

MIGRATION AND WAGE EFFECTS OF TAXING TOP EARNERS: EVIDENCE FROM THE FOREIGNERS' TAX SCHEME IN DENMARK*

HENRIK JACOBSEN KLEVEN
CAMILLE LANDAIS
EMMANUEL SAEZ
ESBEN SCHULTZ

This article analyzes the effects of income taxation on the international migration and earnings of top earners using a Danish preferential foreigner tax scheme and population-wide Danish administrative data. This scheme, introduced in 1991, allows new immigrants with high earnings to be taxed at a preferential flat rate for a duration of three years. We obtain two main results. First, the scheme has doubled the number of highly paid foreigners in Denmark relative to slightly less paid—and therefore ineligible—foreigners. This translates into a very large elasticity of migration with respect to 1 minus the average tax rate on foreigners, between 1.5 and 2. Second, we find compelling evidence of a negative effect of the scheme-induced reduction in the average tax rate on pretax earnings of foreign migrants at the individual level. This finding can be rationalized by a matching frictions model with wage bargaining where there is a gap between pay and marginal productivity. *JEL* Codes: H31, J61.

I. INTRODUCTION

Tax-induced international mobility of talent is a controversial public policy issue, especially when tax rates differ substantially across countries and migration barriers are low, as in the case of the European Union. High top tax rates may induce top earners to migrate to countries where the tax burden is lower, thereby potentially creating tax competition across countries. Many European countries have introduced preferential tax schemes to high-skilled foreigners.¹ This debate raises two

* We thank Raj Chetty, Michael Devereux, Amy Finkelstein, Rick Hornbeck, Caroline Hoxby, Lawrence Katz, Wojciech Kopczuk, Claus Kreiner, Ilyana Kuziemko, Bentley MacLeod, Alan Manning, Jonathan Meer, James Poterba, Andrei Shleifer, three anonymous referees, and numerous seminar participants for very helpful comments and discussions. Financial support from NSF Grant SES-1156240, the Center for Equitable Growth at UC Berkeley, EPRN, and the European Tax Policy Forum is gratefully acknowledged.

1. For example, preferential tax schemes for high-skilled foreign workers have been introduced in Belgium, Denmark, Finland, Netherlands, Portugal, Spain, Sweden, and Switzerland. A summary of all such existing schemes in OECD countries is provided by OECD (2011), Table 4.1, p. 138.

© The Author(s) 2013. Published by Oxford University Press, on behalf of President and Fellows of Harvard College. All rights reserved. For Permissions, please email: journals.permissions@oup.com

The Quarterly Journal of Economics (2014), 333–378. doi:10.1093/qje/qjt033.

Advance Access publication on October 24, 2013.

important questions. First, how responsive is international migration by high-skilled workers to tax differentials across countries? Second, what is the effect of lowering top tax rates on the wages of top earners? In particular, do preferential foreigner tax schemes only benefit highly compensated foreigners, or do local employers capture part of the benefit? Both questions are crucial for evaluating international tax design for top earners, whereas the second question is also key for understanding the functioning of the labor market for top earners.² This article breaks new ground on these questions using quasi-experimental variation created by a Danish preferential tax scheme for high-earning immigrants.

Although an enormous empirical literature has studied labor supply and taxable income responses to taxation (as surveyed by Blundell and MaCurdy 1999; Saez, Slemrod, and Giertz 2012), there is very little empirical work on the effect of taxation on the spatial mobility of individuals, and especially international mobility among high-skilled workers.³ Furthermore, there is also very little work trying to exploit tax variation to cast light on the wage-setting process, particularly among top earners. Most studies assume that taxes do not affect individual wage rates directly.⁴ The wage effects we obtain provide strong empirical

2. An additional justification for such schemes mentioned in policy debates are positive spillover effects of attracting high-skill workers on the economy. Unfortunately, our empirical setting cannot be used to address this question compellingly (see the working paper version [Kleven et al. 2013] for a detailed discussion).

3. A small literature has considered tax-induced mobility of people across local jurisdictions within countries. See, for example, Kirchgassner and Pommerehne (1996) and Liebig, Puhani, and Sousa-Poza (2007) on mobility across Swiss cantons, Bakija and Slemrod (2004) and Young and Varner (2011) on mobility across U.S. states. On international mobility, Kleven, Landais, and Saez (2013) analyze the labor market for professional football players across 14 European Union countries and find evidence of tax-induced mobility responses. However, football players might be substantially more mobile than other high-skilled workers, because they earn most of their lifetime income over a short period and their profession involves little country-specific capital.

4. A few studies have tried to estimate standard demand-driven wage incidence effects (see Fullerton and Metcalf 2002 for a survey). The empirical identification is very difficult, because such incidence effects represent market-level changes in wages. By contrast, the incidence effects we uncover represent individual-level or match-level changes in wages, which are interesting in their own right and can be identified compellingly.

evidence in favor of the labor market model with matching frictions and wage bargaining.

In 1992, Denmark enacted a preferential tax scheme for high-earning foreigners, who sign contracts for work in Denmark after June 1, 1991. Under this scheme, the tax rate on labor earnings is reduced to a flat rate of about 30% for a total period of up to three years. Eligibility for the scheme requires annualized earnings above a threshold (indexed to average earnings growth and equal to about 100,000 euros in 2009), corresponding roughly to the 99th percentile of the distribution of individual earnings in Denmark. The scheme also applies to Danish citizens who have lived abroad with tax residence outside Denmark for a period of at least 3 years (10 years since 2011). This scheme is much more generous than the regular tax system, which imposes a top marginal tax rate of about 62% above 47,000 euros (as of 2009).⁵ Absent the special tax scheme, workers with earnings above the scheme threshold would face average income tax rates of around 55%, almost twice as high as the scheme rate. After the three years of preferential tax treatment, the taxpayer becomes subject to the ordinary tax schedule on subsequent earnings.

The scheme creates large discontinuities in tax liability depending on the contract start date (before and after June 1, 1991), duration of stay in Denmark (three-year rule), and earnings level (earnings eligibility threshold). Hence, the reform generates sharp quasi-experimental variation along several different dimensions and provides a powerful way of identifying the effect of taxation on migration and earnings using bunching and difference-in-differences methods. For this analysis, we have obtained access to matched employer–employee administrative data for the full population of Danish residents (Danish citizens and foreigners) since 1980. The data include detailed information on citizenship, immigration history, income and tax variables, labor market variables, and sociodemographic variables.⁶

Our analysis of the foreigners' tax scheme in Denmark yields two main empirical results.

5. As is well known, Denmark has the highest tax revenue/GDP ratio among OECD countries (OECD 2012, Part I, Table A, p. 19) so that the tax burden on high wage earners is particularly high (OECD 2013).

6. The data were specifically prepared by Statistics Denmark for our research project and securely accessed through a server at the Centre for Economic and Business Research (CEBR).

First, we obtain compelling evidence that the scheme had a very large effect on the number of highly paid foreigners in Denmark. The number of foreigners paid above the eligibility threshold doubles relative to the number of foreigners paid slightly below the threshold after the scheme is introduced. This effect builds up in the first five years of the scheme and remains stable afterward. As a result, the fraction of foreigners in the top 0.5% of the earnings distribution is 7.5% in recent years compared to a 4% counterfactual absent the scheme. This is consistent with a very large elasticity of migration with respect to the net-of-tax rate (defined from now on as 1 minus the average tax rate) among foreigners. The estimated elasticities are between 1.5 and 2. The resulting revenue-maximizing tax rate for a scheme targeting highly paid foreigners is therefore relatively small (about 35%) and corresponds roughly to the current tax rate on foreigners in Denmark once we account for other relevant taxes (VAT and excises).⁷ It can therefore be desirable from a single-country revenue perspective to adopt such preferential schemes for highly paid foreigners. At the same time, these schemes impose negative fiscal externalities on other countries by reducing their capacity to collect taxes from top earners. This tension between country welfare and global welfare in the design of individual income tax policy has loomed large in the debate about tax competition for a long time, but our article provides the first compelling evidence that this is indeed a significant empirical issue.

Second, we find compelling evidence of a negative effect of scheme-induced increases in the net-of-tax rate on pretax earnings at the individual level. We show in a difference-in-differences setting that scheme eligible foreign immigrants experience a 5–10% decline in their pretax earnings relative to noneligible foreign immigrants after the introduction of the scheme. Most important, we find that migrants who stay in Denmark beyond the three-year scheme duration experience a sharp *increase* in their earnings when the scheme elapses. By focusing on a panel of stayers on each side of the three-year discontinuity, we ensure that this result is not driven by selection

7. Importantly, the revenue-maximizing tax rate on natives is much higher, because the elasticity of migration with respect to the net-of-tax rate among natives is much lower. We can partly measure the elasticity for natives by estimating how many expatriate Danish natives come back because of the scheme.

into Denmark or by nontax aspects of wage tenure profiles. In the competitive labor market model where pay equals productivity, we would expect a *decrease* in labor supply and therefore in earnings when the individual is faced with the much higher regular tax rate. Related, we find evidence of bunching in the earnings distribution of immigrants just above the scheme eligibility threshold, but no evidence of a hole below the threshold. This suggests that bunchers are coming from the low-tax side of the threshold, which is inconsistent with the competitive labor supply model, where bunching is driven by individuals on the high-tax side increasing their labor supply to qualify for the scheme. Those two empirical findings can be rationalized by a simple matching frictions model with wage bargaining where there is a gap between pay and marginal productivity. To our knowledge, this is the first time that tax variation has been used to provide evidence of a wedge between pay and productivity and therefore potential for wage bargaining effects, as in the widely influential theory of job search.⁸

Finally, although our study is based on a single country characterized by a certain size, culture, quality of life, immigration tradition, and so on, we argue that the empirical insights have broader relevance. In particular, because Denmark is a small and homogeneous country starting from a small base of highly paid foreigners, the migration elasticity with respect to the net-of-tax rate on foreigners is likely to be larger in Denmark than in countries starting from larger bases of highly paid foreigners. Indeed, an important insight from the theory of tax competition (Kanbur and Keen 1993) is that tax havens tend to be small countries, because they have small tax bases relative to the global economy and therefore feature larger tax base elasticities. Our findings are consistent with this theoretical mechanism. Furthermore, even if such large migration elasticities do not carry over to larger countries, the combined efforts of many small countries in attracting high-skilled labor can have nontrivial consequences for large countries.

