221	1.562
$Female\ head_{iq}$	30,307	0.362	0.481
$Two\ heads_{iq}$	30,307	0.640	0.480
$One\ head,\ dependents_{iq}$	30,307	0.342	0.474
$One\ head,\ no\ dependents_{iq}$	30,307	0.018	0.134
$Labor\ income\ pre-exp_i$	30,307	1398.770	870.305
$Wage\ rate\ pre-exp_i$	30,307	2.892	1.128
$Non-labor\ income_{iq-1}$	30,307	6.036	42.882
$Welfare\ income_{iq-1}$	30,307	90.585	224.465
$Social\ security\ income_{iq-1}$	30,307	40.775	173.082

a family receives, denoted by S in the theoretical model, is parameterized by simply adding up the sample means of total welfare income ($Welfare\ income_{iq-1}$) and social security income ($Social\ security\ income_{iq-1}$) depicted in Table 3.1. The parameterization of T, i.e., the available time that is split up into work time L and leisure F, is somewhat more complicated. The reason for this is that it is not observable in the data and may in theory differ between individuals, depending on their circumstances. We therefore make an assumption, which is that each worker has a total of 60 hours at his or her disposal per week. Multiplying this figure with the sample mean of the number of adults per family (2.112; see Table 3.1) and accounting for the fact that a quarter has 12 weeks, we arrive at a $T = 1520.64$ hours. Finally, we need to make an assumption regarding the model parameter α in the Cobb-Douglas utility function, which defines both the utility elasticity of consumption and leisure. Rather than assuming a fixed value, we calculate the social welfare maximizing keep rate k^* for a range of α. Note, however, that we disregard $\alpha < 0.4$, as for this range we find that the slope of the indirect utility function V is not strictly monotonically increasing in k, which contradicts the theoretical model (see Section 3.2.1) and basic intuition. For the sake of illustration, we set $\alpha = 0.7$ for some graphical depictions of the indirect utility function as well as the social welfare function.

3.4. Results

3.4.1. Nonlinear Labor Supply Estimation

The results of our labor supply estimation are depicted in Table 3.2. We first focus on specification (1), which uses the families' total labor supply in a given quarter in hours ($Labor\ supply_{iq}$) as dependent variable. Note that this specification is also used as basis for our welfare analysis. A central result is that the labor supply corresponding to a hypothetical family which has control variables that are equal to the respective sample means is increasing in the keep-rate k. This finding is generally in line with the prediction of our theoretical

Table 3.2: *PREDICTED AVERAGE LABOR SUPPLY AND LABOR INCOME FOR DIFFERENT KEEP-RATES*

The table presents OLS estimates. The dependent variable of specification (1) is the labor supply of family i in experimental quarter q, *Labor supply*$_{iq}$. The dependent variable of specification (2) is the total labor income of family i in experimental quarter q, *Labor income*$_{iq}$. Both models are estimated without constant and all independent variables except for the $\mathbb{1}(k_{iq} = k)$'s and $No\ treatment_{iq}$ are de-meaned. Robust standard errors are reported in parentheses (clustered at the family level). *** denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level. Definitions and descriptive statistics on all are provided in Section 3.3.2.

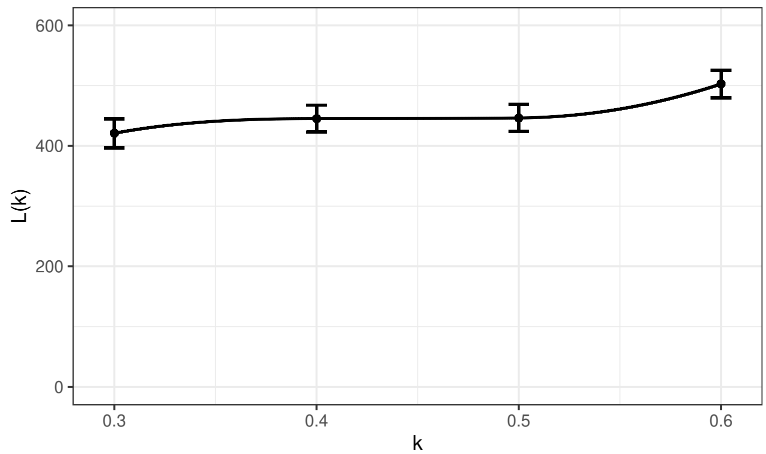
	(1)	(2)
$\mathbb{1}(k_{iq} = 0.3)$	420.919***	1295.262***
	(12.352)	(41.085)
$\mathbb{1}(k_{iq} = 0.4)$	445.200***	1313.494***
	(11.398)	(33.583)
$\mathbb{1}(k_{iq} = 0.5)$	446.289***	1399.793***
	(11.538)	(36.842)
$\mathbb{1}(k_{iq} = 0.6)$	502.634***	1475.266***
	(11.753)	(36.916)
G_{iq}	0.105***	0.445***
	(0.014)	(0.049)
$No\ treatment_{iq}$	497.283***	1635.510***
	(8.932)	(29.876)
$NIT\ payment_{iq-1}$	-0.360***	-1.207***
	(0.012)	(0.040)
$Persons\ 18\ or\ older_{iq}$	6.337	-1.820
	(4.073)	(13.175)
$Persons\ 17\ or\ younger_{iq}$	0.152	11.626
	(2.378)	(8.066)
$Female\ head_{iq}$	-144.413***	-522.334***
	(20.956)	(73.654)
$One\ head,\ dependents_{iq}$	35.302*	176.640**
(base: $Two\ heads_{iq}$)	(20.467)	(73.745)
$One\ head,\ no\ dependents_{iq}$	-32.326	-127.441
(base: $Two\ heads_{iq}$)	(31.151)	(100.406)
$Labor\ income\ pre-exp_i$	0.126***	0.458***
	(0.007)	(0.022)
$Wage\ rate\ pre-exp_i$	-54.196***	60.491***
	(8.138)	(19.622)
$Non-labor\ income_{iq-1}$	0.034	0.252
	(0.063)	(0.198)
$Welfare\ income_{iq-1}$	-0.332***	-1.103***
	(0.017)	(0.050)
$Social\ security\ income_{iq-1}$	-0.202***	-0.720***
	(0.021)	(0.065)
Year FEs	YES	YES
Adjusted R^2	0.830	0.829
Observations	30,307	30,307