The article is organized as follows. Section II describes the Danish foreigner tax scheme and the administrative data we use.

8. There is an enormous structural empirical literature in labor economics using the search framework, but there is much less work that directly tests the validity of the search model against the standard frictionless model (see Mortensen and Pissarides 1999 for a survey).

Section III lays out a simple theoretical framework with matching fractions and wage bargaining and contrasts its implications with the competitive labor market model. Section IV presents the empirical analysis. Section V concludes.

II. INSTITUTIONAL BACKGROUND AND DATA

II.A. *The Foreigner Tax Scheme in Denmark*

In 1992, Denmark enacted a preferential tax scheme for foreign researchers and high-earning foreigners in all other professions who sign contracts for employment in Denmark after June 1, 1991. The scheme is commonly known in Denmark as the *Researchers' Tax Scheme*. In this article, we focus solely on top earners and exclude foreign researchers in the scheme from our analysis. When the scheme was first introduced, it offered a flat income tax rate of 30% in lieu of the regular progressive income tax with a top marginal tax rate of 68% and an average tax rate on high-income workers of around 55%. The scheme rate was reduced to 25% in 1995, but at the same time a payroll tax of 8% was gradually phased in between 1994 and 1997, leaving the total scheme rate roughly unchanged around 30%.⁹ The scheme can be used for a total period of up to 36 months, after which the taxpayer becomes subject to the ordinary income tax schedule. The 36 months do not have to be taken together, but can be divided into multiple spells. As we discuss in more detail in the next section, this form of duration dependence creates a discrete jump in *marginal* lifetime tax liability with respect to duration of stay in Denmark at the three-year cutoff and hence a *kink* in the lifetime budget set as a function of duration.

The scheme imposes three key eligibility requirements. First, the taxpayer cannot have been a Danish tax resident during the three years prior to going on the scheme.¹⁰ Citizenship plays no formal role in determining eligibility, and therefore

9. Because the payroll tax is deductible in the base for the regular income tax, the total scheme rate from 1997 onward can be calculated as $8\% + 0.92 \cdot 25\% = 31\%$.

10. The 3-year nonresidency requirement was changed to 10 years in 2011, but this lies outside our data period. Furthermore, for taxpayers who split scheme take-up into several spells, the 3-year (10-year) nonresidency requirement applies to each spell separately. But in assessing whether a taxpayer has been tax liable in Denmark prior to a given scheme spell, time spent in Denmark under prior scheme spells is not counted.

Danish citizens who have been foreign tax residents for at least three years can also apply for the scheme, a group we refer to as *Danish expatriates*. Second, scheme workers must carry out most of the work within Danish borders. A “reasonable” amount of work traveling is allowed for, but as a guideline at least two-thirds of the work should be carried out within Denmark.¹¹ This ensures that the scheme cannot be used by individuals working outside Denmark but employed by Danish multinational corporations. From now on, we refer to such gaming of the scheme as “reporting effects” (as opposed to real migration effects). Related, scheme rules put additional restrictions on participation by individuals who are (i) recruited by multinational corporations from one of their foreign subsidiaries, (ii) part of the management or ownership of the Danish firm hiring them. The first of these reduces somewhat the scope for multinationals to take advantage of the scheme by moving (e.g., switching) employees between foreign and Danish subsidiaries. Third, unless the worker qualifies as a researcher, annualized wage earnings must be above an eligibility threshold. The threshold grows roughly at the rate of average earnings, and it always lies between the 99.2th and 99.4th percentiles of the Danish wage earnings distribution.¹² It is equal to 765,600 Danish kroner (about 103,000 euros) in 2009. As the preferential scheme rate applies to *total* earnings conditional on eligibility, the earnings requirement creates a discrete jump in *total* annual tax liability at the threshold—a notch in the annual budget set as a function of earnings.

The scheme treatment has to be requested by the employer, who is responsible for showing the tax authorities that the various qualifying requirements are met. Importantly, the earnings eligibility threshold applies only to earnings with the specific employer requesting the scheme; having other sources of income or earnings does not help qualify. The eligibility threshold must be met on an annualized basis, so that a six-month contract, for

11. This rule was relaxed in 2009 (after our data period) and now is no longer an explicit requirement that scheme work is carried out within Denmark.

12. As we shall see, in our empirical analysis, we need to impute a scheme threshold for years before the scheme enactment. We do so by assuming that the threshold to average earnings ratio in years before enactment is the same as 1991, the first year the scheme is in place. Average earnings are estimated on the full population sample, including all workers with any positive earnings. Because of the great stability of the earnings distribution in Denmark, other methods for imputing the threshold before 1990 deliver almost identical results.

example, is subject to a threshold equal to half the annual threshold. Perquisites such as cars or housing allowances are included in earnings eligible for the scheme and are also taxed at the same flat rate.

If the scheme beneficiary has other income besides qualifying earnings, that income is taxed according to the standard progressive income tax schedule independently of scheme earnings. Therefore, scheme earnings are effectively taxed at a flat rate completely independently of the other circumstances of the individual. In particular, when the 36 months of scheme duration ends within a given calendar year, and the individual stays in Denmark, any postscheme earnings will be taxed according to the regular schedule. There is no prorating for nonscheme earnings (or unearned income) taxed according to the regular tax schedule. Spouses of scheme recipients are taxed according to the regular tax (except if they themselves qualify for the scheme). As tax residents of Denmark, scheme earners have access to the same public goods as all other residents, including public health insurance and schooling.

To summarize, the Danish tax scheme creates the following quasi-experimental variation. First, the scheme introduced a much lower tax rate on a specific sample of workers (high-earning immigrants; not tax liable in Denmark three years prior) arriving in Denmark after June 1, 1991. This variation provides a suitable setting for a difference-in-differences analysis of migration effects. Second, the scheme introduced a notch in the individual budget constraint creating very strong incentives for foreigners to have earnings above the eligibility threshold. Third, the scheme introduced a three-year duration kink among those who migrate to Denmark, providing sharp quasi-experimental variation that can be used to study the effects of taxation on duration of stay and on the tax incidence on individual earnings for workers who stay beyond the three-year scheme duration.

II.B. Data and Summary Statistics

1. Administrative Data. The empirical analysis is based on an administrative data set that includes tax and payroll records for the full population of resident individuals in Denmark (both Danish citizens and foreigners) since 1980. The data includes detailed information about earnings, tax variables, labor market variables, and sociodemographic variables at an annual

frequency. Most important, the data contain detailed citizenship and migration information such as daily dates of entry and exit. Individuals working in Denmark must obtain a personal identification number (CPR) to pay withholding taxes, rent an apartment, register with health insurance, and soon. The application for a CPR contains detailed questions about citizenship, country of origin, and date of entry in Denmark. The registry administration updates this information in case an individual leaves the country. The data also contain detailed information specifically for scheme beneficiaries, such as the start and end dates of labor contracts. Unfortunately, because this information was not computerized for the first years of the scheme, we do not have individual earnings information available for scheme beneficiaries from 1991 to 1994 (but all scheme beneficiaries are in the data).

2. Summary Statistics. Table I presents summary statistics. As described above, the scheme law (during our data period) stipulates that participants carry out most of the work within Danish borders and therefore take up actual residency in Denmark. First, the table shows that 94.9% of all scheme workers entering between 1991 and 2010 have obtained a CPR. According to CPR legislation, to obtain such a number it is a requirement that the individual resides in Denmark for at least three months and, for non-Nordic immigrants, that the individual holds a Danish residence permit. Second, it is a legal requirement for someone moving to Denmark to report his or her address to the local municipality. This information is only available for entrants up to 2006. Among foreigners taking up the preferential tax scheme between 1991 and 2006 and having a CPR, Table I shows that 98.5% have an actual address in Denmark that the local municipality knows about.¹³ To capture real migration effects as opposed to reporting effects, we exclude from our analysis scheme earners with missing CPRs as well as scheme earners with missing local address information (for 1991–2006 entrants).

Table I shows the following. First, the average duration of stay by scheme participants (including both time under the scheme and time after the scheme elapses) is 2.35 years, with

13. CPR observations are not a perfect subset of those with municipal addresses (1.5% of CPR observations have no address observation), because it is possible to satisfy the requirements for CPR at a point in time and then move out without necessarily losing the CPR number immediately.