model in Section 3.2. However, the fact that the increase in labor supply between $k = 0.50$ and $k = 0.60$ is larger than the increase between $k = 0.40$ and $k = 0.50$ contradicts the Marshallian labor supply function depicted in equation (3.5) in Section 3.2, which suggests that the optimal labor supply curve is concave, i.e., becomes flatter as k increases. The contradiction becomes more apparent in the graphical representation of the labor supply estimation as function of k, $\hat{L}(k)$, in Figure 3.1. Note that in the figure, the areas between the observed k's are imputed using a cubic spline monotonic interpolation approach (Dougherty et al., 1989; Forsythe et al., 1977; Hyman, 1983). Overall, we find that the difference in labor supply between the most generous keep-rate that was tested ($k = 0.60$) and the smallest keep-rate that was tested ($k = 0.30$) amounts to economically significant $503 - 421 = 82$ hours. Regarding our alternative specification in column (2) of Table 3.2 which analyzes labor income ($Labor\ income_{iq}$) instead of labor supply, we also find a positive relationship between the dependent variable and the different keep-rates. Note, however, that the labor income response does not exhibit the same pattern of differences between adjacent observed levels of k. This is due to the fact that – unlike in the theoretical model – wages are varying across families and time, which results in the labor supply and the labor income responses not being directly comparable.

Before we turn to the welfare analysis, let us briefly discuss the coefficient estimates on the other control variables. We find that labor supply is increasing in the guaranteed income level G_{iq}. This is surprising, given that our theoretical model predicts a negative relationship between labor supply and G_{iq} (see equation (3.5) in Section 3.2). However, in deciding on their labor supply for the current quarter, it can be argued that the actual received NIT payment from the previous quarter ($NIT\ payment_{iq-1}$), which also equals the NIT payment of the current quarter if the total family income remains unchanged, is the more relevant measure for families compared to G_{iq}, which states the transfer amount that is paid out only in the rare case where labor supply is equal to zero. The negative coefficient estimate on $NIT\ payment_{iq-1}$, which in absolute terms is three times as large as the estimate on G_{iq},

Figure 3.1: ***NONLINEAR LABOR SUPPLY ESTIMATION FOR DIFFERENT KEEP-RATES***

The figure depicts the nonlinear labor supply estimation for different keep-rates k. The black dots are the OLS coefficient estimates on different keep-rates k (see Table 3.2) and are depicted with corresponding 95% confidence intervals that are based on robust standard errors (clustered at the family level). The line connecting the OLS estimates is imputed using a cubic spline monotonic interpolation approach (Dougherty et al., 1989; Forsythe et al., 1977; Hyman, 1983).



matches the work-disincentive theory that more generous NIT plans reduce labor supply. Also in line with work-disincentive argument regarding the NIT is the positive coefficient on the indicator $No\ treatment_{iq}$, which suggest that untreated families exhibit a substantially higher labor supply. However, it has to be stressed that $No\ treatment_{iq}$ by definition is also equal to unity for families that are technically assigned to an NIT plan but did not receive any NIT payments in the previous period, as their total income exceeded the break-even threshold. Assuming that families with such high income also exhibit above average labor supply and assuming that their labor supply is somewhat constant across quarters, a part of the large coefficient on $No\ treatment_{iq}$ can be explained by a mechanical effect of high income earners being assigned to the control group.

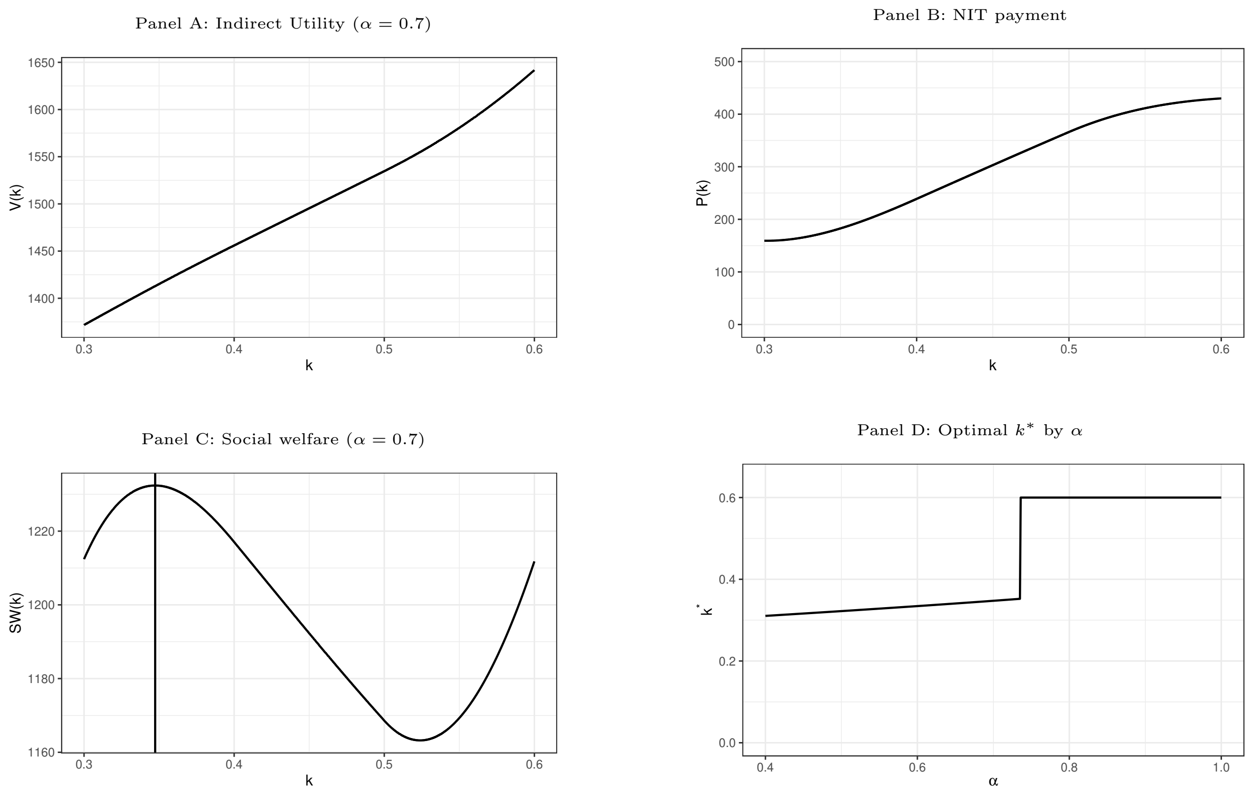
Regarding the family characteristics, the coefficients on the variables concerning the size and type of families suggest that families with a female head supply less labor and that labor supply is increasing with the amount of children. Furthermore, we find that families that received higher welfare and social security payments in the previous quarter work fewer hours on average – again, however, note that these effects may be in parts mechanical, as these payments often depend on and are decreasing in labor income. The coefficient on non-labor income is not significantly different from zero, which may be due to the fact that most families in the sample had no or negligibly little non-labor income and consequently the coefficient is identified from very little variation. As one would expect, we find that families with high pre-experimental income provide more labor on average. The negative coefficient on the pre-experimental wage rate, however, at first may seem surprising. A possible explanation may be that treated families could purposely reduce their labor supply to a level such that they become eligible for NIT payments. In this context, families with high wage rates need to reduce their labor supply to lower levels than families with low wage rates to obtain the same NIT transfer. Alternatively, the sign could be explained by the fact that high wage earners need to work less than low wage earners to generate the same income required to cover basic expenses. This latter mechanism is also valid for untreated families. A result that stands out in Table 3.2 is that the wage rate is the only control variable for which the coefficient estimate is both statistically significant from zero and differs in sign between our two model specifications. However, in the context of specification (2) which uses labor income rather than labor supply as dependent variable, the positive coefficient on the wage rate can simply be explained by the mechanical effect of income which is, holding work hours constant, increasing in the wage rate.

3.4.2. Optimal Social Welfare

We now turn to the analysis of social welfare, for which we use the nonlinear labor supply estimation $L^* = \hat{L}(k)$ that is depicted in Figure 3.1. We start out by analyzing the indirect utility function (using $\alpha = 0.7$), which is obtained by inserting $L^* = \hat{L}(k)$ into the first row of equation (3.9) (see Panel A of Figure 3.2). It can be seen that utility increases in the keep-rate k, which is in line with the theoretical model in Section 3.2. Moreover, the slope is almost perfectly linear, with only the section between $k = 0.50$ and $k = 0.60$ exhibiting a slightly steeper slope.