TABLE I
DESCRIPTIVE STATISTICS

	(1) Mean	(2) Std. dev.
Scheme employees 1991–2010		
Number of scheme spells = 11,642		
Fraction with CPR (full sample 1991–2010 entrants)	.949	.220
Fraction with CPR (among 1991–2006 entrants only)	.945	.226
<i>among which fraction with reported address (KOM)</i>	.985	.122
Duration of stay (years)	2.345	1.448
<i>fraction > 3 years</i>	.251	.434
Scheme earnings (in '000s of 2009 euros)	163.5	281.2
<i>as a fraction of threshold</i>	1.53	2.67
Average tax rate	.308	.003
Age	39.97	8.80
Scheme take-up rate: .81		
Firms 1991–2010		
<i>N</i> = 2,235		
Number of employees	438.06	2,316.47
average gross wage (in '000s of 2009 euros)	58.7	40.9
Number of scheme employees	1.8	3.2
Number of native employees above threshold	13.99	61.65

Notes. This table presents a number of summary statistics for all individuals who used the foreigners' tax scheme in Denmark from 1991 to 2010 (we exclude foreigners who qualify as researchers). The top panel reports spell-level summary statistics. Multiple contracts for the same individual with the same employer are counted as one unique spell. 11,642 distinct spells have been recorded in the scheme. For 94.9% of spells, we were able to find a CPR from the registry files. For entrants in years 1991–2006, we were also able to check that among the 94.5% of individuals with CPR, 98.5% also have an address reported to the local municipal authority (KOM). KOM data are not available after 2006. All other statistics in the table (and in the subsequent analysis of the article) are restricted to spells with proof of residence (both KOM and CPR for entrants until 2006, and CPR for entrants in 2007–2010). The average tax rate is 30% before 1995 and $25\%(1 - \text{AMB}) + \text{AMB} = 31\%$ after 1995 where AMB is a flat payroll tax rate of 8%. The bottom panel reports firm-level summary statistics. 2,235 distinct firms hired scheme workers at some point between 1991 and 2010. The statistics report the average characteristics of the firms at the time they have at least one scheme employee. Scheme-participating firms are usually large firms with an average of 438 employees in total. The average yearly gross wage in scheme-participating firms (including scheme employees is 58,700 euros in 2009).

about a quarter of scheme beneficiaries staying beyond the maximum scheme duration of 36 months. Second, average scheme earnings equal 153% of the threshold (about 163,500 euros in 2009) and the average tax rate of scheme earners is around 30%. Third, scheme earners are relatively young (40 years on average). Finally, scheme beneficiaries work in large firms (440 employees on average) that pay relatively well—on average 59,000 euros per employee in 2009 compared to the average of 36,500 euros for all employees in Denmark. These firms employ relatively few scheme workers (1.8 on average) even though they employ on average 14 workers paid above the threshold.

Online Appendix Figure A1 reports the composition of scheme participants (excluding researchers) by country of citizenship (Panel A) and industrial sector (Panel B) from 1991 to 2006. The vast majority of scheme workers come from advanced economies: 25% are from Nordic countries (outside Denmark), 10% are Danish citizens (who qualify by being foreign tax residents for at least three years), 19% come from the United Kingdom or Ireland, 10% come from North America, and about 20% come from Germany, France, and Benelux (Belgium, Netherlands, Luxembourg) combined. The composition by industrial sector reveals that all sectors made use of the scheme but with an overrepresentation of the financial sector and the sports/entertainment industry.

We can compute a scheme take-up rate as the share of scheme foreigners among all foreigners arriving in Denmark who have not been Danish residents for at least three years and who have (annualized) earnings above the eligibility threshold. This take-up rate is 81%, high but still significantly below 100% for two main reasons. First, the take-up rate is a lower bound by construction, because it does not account for all of the eligibility rules described in the previous section.¹⁴ Furthermore, it is possible that some apparently scheme-eligible foreign immigrants were not in compliance with the earnings eligibility threshold, because they did not have one labor contract specifying earnings above the threshold *ex ante* even if they ended up having total earnings above the threshold *ex post*. Second, companies have to file a scheme application for each eligible employee and also meet certain firm-level requirements (such as tax withholding). Some companies may have been unwilling to bear these administrative burdens. Lack of awareness of the scheme by Danish employers is an unlikely explanation for imperfect take-up given the salience of this scheme in Denmark.

III. CONCEPTUAL FRAMEWORK

In this section, we contrast two labor market models that yield different predictions of the effects of the scheme on migration and earnings. We start with the competitive labor market

14. This includes the rule that scheme work must be carried out in Denmark and the restrictions on hiring by multinationals and of individuals involved in the management or ownership of the firm.

model conventionally used in empirical labor supply and tax studies, which we refer to as the competitive labor supply model. Since some of our empirical results cannot be easily rationalized within this model, we set out an alternative simple model with matching frictions and wage bargaining that can account for all our empirical findings. In both models, we simplify the exposition by abstracting from standard wage incidence effects whereby an influx of high-skilled workers creates a labor market-wide reduction in the wages of substitutable workers. We assume away such incidence effects because our empirical analysis always compares high-skilled foreign workers in the *same* labor market and is therefore not affected by market-wide wage changes.¹⁵

III.A. *Competitive Labor Supply Model*

In the standard frictionless labor supply model, workers receive a wage equal to their marginal product and choose location and labor supply to maximize a utility function that depends on net-of-tax earnings, labor supply, and location. The scheme affects behavior along three dimensions: (i) the migration decision, (ii) the duration of stay conditional on migrating, and (iii) labor supply and earnings among migrants. Let us review each dimension in turn.

1. Migration. The scheme reduces the average tax rate on high-earning immigrants to Denmark, which increases immigration by eligible workers to this country. Conditional on earnings, scheme-induced immigrants have weaker preferences for Denmark (but not necessarily lower productivity) than those who would have come anyway. Importantly, the scheme incentivizes only *real* migration responses (as opposed to tax avoidance) by requiring that scheme work is carried out mostly within Danish borders. As explained in Section II, the scheme also restricts the degree to which multinational firms can take advantage of the scheme by moving (e.g., switching) employees between foreign and Danish subsidiaries of the same firm.

15. Importantly, we assume away imperfect substitutability across groups of foreign workers by year of entry. As discussed in Section IV.B, relaxing this assumption could rationalize our empirical findings on earnings within the competitive labor market model.

2. *Duration.* The lower scheme tax rate during the first three years of stay in Denmark should create bunching in durations precisely at three years. Following Saez (2010), with smooth and convex preferences over lifetime income and time spent in Denmark (an assumption relaxed below), Figure I, Panel A illustrates duration responses to the introduction of a lower tax rate below the duration threshold D^* . The discrete jump in the tax rate at D^* produces a *kink* in the *lifetime* budget constraint of duration of stay (x -axis) versus disposable lifetime income (y -axis). This creates excess bunching at D^* by everybody with durations in the interval $[D^* - \Delta D^*, D^*]$ absent the scheme, and increases durations in the interior of the scheme period by those initially below $D^* - \Delta D^*$.

In practice, preferences may not be smoothly convex with respect to time spent in Denmark. For example, constraints arising from children's schooling or family/work obligations may create discontinuous preferences for specific time spells in Denmark (e.g., one year abroad only), creating lumpiness in the distribution of durations with the following implications for our duration analysis. First, there will be excess bunching around every yearly threshold, and not just at three years. Second, *tax-driven* bunching at the three-year threshold (as opposed to natural preference-driven bunching) will be attenuated by the lumpiness of duration preferences, since there will be fewer potential bunchers located in the continuous interval just below three years absent the scheme. Depending on preferences, bunching at three years will not necessarily be larger than bunching at one and two years, and so a bunching strategy is not ideally suited to estimate duration responses. Third, the new migrants that come to Denmark induced by the lower scheme tax rate may not stay for the full duration of the scheme, but will be distributed everywhere between zero and three years.¹⁶ Hence, the post-scheme duration distribution will feature excess mass everywhere below three years compared with the prescheme duration distribution. This implies that the average duration among all migrants may be lower with the scheme than without (even if each potential migrant stays longer than she otherwise would have).

16. Some of them may also stay longer than three years in the presence of path-dependent location preferences (such that the utility from location i in year t depends positively on being in location i in year $t - 1$).