Figure 3.2: *WELFARE ANALYSIS WITH NONLINEAR LABOR SUPPLY*

The figure depicts the welfare analysis based on the nonlinear estimation of the labor supply, $L^* = \hat{L}(k)$ depicted in Figure 3.1. Panel A depicts the indirect utility, which is obtained by plugging $L^* = \hat{L}(k)$ into (the first row of) (3.9). α is set to 0.7. Panel B depicts the NIT payment, which is obtained by plugging $L^* = \hat{L}(k)$ into (3.11). Panel C depicts the social welfare, which is obtained by plugging $L^* = \hat{L}(k)$ into (3.13). Note that the line depicted in Panel C gives the difference between the lines depicted in the Panels A and B. The vertical line depicts the social welfare maximizing keep-rate k^*, which is equal to 0.347 for setting $\alpha = 0.7$. Finally, Panel D depicts the welfare maximizing keep-rate k^* for different levels of α. For the parameterization of the other parameters, see Section 3.3.2.



Panel B of the same figure depicts the NIT payments based on the nonlinear labor supply estimation. It is obtained by plugging $L^* = \hat{L}(k)$ into equation (3.11). The shape of the NIT payment curve is a direct reflection of the curvature of L^* depicted in Figure 3.1: a strong labor supply increase associated with a marginal increase in k, as observable in the areas between $k = 0.30$ and $k = 0.40$ as well as between $k = 0.50$ and $k = 0.60$, counteracts the mechanical increase in the NIT payments substantially, which explains the flatness in these areas. A more moderate labor supply response, on the other hand, as seen between $k = 0.40$ and $k = 0.50$, is associated with a smaller reduction of the same mechanical increase in the NIT payments, which explains the steeper increase of the NIT payment curve in this range.

Finally, Panel C depicts our notion of social welfare as function of the keep-rate k. Social welfare is simply defined as the difference between the indirect utility (Panel A) and the NIT payment (Panel B) (see equation (3.13) in Section 3.2). Regarding the shape of the function, it stands out that social welfare is increasing only in the two areas with the smallest labor supply response, that is, at the bottom and the top of our observed range of k. In between these areas, social welfare is decreasing in k, which can be explained with the moderate labor supply response and its aforementioned relationship with the NIT payments. We find a local maximum at around $k^* = 0.347$, which is at the lower end of our observed range of k. However, we need to discuss two major caveats regarding the determination of k^*. First, we cannot rule out that there is a global maximum outside of the observed range $[0.30,\ 0.60]$. Second, the parameter α, which defines the utility elasticities of consumption and leisure, is a key determinant of the shape of the indirect utility function and therefore also the social welfare function. Due to data availability, there is nothing that can be done about the first point. However, the second point, i.e., the dependence of k^* on α, can easily be investigated in more detail. Panel D depicts for each α the corresponding social welfare maximizing level of k^*.[27] We find that up until about $\alpha = 0.736$ the corresponding optimal local maxima lie in the range between $k^* = 0.310$ and $k^* = 0.352$, with the k^* linearly increasing in α.

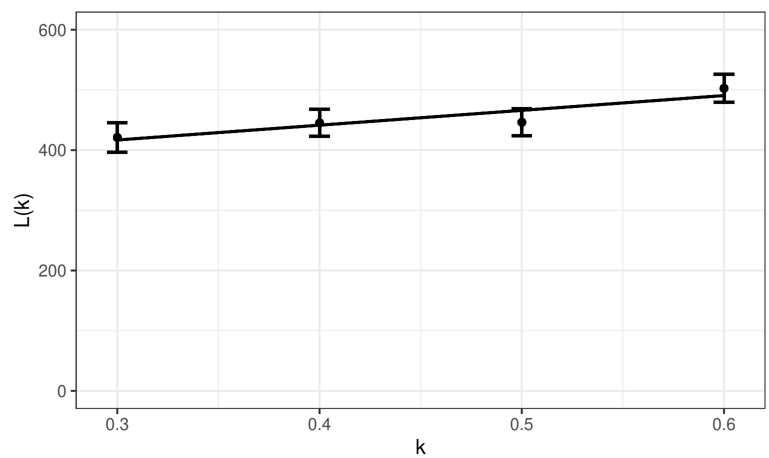
[27]Note again that we only consider $\alpha \geq 0.4$, as for $\alpha < 0.4$ the indirect utility function is not strictly monotonically increasing in k.

This small amount of variation in k^* within a rather large range of α shows that the finding with α fixed at 0.7 is quite robust in the sense that small deviations in α lead to similar optimal keep-rates k^*. However, at around $\alpha = 0.736$ there is a jump in k^*, which then has an optimal value of 0.60, i.e., the highest observed value of k in our analysis, for all α's above this threshold.

3.4.3. Comparison to Social Welfare Analysis with Linear Labor Supply

Figure 3.3: ***LINEAR LABOR SUPPLY ESTIMATION FOR DIFFERENT KEEP-RATES***

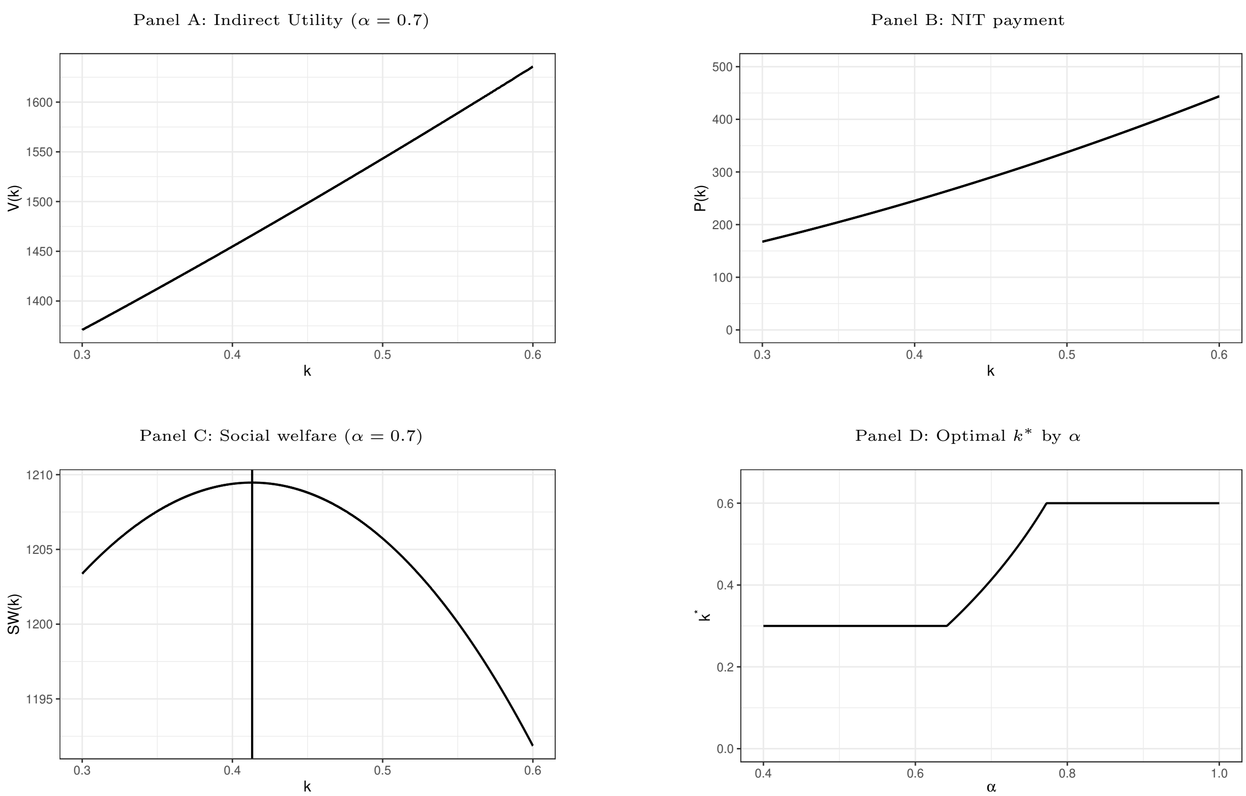
The figure depicts the linear labor supply estimation for different keep-rates k. The black dots are the OLS coefficient estimates on different keep-rates k (see Table 3.2) and are depicted with corresponding 95% confidence intervals that are based on robust standard errors (clustered at the family level). The line is fitted using an OLS regression with a constant.



In a last step of our analysis, to illustrate the importance of accounting for nonlinearities in the labor supply, we redo our welfare analysis on the basis of a linear labor supply function. This is supposed to mimic standard sufficient statistics welfare analysis where the curvature of the key behavioral response margin is summarized by a single parameter. In detail, the linear labor supply function is obtained by fitting a straight line between the four coefficients corresponding to the observed levels of k (see Table 3.2) using OLS. The result of this exercise is depicted in Figure 3.3. The linear estimate of $L^* = L(k)$ is then used to calculate indirect utility, the NIT payment, as well as social welfare using an otherwise identical