3. *Labor Supply and Earnings.* The scheme reduces sharply the average tax rate for immigrants with earnings above the eligibility threshold. This creates an upward *notch* in the *annual* budget constraint of pretax earnings (x -axis) versus disposable post-tax earnings (y -axis) at the scheme eligibility threshold z^* as depicted in Figure I, Panel B. According to the competitive labor supply model, such a notch induces foreign workers in an interval $[z^* - \Delta z^*, z^*)$ to increase earnings to (at least) the

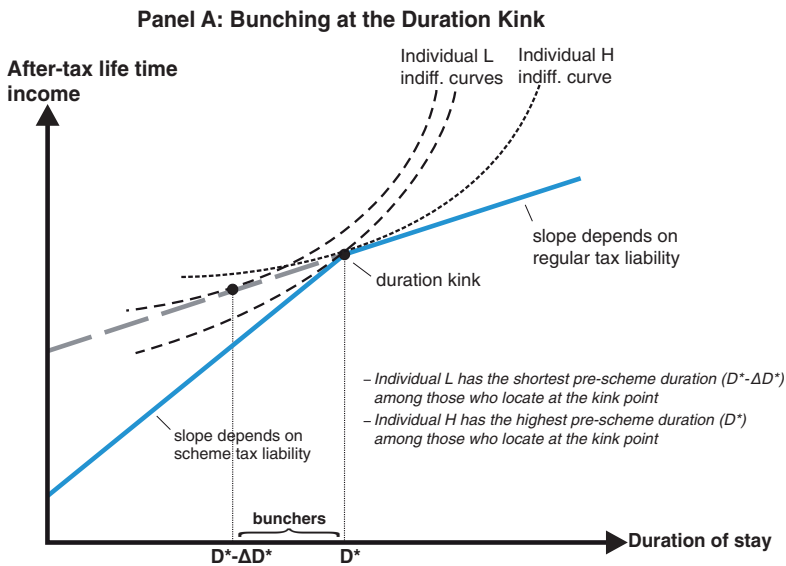
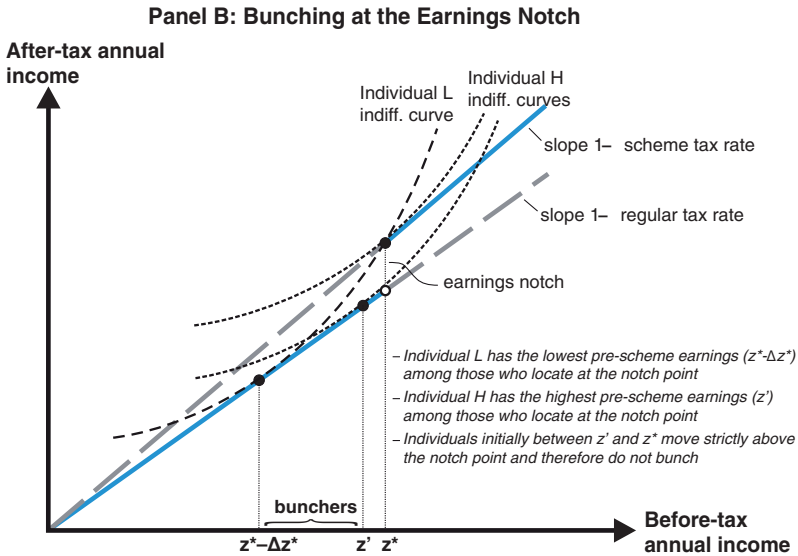


FIGURE I

Responses to the Tax Scheme in the Competitive Labor Supply Model

The figure illustrates behavioral responses to the duration kink (Panel A) and the earnings notch (Panel B) in the standard labor supply model of Section III.A. Panel A depicts the kink created by the scheme in the lifetime budget defined over duration D of stay in Denmark and lifetime consumption. The solid line represents the budget under the scheme with maximum duration D^* ($= 3$ years). The dashed gray line represents the budget when there is no scheme (prescheme). A scheme with finite duration generates bunching at the threshold D^* as all individuals with preferences in between individual L and individual H bunch at the kink. Panel B depicts the notch created by the scheme at the earnings eligibility threshold z^* in the annual budget set of individuals (solid line). The notch creates excess bunching at the eligibility threshold and a corresponding hole in the density distribution just below the threshold. Above the notch, the scheme increases the net-of-tax wage rate and therefore increases labor supply and earnings if the uncompensated labor supply elasticity is positive.

(continued)



threshold z^* to qualify for the scheme, thereby producing a hole in the earnings distribution just below the notch point and excess bunching in the earnings distribution just at the notch point.¹⁷ Assuming a positive uncompensated labor supply elasticity, workers who are sufficiently close to z^* to begin with (the interval $(z', z^*]$ in the figure) move *strictly above* the threshold, whereas those in the interval $[z^* - \Delta z^*, z']$ move exactly to the threshold and create sharp bunching. Furthermore, for foreign workers above the threshold, the uncompensated net-of-tax wage increase stimulates labor supply and earnings assuming the uncompensated labor supply elasticity is positive.¹⁸ Conversely, when the

17. As analyzed by Kleven and Waseem (2013), the hole may not be very sharp in the presence of optimization frictions, such as imperfect information and adjustment costs, but it remains the case that any excess mass above the notch point should be accompanied by an equal amount of missing mass below the notch point.

18. This assumption is consistent with a large number of empirical labor supply studies across many countries (see, e.g., Blundell and MaCurdy 1999). In particular, Kleven and Schultz (2012) find positive uncompensated labor earnings elasticities throughout the Danish income distribution (including the top decile), suggesting that the labor supply curve does not bend backward for high-income people in Denmark.

scheme elapses after its three-year duration, the net-of-tax wage rate decreases and we should observe a decrease in labor supply and earnings of immigrants who stay in Denmark beyond three years.

Some of those responses could take place through tax avoidance rather than pure labor supply without affecting the theory and predictions as in Feldstein (1999).¹⁹ Such avoidance effects simply magnify the behavioral responses of the competitive labor supply model we have described. Related to tax avoidance, there might also be intertemporal substitution in earnings to take advantage of the scheme before it expires. If the employee and employer agree to continue the work contract after the three-year scheme duration, it will be advantageous for them to increase pay while the worker is still in the scheme and reduce pay when the worker is no longer in the scheme and faces the regular income tax. Hence, we should observe excess pay during the last year a worker is on the scheme and depressed pay during the first year off the scheme. Naturally, as employees can always choose to leave a job, such agreements cannot be formally enforced, and this may limit the scope for such intertemporal substitution.²⁰ Nevertheless, such intertemporal substitution in earnings is in fact observed in the data, as we shall see.

III.B. Matching Frictions and Wage Bargaining Model

As we shall see, our empirical findings are not in line some of the predictions of the basic competitive labor supply model. Hence, we consider an alternative model with matching frictions and wage bargaining that can rationalize all of our empirical findings and which nests the competitive model. To simplify, we

19. For example, employees can give up some nontaxable benefits or on-the-job amenities (such as better offices, flexibility in the work schedule) to get higher taxable pay and qualify for the scheme. As mentioned, most perquisites such as company cars, mobile phone bills, or company-provided lodging are by law included in taxable earnings for the scheme purposes. It is very unlikely that the large and sophisticated firms hiring those highly skilled employees would engage in outright tax evasion, for example, colluding with the employee to fake earnings to meet the eligibility threshold.

20. Such intertemporal substitution could also be due to real labor supply changes, for example, workers doing overtime at the end of the scheme and reducing hours after the scheme elapses, in which case no commitment or enforcement is required.

focus on the case with inelastic labor supply (conditional on migration).

A worker contemplating migration to Denmark has marginal product y for a prospective employer in Denmark. In the standard competitive model used here, the pay offered by the employer is equal to y , and in the absence of labor supply responses this level of pay is not affected by the scheme tax rate (abstracting from general equilibrium effects on wages as discussed earlier). In a model with matching frictions, prospective immigrant workers and employers in Denmark need to expend resources to create a match as in the search-and-matching framework of Diamond-Mortensen-Pissarides. The worker has a pretax reservation wage equal to y_0 such that, conditional on a match, the worker is willing to migrate to Denmark and work for the employer if she is paid at least y_0 . Conditional on a match, the employer values the worker at her marginal product y and is therefore willing to pay up to y . If $y_0 \leq y$, any wage $z \in [y_0, y]$ will be acceptable to both the worker and the firm. In such search models, the wage z is therefore not determined and can be set anywhere in the acceptable band $[y_0, y]$ (Howitt and McAfee 1987; Hall 2005). If $y_0 > y$, then no wage can satisfy both the employee and the employer and the match cannot proceed.

1. Migration. A worker migrates if and only if $y_0 \leq y$, that is, when there exists a band of wages for which the move is mutually beneficial to the worker and the employer. The width of this acceptable wage band depends on the tax system in Denmark and in the home country. Denoting by z_h the earnings of the worker in her home country, by τ_h the average tax rate in her home country, and by ν the net cost of migration (the moving cost plus the differential value of living in her home country vs. Denmark), the worker needs to be paid net-of-tax earnings of at least $z_h \cdot (1 - \tau_h) + \nu$ to be willing to make the move to Denmark. Denoting by τ the average tax rate in Denmark, the pretax reservation wage y_0 must satisfy $y_0 \cdot (1 - \tau) = z_h \cdot (1 - \tau_h) + \nu$, that is,

$$y_0 = \frac{z_h \cdot (1 - \tau_h) + \nu}{1 - \tau} = \frac{y_0^{\tau=0}}{1 - \tau},$$

where $y_0^{\tau=0} \equiv z_h \cdot (1 - \tau_h) + \nu$ is the reservation wage with zero taxes on earnings in Denmark.

The scheme lowers the tax rate in Denmark from the regular rate τ^D to the scheme rate τ^S for migrants paid above the eligibility threshold z^* . Hence, there are two reservation wages y_0^S, y_0^D based on whether the worker is eligible for the scheme:

$$y_0^S = \frac{z_h \cdot (1 - \tau_h) + v}{1 - \tau^S} < \frac{z_h \cdot (1 - \tau_h) + v}{1 - \tau^D} = y_0^D.$$

Workers who can use the scheme have a lower pretax reservation wage. Hence, the scheme widens the band $[y_0, y]$ of acceptable wages and induces migration when $y_0^S \leq y < y_0^D$.