Figure 3.4: *WELFARE ANALYSIS WITH LINEAR LABOR SUPPLY*

The figure depicts the welfare analysis based on the linear estimation of the labor supply, $L^* = \hat{L}(k)$ depicted in Figure 3.3. Panel A depicts the indirect utility, which is obtained by plugging $L^* = \hat{L}(k)$ into (the first row of) (3.9). α is set to 0.7. Panel B depicts the NIT payment, which is obtained by plugging $L^* = \hat{L}(k)$ into (3.11). Panel C depicts the social welfare, which is obtained by plugging $L^* = \hat{L}(k)$ into (3.13). Note that the line depicted in Panel C gives the difference between the lines depicted in the Panels A and B. The vertical line depicts the social welfare maximizing keep-rate k^*, which is equal to 0.413 for setting $\alpha = 0.7$. Finally, Panel D depicts the welfare maximizing keep-rate k^* for different levels of α. For the parameterization of the other parameters, see Section 3.3.2..



parameterization compared to above. The results are depicted in Figure 3.4. We find a, compared to the analysis using the nonlinear labor supply, similarly linearly shaped indirect utility function (Panel A). However, regarding the NIT payments (Panel B), we now find a perfectly linear curve, which is a stark contrast compared to the analysis with nonlinear labor supply. As discussed above, the curvature of the NIT payment function is directly corresponding to the labor supply function, which is now linear. Finally, Panel C depicts the social welfare, which is now a strictly concave function of k, with a local maximum at $k^* = 0.413$. Note that this social welfare maximizing k^* is substantially larger than the one in the analysis using nonlinear labor supply above, which amounts to 0.347. We further find that also for other levels of α the welfare-maximizing levels of k^* substantially differ between the different methods of computing labor supply (compare Panels D of Figures 3.2 and 3.4). Note also that the linear labor supply yields larger ranges of α for which either the minimum or the maximum observed values of k are optimal. This underlines the importance of accounting for nonlinearities in key behavioral responses to policy parameters and suggests that conventional social welfare analysis that disregards nonlinearities may potentially be very imprecise.

3.5. Conclusions

We conduct a welfare analysis of the NIT that explicitly allows for nonlinearities in the labor supply response to changes in the take-back rate. Our key finding is similar to that of Kasy (2018), which is that welfare optimizing levels of the policy parameter differ between analyses accounting for nonlinearities or not. We think that this finding should be considered in future research and in the interpretation of results from standard sufficient statistics analyses. However, regarding the magnitude of the optimal levels of the keep-rate, we refrain from formulating any policy implications. The reason for this is threefold, as (i) the external validity of the data is likely low, as the NIT experiments that are analyzed were conducted about 50 years ago; (ii) the data quality is poor, which is mainly due to the fact that treatment assignment was not random, that specific NIT plans were only tested in one of the two experiments considered, and that there was misreporting;[28] and (iii) only a limited number of guaranteed income levels and take-back rates was tested, with the latter only covering a range of 30 percentage points. Ideally, a potential future NIT experiment is conducted as RCT using a wider range of tested plans. The formulas derived in this paper could serve as basis for calculating social welfare with data from such an RCT.

[28]Note that the ambiguity in empirical results based on the NIT experiments data is a much discussed topic in the empirical NIT literature, see, e.g., Ashenfelter and Plant (1990), Hum and Simpson (1993), or Widerquist (2005).

3.6. Appendix

3.6.1. Proof that $\partial V^*/\partial k > 0$

In the following, we proof that $\partial V^*/\partial k > 0$. First, rewrite the indirect utility function as

$$V^* = Y^{*\alpha} \left(T - L^*\right)^{1-\alpha}. \tag{3.17}$$

Taking the first derivative of (3.17) with respect to k yields

$$\begin{aligned}\frac{\partial V^*}{\partial k} &= \alpha Y^{*\alpha-1} \frac{\partial Y^*}{\partial k} \left(T - L^*\right)^{1-\alpha} + Y^{*\alpha}(1-\alpha)\left(T - L^*\right)^{-\alpha}(-1)\frac{\partial L^*}{\partial k} \\ &= \underbrace{Y^{*\alpha}(T-L^*)^{-\alpha}}_{>0} \underbrace{\left[\alpha Y^{*-1}\frac{\partial Y^*}{\partial k}\left(T - L^*\right) - (1-\alpha)\frac{\partial L^*}{\partial k}\right]}_{\equiv Z}.\end{aligned} \tag{3.18}$$

Since we assume that $T > L > 0$ and $Y = G + kwL$ with $G > 0$, $k > 0$, and $w > 0$ (see Section 3.2.1), it follows that both consumption and leisure are always strictly positive and therefore $Y^{*\alpha}(T - L^*)^{-\alpha} > 0$. The sign of $\partial V^*/\partial k$ therefore equals the sign of Z. Using the results for Y^* and L^* as well as their first derivatives from Section 3.2.1, we can write Z as follows:

$$\begin{aligned} Z &= \alpha \frac{1}{\alpha(kwT+G)} \alpha w T \left[T - \left(\alpha T - \frac{G(1-\alpha)}{kw}\right)\right] - (1-\alpha)\frac{G(1-\alpha)}{k^2 w} \\ &= \frac{\alpha w T}{kwT+G}\left[(1-\alpha)\frac{kwT+G}{kw}\right] - (1-\alpha)\frac{G(1-\alpha)}{k^2 w} \\ &= (1-\alpha)\frac{\alpha T}{k} - (1-\alpha)\frac{G(1-\alpha)}{k^2 w} \\ &= \frac{(1-\alpha)}{k}\underbrace{\left[\alpha T - \frac{G(1-\alpha)}{kw}\right]}_{=L^*}. \end{aligned} \tag{3.19}$$

$(1-\alpha)/k > 0$, as we assume $\alpha, k \in (0,1)$ (see Section 3.2.1). The hours worked in the optimum, L^*, must be strictly positive, too, as we assume $T > L > 0$. Therefore, $Z > 0$, and as a result, $\partial V^*/\partial k > 0$. ■

3.6.2. Envelope Theorem

In the following, we demonstrate the Envelope Theorem that states that the effect of a marginal change in k on utility in the optimum equals its direct derivative.[29] Or, in other words, changes in the choice variables (i.e., Y, L, and λ) induced by the change in k do not affect utility in the optimum. For the sake of notational clarity, we restate the utility maximization problem from (3.4) using the placeholder $U(Y, T-L)$ for the Cobb-Douglas utility function:

$$V(k, w, \alpha, T, G) = \max_{Y,L} \left[U(Y, T-L) + \lambda(G + kwL - Y) \right]. \tag{3.20}$$

The first-order conditions from the utility maximization problem in (3.20) state as follows (with $U_Y = \partial U / \partial Y$ and $U_L = \partial U / \partial L$):

$$Y: \quad U_Y - \lambda \stackrel{!}{=} 0, \tag{3.21a}$$

$$L: \quad -U_L + \lambda kw \stackrel{!}{=} 0, \tag{3.21b}$$

$$\lambda: \quad G + kwL - Y \stackrel{!}{=} 0. \tag{3.21c}$$

Solving the first-order conditions yields value functions for the optimal labor supply Y^* and L^* (depicted in detail in Section 3.2.1) as well as the household's marginal, or shadow, value of income in the optimum (see, e.g., Mas-Colell et al., 1995):

$$\lambda^* = U_Y = \alpha Y^{*(\alpha-1)} (T - L^*)^{(1-\alpha)}. \tag{3.22}$$

[29] Note that we borrow the notation from Keuschnigg and Wamser (2024).

The indirect utility function can then be obtained by plugging the value functions for Y, L, and λ into (3.20):

$$V^* = V(k, w, \alpha, T, G) = U(Y^*, T - L^*) + \lambda^*(G + kwL^* - Y^*). \tag{3.23}$$

In a next step, we take the derivative of (3.23) with respect to k:

$$\frac{\partial V^*}{\partial k} = \lambda^* w L^* + \underbrace{(U_Y - \lambda^*)}_{=0 \text{ (see (3.21a))}} \frac{\partial Y^*}{\partial k} + \underbrace{(\lambda^* kw - U_L)}_{=0 \text{ (see (3.21b))}} \frac{\partial L^*}{\partial k} + \underbrace{(G + kwL^* - Y^*)}_{=0 \text{ (see (3.21c))}} \frac{\partial \lambda^*}{\partial k}. \tag{3.24}$$

The way we arranged the derivative shows that the second, third, and fourth terms are equal to zero as the optimality conditions (3.21a)-(3.21c) must be fulfilled. The partial derivative hence equals $\lambda^* w L^*$, which equals equation (3.10). The latter can be shown by substituting the value functions Y^* (see (3.8)), L^* (see (3.5)), and λ^* (see (3.22)) into (3.24) and (3.10):

$$\begin{aligned}
& \underbrace{\lambda^* w L^*}_{(3.24)} = \underbrace{\frac{(1-\alpha)}{k} Y^{*\alpha} (T - L^*)^{-\alpha} L^*}_{(3.10)} \\
\leftrightarrow \quad & \underbrace{\alpha Y^{*(\alpha-1)} (T - L^*)^{(1-\alpha)}}_{=\lambda^* \text{ (see (3.22))}} w = \frac{(1-\alpha)}{k} Y^{*\alpha} (T - L^*)^{-\alpha} \\
\leftrightarrow \quad & \alpha Y^{*(-1)} (T - L^*) w = \frac{(1-\alpha)}{k} \\
\leftrightarrow \quad & \alpha \frac{1}{\alpha(kwT + G)} \left(T - \alpha T + \frac{G(1-\alpha)}{kw} \right) w = \frac{(1-\alpha)}{k} \\
\leftrightarrow \quad & \frac{1}{(kwT + G)} \left(\frac{(1-\alpha)(kwT + G)}{kw} \right) w = \frac{(1-\alpha)}{k} \\
\leftrightarrow \quad & \frac{(1-\alpha)}{k} = \frac{(1-\alpha)}{k}. \ \blacksquare
\end{aligned} \tag{3.25}$$

3.7. Acknowledgements

I would like to thank Giacomo Brusco, Anna Gumpert, Jaqueline Hansen, Michael Knaus, Wilhelm Kohler, Martin Ruf, and Georg Wamser for helpful comments and suggestions. Furthermore, I am thankful to Margaret Townsend from the Institute for Research on Poverty at the University of Wisconsin-Madison for her support regarding the data on the negative income tax experiments. Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – FOR2783 – is gratefully acknowledged.