2. Duration. The effect of the scheme on duration of stay is qualitatively similar to the competitive labor supply model. Conditional on migration, the scheme creates an incentive for longer durations up to three years and produces excess bunching at the three-year threshold as shown in Figure I, Panel A.²¹ As described, with idiosyncratic time variation in location preferences (here captured by v) featuring preferences for short spells by some migrants and/or spells in full-year increments, there may be bunching at every yearly threshold and tax-driven bunching at three years may be relatively small. Furthermore, such preferences imply that scheme-induced migrants may choose durations strictly less than the three-year scheme period, which produces excess mass everywhere between zero and three years in the postscheme duration distribution.

3. Wage Determination. Within the band $[y_0, y]$ of acceptable wages, how is the equilibrium wage determined? Although there are many potential models of wage determination in this type of setting, the most widely used model assumes that the pretax wage z splits the surplus between the worker and the firm through a Nash bargaining process in which an exogenous parameter $0 \leq \beta \leq 1$ captures bargaining power of the worker (and $1 - \beta$ captures bargaining power of the employer). Nash bargaining is particularly useful for solving the model in the presence of *discontinuous* incentive schemes (which arise here because of the scheme eligibility threshold).

21. The depiction in Online Appendix Figure 1A corresponds to the case where the preference parameter v (net utility cost of living in Denmark) is smoothly increasing in total time spent in Denmark.

Formally, given the tax rate τ , pay z is set to maximize $W = (y - z)^{1-\beta}((1 - \tau)z - y_0^{\tau=0})^\beta$ where $y - z$ is the firm's surplus and $(1 - \tau)z - y_0^{\tau=0}$ is the worker's surplus. Ignoring first the notch (i.e., the discrete jump in τ) at threshold z^* , the solution to the bargaining problem is

$$(1) \quad z = \beta y + (1 - \beta)y_0 \quad \text{with} \quad 0 \leq \beta \leq 1.$$

Note that this model nests the standard frictionless case when $\beta = 1$. From condition (1), earnings under the scheme tax rate τ^S are equal to $z^S = (1 - \beta)y_0^S + \beta y$, whereas earnings under the regular Danish tax rate τ^D are equal to $z^D = (1 - \beta)y_0^D + \beta y$. Hence, the scheme *reduces* pretax earnings as long as firms have some bargaining power so that $1 - \beta > 0$ (i.e., when we are not in the standard competitive model with $\beta = 1$), because the lower reservation wage induced by the preferential tax scheme allows firms to bargain down wages.

Let us now consider the implications of the tax notch at the eligibility threshold z^* . There will be bunching at the threshold as in the competitive labor supply model, but a conceptual difference is that the matching model predicts that bunchers may be coming from both the low-tax and high-tax sides of the threshold. This can be understood as follows. First, consider workers with earnings above the threshold z^* in the absence of the scheme, that is, workers for whom $z^D > z^*$. As shown above, the introduction of the preferential scheme rate reduces pretax earnings inside the eligible range when firms have positive bargaining power, so we have $z^S < z^D$ for these workers. However, if earnings were sufficiently close to the threshold to begin with, a situation arises wherein $z^S < z^* < z^D$, which is inconsistent with an interior solution in either tax bracket. Such workers will therefore bunch from above and the amount of this bunching is increasing in the bargaining power of firms $1 - \beta$. Second, consider workers with earnings below the threshold z^* in the absence of the scheme, that is, workers for whom $z^D < z^*$. Even though the tax rate has not changed in this range, the introduction of the notch at z^* may allow these workers to push up their wages provided that they have positive bargaining power β . In particular, for a worker with earnings just below the threshold absent the scheme, a small increase in pay produces a large gain for the worker at a small cost for the firm, and such a pay increase will be the equilibrium outcome under Nash bargaining with positive bargaining power for

workers. Hence, there will also be bunching from below and the amount of this bunching is increasing in the bargaining power of workers β . Finally, extensive migration responses also contribute to bunching as marginal entrants have an excess tendency to locate at z^* . Again, this reflects bunching either from above (those for whom $z^* < z^D < y_0^D$ without the scheme and $z^S = z^* > y_0^S$ with the scheme) or from below (those for whom $z^D < z^*, y_0^D$ without the scheme and $z^S = z^* > y_0^S$ with the scheme).

Importantly, bunching from below (created by $\beta > 0$) is associated with a hole in the earnings distribution below the threshold, while bunching from above (created by $1 - \beta > 0$) is generated by a shift in the entire distribution above the threshold z^* and is therefore not associated with any hole. Hence, for any size of bunching, the size of the hole below the eligibility threshold is informative of the bargaining power of workers β . When $\beta = 1$ (standard frictionless model), all bunching is coming from below and creates a hole on this side of the threshold. When $\beta = 0$, all bunching is coming from above and creates no hole on either side of the threshold.

Figure II illustrates these theoretical results in a density distribution diagram. Panel A shows how bunching and the hole is created by bargaining responses (conditional on migration) from below and above depending on the bargaining power of workers and firms, and Panel B shows how migration responses affect the distribution and adds to bunching. We may summarize the key predictions of the matching frictions model as follows.

PREDICTION 1: *Migration*. All workers for whom $y_0^S < y \leq y_0^D$ and $z^* \leq y$ migrate into the country because of the scheme. This lifts up the density of foreign migrants above z^* . A fraction of those migrants will bunch at z^* .

PREDICTION 2: *Duration*. The scheme produces bunching at the threshold D^* in the duration density. Scheme-induced migrants have an excess propensity for durations between 0 and D^* , which creates excess mass on $(0, D^*)$ in the post-scheme duration density.

PREDICTION 3: *Wages*. Among migrant workers paid above the eligibility threshold z^* , the scheme *reduces* pretax earnings.

PREDICTION 4: *Bunching and Hole*. There is bunching at z^* from above when firms have bargaining power ($1 - \beta > 0$). There is

bunching at z^* from below when workers have bargaining power ($\beta > 0$). There is a hole below z^* only when workers have bargaining power ($\beta > 0$).

IV. EMPIRICAL EVIDENCE

In the following, we estimate the empirical effects of the scheme on migration and duration (Section IV.A) and wages (Section IV.B), using the framework set out before. We show

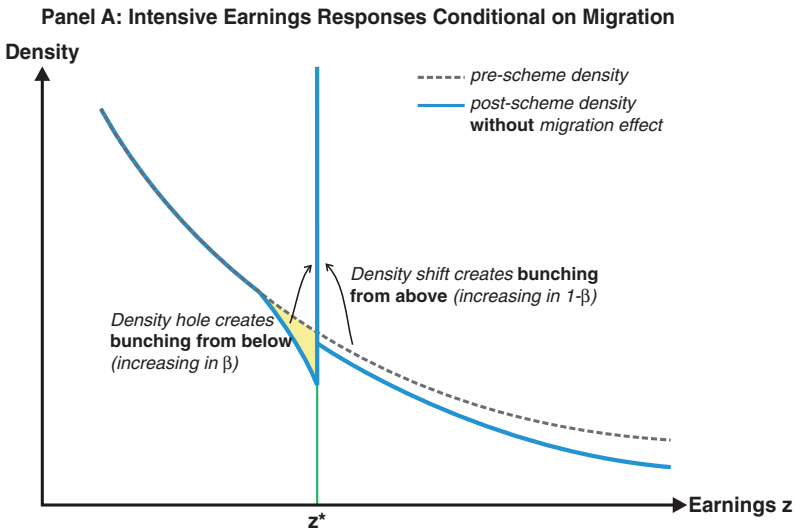


FIGURE II

Responses to the Tax Scheme in the Matching Frictions Model

The figure shows earnings density diagrams depicting intensive earnings responses (Panel A) and extensive migration responses (Panel B) to the scheme in the matching friction model of Section III.B. Absent the scheme, the earnings density of immigrants is smoothly decreasing (short-dashed line). The top panel shows that the scheme reduces earnings for all workers above the eligibility threshold z^* if employers have positive bargaining power ($1 - \beta > 0$). This shifts the full density distribution above z^* to the left (to the solid line) and creates bunching at the threshold from above. This panel also shows that there will be bunching from below if workers have positive bargaining power ($\beta > 0$), creating a hole in the density distribution just below z^* . The bottom panel shows that the scheme induces migration, which shifts the density distribution upward above the eligibility threshold z^* and adds to bunching at z^* .

(continued)

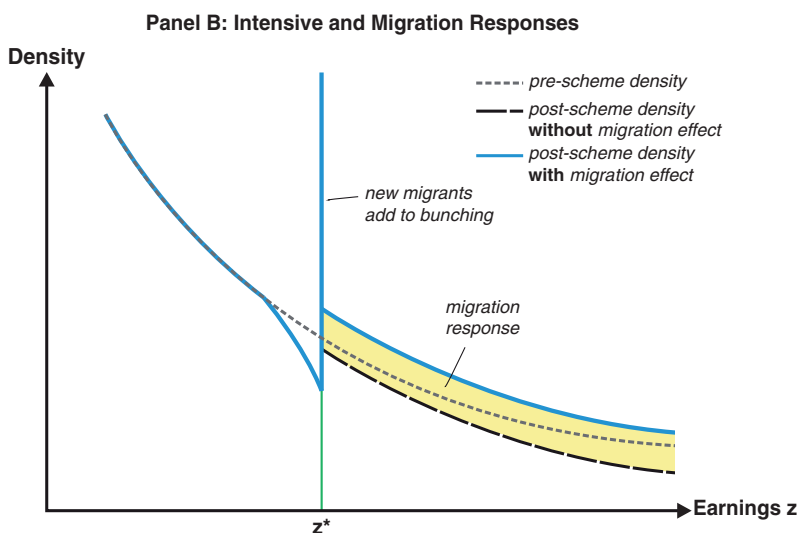


FIGURE II

Continued

that each of the four predictions of the matching frictions model are borne out by the data.