3.8. References

Aboudi, R., Thon, D., & Zheng, M. (2014). Designing a basic income system with a social welfare function. *Journal of Public Economic Theory*, 16(4), 546-568.

Ashenfelter, O., & Plant, M. W. (1990). Nonparametric estimates of the labor-supply effects of negative income tax programs. *Journal of Labor Economics*, 8(1, Part 2), S396-S415.

Banerjee, A. V., Hanna, R., Kreindler, G. E., & Olken, B. A. (2017). Debunking the stereotype of the lazy welfare recipient: Evidence from cash transfer programs. *The World Bank Research Observer*, 32(2), 155-184.

Chetty, R. (2009). Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annual Review of Economics*, 1(1), 451-488.

Conlisk, J., & Watts, H. (1979). A model for optimizing experimental designs for estimating response surfaces. *Journal of Econometrics*, 11(1), 27-42.

Dougherty, R. L., Edelman, A. S., & Hyman, J. M. (1989). Nonnegativity-, monotonicity-, or convexity-preserving cubic and quintic Hermite interpolation. *Mathematics of Computation*, 52(186), 471-494.

Forsythe, G. E., Malcolm, M. A., & Moler, C. B. (1977). *Computer Methods for Mathematical Computations*. Prentice Hall.

Fuest, C., Greil, S., Hugger, F., & Neumeier, F. (2022). Global Profit Shifting of Multinational Companies: Evidence from CbCR Micro Data. CESifo Working Paper No. 9757.

Greenberg, D., Moffitt, R., & Friedmann, J. (1981). Underreporting and experimental effects on work effort: evidence from the Gary income maintenance experiment. *The Review of Economics and Statistics*, 581-589.

Greenberg, D., & Halsey, H. (1983). Systematic misreporting and effects of income maintenance experiments on work effort: evidence from the Seattle-Denver experiment. *Journal of Labor Economics*, 1(4), 380-407.

Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Vol. 2). Springer.

Hausman, J. A. (1985). Taxes and labor supply. In A. J. Auerbach & M. Feldstein (Eds.), *Handbook of Public Economics* (Vol. 1, pp. 213-263). Elsevier.

Heckman, J. J. (1993). What has been learned about labor supply in the past twenty years? *The American Economic Review*, 83(2), 116-121.

Hum, D., & Simpson, W. (1993). Economic response to a guaranteed annual income: Experience from Canada and the United States. *Journal of Labor Economics*, 11(1, Part 2), S263-S296.

Hyman, J. M. (1983). Accurate monotonicity preserving cubic interpolation. *SIAM Journal on Scientific and Statistical Computing*, 4(4), 645-654.

Kasy, M. (2018). Optimal taxation and insurance using machine learning–Sufficient statistics and beyond. *Journal of Public Economics*, 167, 205-219.

Keeley, M. C., & Robins, P. K. (1978). The Design of Social Experiments: A Critique of the Conlisk-Watts Assignment Model. Research Memorandum no. 57. Stanford Research Institute, Center for the Study of Welfare Policy.

Keeley, M. C. (1981). *Labor Supply and Public Policy: A critical Review*. Academic Press.

Keuschnigg, C., & Wamser, G. (2024). *Economics of Taxation*. Mohr Siebeck. Forthcoming.

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic Theory*. Oxford University Press.

Mathematica Policy Research, Inc., & Social & Scientific Systems, Inc. (1980, June 30). Record Format Description for the Cross-Site Analysis File (revised April 1981).

Mathematica Policy Research, Inc. (1980, October). User's guide for the cross-site analysis file from the New Jersey, Gary, Seattle, and Denver income maintenance experiments (revised January 1981).

Potdar, K., Pardawala, T. S., & Pai, C. D. (2017). A comparative study of categorical variable encoding techniques for neural network classifiers. *International Journal of Computer Applications*, 175(4), 7-9.

Robins, P. K. (1985). A comparison of the labor supply findings from the four negative income tax experiments. *Journal of Human Resources*, 567-582.

Saez, E. (2002). Optimal income transfer programs: intensive versus extensive labor supply responses. *The Quarterly Journal of Economics*, 117(3), 1039-1073.

Verho, J., Hämäläinen, K., & Kanninen, O. (2022). Removing welfare traps: Employment responses in the Finnish basic income experiment. *American Economic Journal: Economic Policy*, 14(1), 501-22.

Widerquist, K. (2005). A failure to communicate: What (if anything) can we learn from the negative income tax experiments?. *The Journal of Socio-Economics*, 34(1), 49-81.