IV.A. Migration Effects

1. Differences-in-Differences Approach. As a first step in testing whether the Danish tax scheme had an effect on high-skilled migration, we consider the evolution of the number of foreigners with earnings above the scheme threshold between 1980 and 2005 in Figure III, Panel A.²² This series (labeled “treatment” in the figure) shows that the number of highly paid foreigners was fairly stable around 800 in the prescheme period from 1980 to 1990. After the scheme is introduced in 1991, demarcated by a vertical line in the figure, there is a steady increase in the number of highly paid foreigners. The number reaches 2,000 by 1997 and is close to 3,000 by 2005. Naturally, the number of highly paid foreigners could have increased even in the absence of the scheme. For example, European Union labor market integration following the Single European Act (taking effect from 1987) and

22. Earnings are annualized based on duration of stay in the year for part-year residents. Duration of stay in the year is measured using the migration database.

the Maastricht Treaty (taking effect from 1993) could have increased labor mobility across European countries. The simplest way to control for such trends is to plot the number of highly paid foreigners just below the eligibility threshold for the scheme. Hence, Panel A also shows the number of foreigners in Denmark with earnings between 80% and 90% of the threshold

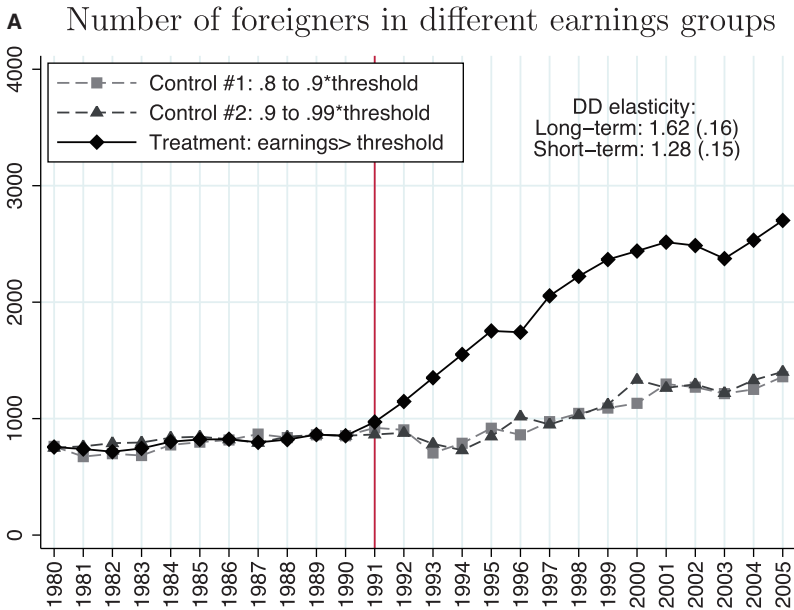


FIGURE III

Migration Effects of the Tax Scheme: Difference-in-Differences Evidence

Panel A reports the number of foreigners with earnings above the scheme eligibility threshold (treatment series) from 1980 to 2005. As control groups, it reports the number of foreigners in Denmark with earnings between 80% and 90% of the threshold (control 1) and with earnings between 90% and 99.5% of the threshold (control 2). Both control series are normalized so that they match the treatment series in 1990—the year before the scheme was first implemented. The vertical line at year 1991 denotes the year the scheme was first implemented (the scheme was enacted in 1992 and applied retrospectively to all contracts starting after June 1, 1991). All numbers are weighted by duration of stay during the year for part-year foreign residents. Earnings are also annualized for part-year residents. Panel B repeats the same series with the flow number of arrivals of foreigners in each year instead of the total stock number of foreigners (earnings of immigrants are annualized to classify them into earnings groups). The corresponding elasticity estimates are reported (see Table II).

(continued)

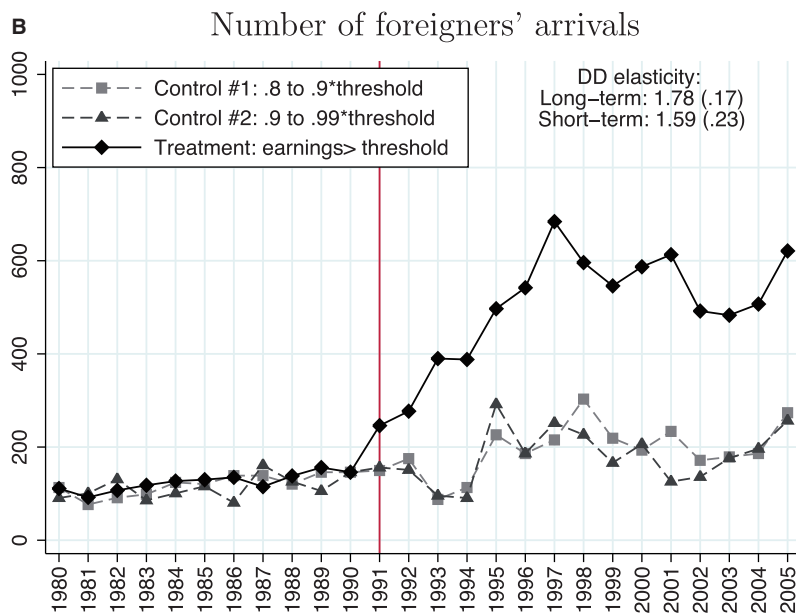


FIGURE III
Continued

(control 1) and with earnings between 90% and 99.5% of the threshold (control 2). Both series are normalized so that they match the treatment series in 1990, the year before the scheme came into force. Before we address potential confounders to this simple differences-in-differences strategy, two lessons emerge from the use of these controls.

First, the control series follow the treatment series extremely closely in the period before the scheme is introduced. The remarkable similarity and stability of the three series lend credibility to our assumption that foreigners just below the threshold are good control groups for the treated foreigners above the threshold. Second, after the scheme is implemented, the control groups series only increase modestly in the first five years. By 1995, the control series are virtually identical to 1990 levels while the treatment series have almost doubled. After 1995, the control series increase steadily over time at about the same rate as the treatment series. Indeed, after 1995, the treatment series are consistently about twice as high as the control series.

Next, we zoom in on the flow of arrivals of highly paid foreigners in Denmark (instead of focusing on the stock). Figure III, Panel B reports the number of foreigners with annualized earnings above the scheme eligibility threshold (treatment series) *arriving* each year in Denmark from 1980 to 2006. As control groups, we again consider the number of foreigners arriving in Denmark with annualized earnings between 80% and 90% of the threshold (control 1) and with earnings between 90% and 99.5% of the threshold (control 2).²³ This panel is consistent with the picture provided by the previous panel for the stock of foreigners. It shows that the number of arrivals of foreigners above the threshold relative to foreigners below the threshold more than doubles quickly after the scheme is put in place.

Table II summarizes the graphical evidence described above by presenting elasticity estimates. The three columns consider different migration elasticity concepts (all defined with respect to the average net-of-tax rate): (i) the elasticity of the total number of foreigners (as in Panel A of Figure III), (ii) the elasticity of the number of arrivals of foreigners (as in Panel B of Figure III), and (iii) the elasticity of the number of foreigners with less than three years of presence in Denmark (as the scheme applies only for a duration of three years). These elasticities are estimated using a two-stage least squares (2SLS) regression specification of the form

$$(2) \quad \log N_{gt} = \alpha_0 + \beta \cdot \mathbb{1}[g = 1] + \gamma_t + e \cdot \log(1 - \tau_{gt}) + \nu_{gt},$$

where $g=0,1$ denotes control and treatment group, t denotes year, N_{gt} is the number of foreigners in group i and year t (corresponding to each of the outcomes in columns (1)–(3)), τ_{gt} is the average tax rate in group g and year t , $\mathbb{1}[g = 1]$ is the treatment group dummy, and γ_t are year fixed effects. The key variable of interest $\log(1 - \tau_{gt})$ is instrumented by the interaction $\mathbb{1}[g = 1] \cdot \mathbb{1}[t > 1991]$. As a baseline, we compute τ_{gt} assuming a 100% take-up rate in which case the results should be interpreted as intent-to-treat effects.

The treatment group is defined as foreigners with earnings above the eligibility threshold, and the control group is defined as foreigners with earnings between 80% and 99% of the eligibility threshold. Effectively, the elasticity estimate e is the Wald ratio

23. Again, both control series are normalized so that they match the treatment series in 1990.

TABLE II
MIGRATION ELASTICITY ESTIMATES

	(1)	(2)	(3)
	Total no. foreigners	No. arrivals	No. foreigners with less than 3 years of presence
Panel A: Treatment: earnings above threshold, control: earnings between 80% and 99% of threshold			
A1. Baseline			
ε_{lt} (long-term)	1.625*** (0.162)	1.779*** (0.168)	2.049*** (0.148)
ε_{st} (short-term)	1.280*** (0.151)	1.590*** (0.228)	1.756*** (0.170)
A2. Control for pre-existing trends			
ε_{lt} (long-term)	1.756*** (0.176)	1.771*** (0.168)	2.152*** (0.158)
A3. Placebo			
ε_{lt} (long-term)	-0.0602 (0.0823)	-0.0101 (0.245)	0.0796 (0.161)
A4. Control for imperfect take-up (IV)			
ε_{lt} (long-term)	2.892*** (0.232)	1.945*** (0.167)	2.392*** (0.138)
A5. Nordic countries			
ε_{lt} (long-term)	1.442*** (0.166)	1.805*** (0.287)	2.208*** (0.257)
A6. English-speaking countries			
ε_{lt} (long-term)	1.852*** (0.222)	2.186*** (0.246)	2.281*** (0.206)
A7. Danish expatriates			
ε_{lt} (long-term)	0.0185 (0.0280)	-0.0913 (0.0708)	-0.0998 (0.0613)
Panel B: Treatment: percentile 99.5-100, Control: percentile 95-99			
B1. Baseline			
ε_{lt} (long-term)	1.133*** (0.0802)	1.015*** (0.169)	1.257*** (0.109)

Notes. Robust standard errors in parentheses, * $p < .05$, ** $p < .01$, *** $p < .001$. The table displays elasticity estimates based on equation (2). Number of arrivals is the number of foreign individuals entering Denmark in a given year. Number of foreigners with less than three years of presence are foreigners who are eligible for the scheme based on all rules except the income threshold rule. The long-term (short-term) elasticity refers to a specification that includes years 1992–2005 (1992–1996) as the postreform period. We exclude 1991 from the specification, because the reform was enacted in 1992 but applied retroactively starting in mid-1991. Panel A displays estimates where the control group is defined as foreigners with (annualized) earnings between 80% and 99% of the threshold and the treatment group is foreigners with (annualized) earnings above the eligibility threshold. Panel A1 is the baseline estimate. Panel A2 controls for differential preexisting trends specific to the control and treatment groups. Panel A3 is a placebo where the control group is foreigners with earnings between 80% and 90% of the threshold while the treatment group is foreigners with earnings between 90% and 99% of the threshold (we assume that the scheme tax rate applies to the treatment group when estimating the elasticity). Panel A4 controls for imperfect take-up, instrumenting the actual average tax rate (given actual take-up) by the intention-to-treat average tax rate. Panels A5 and A6 break down the elasticity by countries of citizenship among foreigners. Nordic countries=Finland, Iceland, Norway, Sweden. English-speaking countries=Australia, Canada, Ireland, New Zealand, South Africa, United Kingdom, United States. Panel A7 looks at the behavioral response of Danish expatriates (also eligible for the scheme). In Panel B, the control group is all foreigners with (annualized) earnings between the 95th and 99th percentiles of the earnings distribution of natives, and the treatment group is all foreigners with (annualized) earnings above the 99.5th percentile.

of the differences-in-differences of the log number of foreigners to the differences-in-differences of the log net-of-tax rate. We always exclude the year 1991 from the regression, because the reform was enacted in 1992 but applied retroactively starting in mid-1991. We consider two time horizons for the migration response. The long-term elasticity refers to a specification that includes 1992–2005 as the postreform period, while the short-term elasticity includes only 1992–1996 as the postreform period. All specifications include 1980–1990 as the prereform period.

Our baseline estimates in Panel A1 of Table II show that elasticities are large and precisely estimated, between 1.5 and 2 across the different elasticity definitions. The large magnitude of elasticities can be understood directly from Figure III: the scheme slightly more than doubles the number of highly paid foreigners while increasing the average net-of-tax rate from about 0.4 to about 0.7, which translates into an elasticity of $\frac{\log(2.2)}{\log(0.7/0.4)} \simeq 1.5$. The short-term elasticities are somewhat smaller than the long-term elasticities as the migration effect builds gradually after the reform. However, in the case of the number of arrivals, short-term and long-term elasticities are extremely close, suggesting that the response to the scheme was fast. Naturally, the elasticity of the number of foreigners with less than three years of presence is larger (close to 2) than for all foreigners as the scheme targets foreigners during their first three years of stay, an important point to which we return later.

The following panels consider various robustness checks. Panel A2 controls for a potential difference in preexisting trends between the treatment and control groups where we first regress $\log N_{gt}$ for all years prior to the reform on group fixed effects and two group specific time trends and then use the residuals as the outcome in the regression specification (2). The elasticity estimates are virtually unchanged compared to the baseline.

Panel A3 presents a placebo specification, where the treatment (control) group is defined as foreigners with income between 90% and 99% (80% and 90%) of the threshold. The net-of-tax rate variable—the denominator of the elasticity—remains the same as in Panel A1. This specification also tests for shifting around the eligibility threshold (via earnings or avoidance responses) to qualify for the scheme, because such shifting would produce a dip in the number of foreigners just below the threshold relative to the number of foreigners further down. The elasticities are

small (0.1 or less) and insignificant, which confirms the graphical evidence.

Panel A4 controls for imperfect take-up using actual tax rates (given actual take-up) and instrumenting the actual tax rates by the intent-to-treat tax rates. This correction has a small impact on the estimates for the number of arrivals and the number of foreigners in their first three years of stay (columns (2) and (3)), reflecting that scheme take-up rate is high as discussed earlier (around 80–85%). It has a bigger effect on the estimate for the total number of foreigners in column (1) because foreigners who stay beyond three years are no longer eligible for the scheme.

Panels A5 and A6 estimate the effects of the scheme by country of citizenship for two groups: foreigners coming from Nordic countries (Finland, Iceland, Norway, Sweden) in Panel A5, and foreigners coming from English-speaking countries (Australia, Canada, Ireland, New Zealand, South Africa, United Kingdom, United States) in Panel A6. Elasticity estimates are slightly larger for English-speaking countries, especially for the total stock of foreigners in column (1) due to the fact that the stock of immigrants from those countries was considerably smaller to begin with.

Finally, Panel A7 reports estimates for expatriates, that is, Danish citizens returning to Denmark after a spell of at least three years abroad (so that they qualify for the scheme on their return). Panel A7 shows small and insignificant estimates. This implies that the scheme was not successful at bringing back highly paid Danish expatriates to their home country. This result is very important for the policy debate on taxes and mobility as we discuss later.

2. Potential Confounders. Our simple graphical differences-in-differences analysis relies on the standard parallel trend assumption. That is, absent the scheme, the trend in the number of foreigners above the threshold would have been parallel to the trend in the number of foreigners slightly below the threshold. There are two potential confounders: (i) a fanning-out of the earnings distribution after 1990, and (ii) an endogenous earnings response to the scheme threshold (notch) as analyzed in the conceptual framework of Section III. Let us address them in turn.

Confounder 1: Fanning out of the earnings distribution. A fanning out of the earnings distribution would increase the

number of workers (natives and foreigners) paid above the eligibility threshold, thereby creating a divergence in the number of foreigners in the treatment and control groups even without the scheme. Online Appendix Figure A2 addresses this issue by plotting the fraction of foreigners in different percentiles of the earnings distribution. Since the threshold for scheme eligibility is always between the 99.2th and the 99.4th percentile of the full earnings distribution among Danish adults with positive earnings, Figure A2 compares the fraction of foreigners in the 99.5–100th percentile (treatment) to the fraction of foreigners in the 95–97th percentile (control 1) and the 97–99th percentile (control 2).²⁴ The figure shows that the fraction of foreigners in each percentile group is extremely stable before 1991. After 1991, the fraction of foreigners increases much more rapidly—in absolute as well as percentage terms—above the 99.5th percentile, where the scheme applies. Consistent with Figure III, there is a doubling of the fraction foreigners above the 99.5th percentile relative to percentile groups just below the scheme eligibility threshold. This graphical result is confirmed in Table II, Panel B, where we define the treatment (control) group as individuals with earnings above the 99.5th percentile (between the 95th and 99th percentiles). The elasticity estimates are slightly attenuated relative to our baseline specification of Panel A1, but remain very large around 1.2.

Confounder 2: Intensive earnings response to the scheme. The second confounder is that foreigners above the eligibility threshold might be displacing foreigners slightly below the threshold through intensive earnings responses as we described in the theory section. Such shifting should produce a dip in the number of foreigners just below the threshold relative to the number of foreigners further down. The completely parallel trends of the two different control groups in Figure III (those between 90% and 99% of the threshold and those between 80% and 90% of the threshold) along with the placebo estimates in Panel A3 of Table II suggest that this dip effect was not significant. To cast further light on this and understand the nature of the behavioral response, it is fruitful to look directly at the density of earnings among foreigners.

Figure IV plots such densities before the scheme was introduced (1980–1990 in dashed gray) and after the scheme was

24. Years 1991–1994 are omitted due to lack of earnings data for these years.

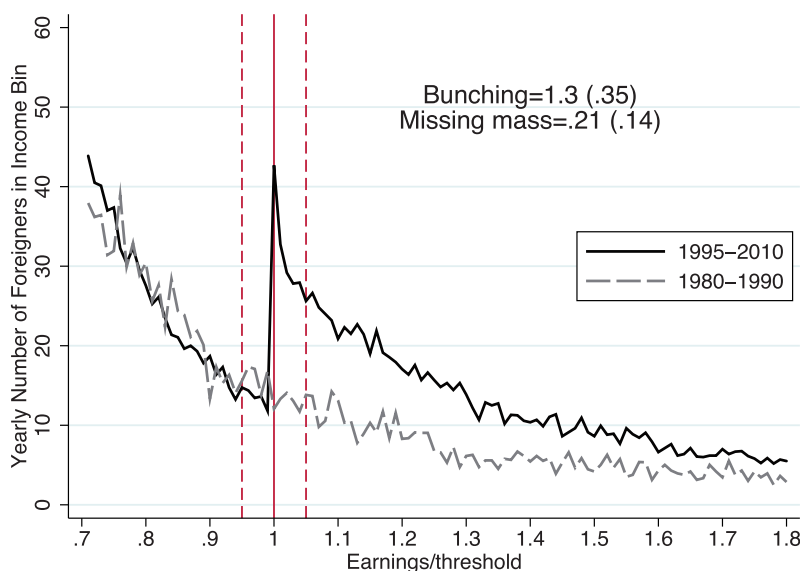


FIGURE IV

Earnings Density for Foreigners before and after Scheme Introduction

The figure reports the density of the earnings distribution of foreigners around the eligibility threshold (denoted by the vertical line) in 1995–2010 after scheme implementation (solid dark line) and in 1980–1990 before scheme implementation (dashed gray line). The sample is restricted to individuals in their first and second full calendar year of presence in Denmark (to avoid having to correct for part-year earnings or part-year scheme eligibility). The 1980–1990 density is reweighed so that it matches the 1995–2010 density on average between 70% and 90% of the scheme eligibility threshold. The graph shows that the scheme almost doubled the density above the threshold due to extensive migration responses and also created bunching at the eligibility threshold due to an intensive margin earnings response. There is no evidence of a hole in the density of earnings below the scheme eligibility threshold. Nonparametric estimates of excess bunching at the threshold and missing mass below the eligibility threshold are reported. They are estimated using the method of Chetty et al. (2011) by fitting polynomials for the densities of the left and right of the vertical dashed lines (see Online Appendix A.1 for details of the estimation method).

introduced (1995–2010 in solid black). Earnings are measured in proportion to the eligibility threshold such that 1 corresponds to the threshold, demarcated by a solid vertical line. The post-scheme density is normalized so that the average level of the density between 70% and 90% of the threshold is the same as in the prescheme period. The figure limits the sample to foreigners in their first and second full calendar years in Denmark. This is

done to avoid using years where the person is either a part-year resident (year of arrival) or a part-year scheme beneficiary (as the scheme elapses at some point during the third full calendar year in Denmark), because annualizing earnings for such observations introduces noise.

The density is smooth around the threshold before the introduction of the scheme. After the scheme is introduced, the density is virtually identical below the threshold, but two differences appear above the threshold. First, the density is everywhere higher above the threshold confirming the strong migration response in Figure III and showing that this response occurs at all earnings levels above the threshold. Second, there is clear bunching just above the threshold (notch point), but no discernible hole below the threshold. The excess mass due to bunching is statistically significant while the missing mass on the left of the threshold is not. Excess bunching and missing mass is estimated following the method developed by Chetty et al. (2011) and adapting it to a differences-in-differences setting to take advantage of the counterfactual distribution before the introduction of the tax scheme (see Online Appendix for details).

The presence of bunching is consistent with both the competitive labor supply model and the matching friction model already presented. Even though the bunching is clearly visible in the figure and therefore provides compelling evidence of an intensive earnings response, it is in reality very modest when compared to the extremely large notch in the budget set created by the scheme (Figure I, Panel B). Using the method developed by Kleven and Waseem (2013), the implied labor supply elasticity in the competitive model of Section III.A would be extremely small, less than 0.01 (in the case where there are no frictions due to imperfect information or costly labor supply adjustment). The fact that no hole or missing mass is discernible below the notch is inconsistent with the competitive labor supply model (where bunching is coming from below), but is consistent with the matching frictions model when employers have most of the bargaining power (in which case bunching is coming from above). It is important to note though that our ability to detect a hole is limited, because such holes are not as visible as bunching spikes in a world with optimization frictions (see Kleven and Waseem 2013 for an analysis of how frictions affect both bunching and holes in the competitive labor supply model).

3. *Duration.* We analyze next the effect of the scheme on the duration of stay for immigrants by plotting the duration densities for foreigners (below and above the earnings eligibility threshold; before and after the scheme enactment) in Figure V. Panel A focuses on the prereform period 1980–1990 and compares duration densities for foreigners just below the earnings threshold (96–99th percentile) to foreigners above the threshold (99.5–100th percentile). Vertical lines demarcate year thresholds, with the solid vertical line representing the three-year threshold where the scheme elapses. In Panel A, the P99.5–100 series are normalized to be equal to the P96–99 series on average so that both series are comparable in levels. This placebo panel shows no noticeable difference between duration distributions below and above the scheme threshold prior to the introduction of the scheme.

Panel B focuses on the postreform period 1991–2006, but is otherwise constructed as the top panel. In Panel B, the P96–99 series are normalized to be equal to the P96–99 series from Panel A on average (so that series P96–99 are comparable across Panels A and B). In Panel B, the P99.5–100 are normalized twice using the product of the Panel A normalization factor for P99.5–100 and the Panel B normalization factor for P96–99. As a result, the excess density of P99.5–100 relative to P96–99 in Panel B can be interpreted as the extensive migration response. Two clear changes emerge after the introduction of the scheme. First, there is a jump in the duration density for the treatment group in the interval below three years compared to the interval above three years, confirming that the scheme encourages durations of at most three years. Using a differences-in-differences specification, the scheme has reduced the probability of staying more than three years in Denmark by about 15 percentage points (among all migrants). Second, large and sharp bunching emerges in the duration density precisely at the three-year threshold consistent with the conceptual framework. Interestingly, bunching also emerges at the one-year and two-year thresholds (and to a small extent at the four-year and five-year thresholds), which shows that scheme foreigners tend to negotiate work contracts in full years. Excess bunching at the three-year threshold is larger than excess bunching at all the other year thresholds as one would expect. Using the rounders method of Kleven and Waseem (2013), we show on Panel B that excess bunching at year 3 is significantly larger than excess bunching at years 1

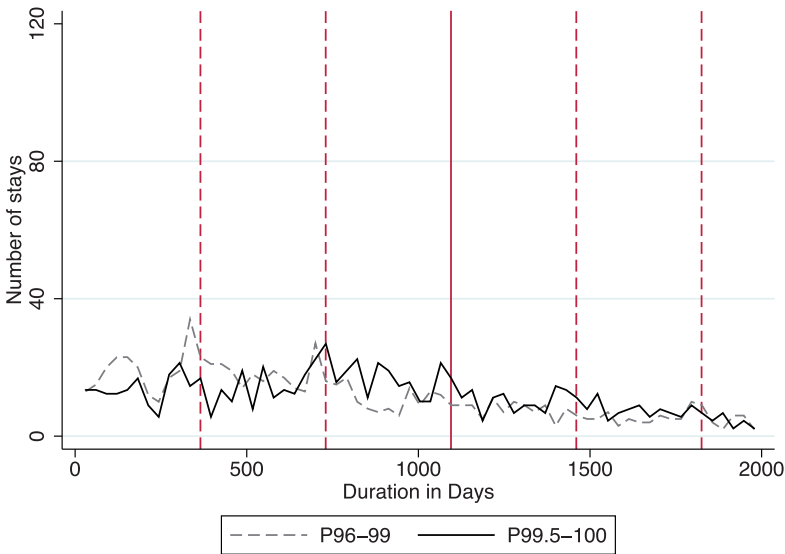
A Before introduction of the scheme, 1980-1990

FIGURE V

Density of the Duration of Stay of Foreigners

The figure reports the density distribution of stays by duration of foreigners with (annualized) earnings above percentile 99.5 (P99.5–100) in solid dark line and earnings between percentile 96 and percentile 99 (P96–99) in dashed gray line. The P99.5–100 group is always above the eligibility threshold for the scheme and the P96–99 group is always below the eligibility threshold for the scheme. Panel A is for years 1980–1990 (before the scheme was implemented) and Panel B is for years 1991–2006 (after the scheme was implemented). Vertical lines demarcate year thresholds, with the solid vertical line representing the three-year threshold where the scheme elapses. In Panel A, the P99.5–100 series are normalized to be equal to the P96–99 series on average (so that both series are the same in levels). In Panel B, the P96–99 series are normalized to be equal to the P96–99 series from Panel A on average (so that series P96–99 are the same in levels across Panels A and B). In Panel B, the P99.5–100 series are then doubly normalized using both the Panel A normalization factor for P99.5–100 and the Panel B normalization factor for P96–99 (so that the excess density of P99.5–100 relative to P96–99 in Panel B can be interpreted as the extensive migration response). Both panels use the same y-axis scale for direct comparison purposes. In Panel B, we show that (a) the difference-in-differences estimate of the probability of staying more than three years falls for P99.5–100 after scheme enactment, and (b) the excess bunching estimate at year 3 using the rounders method of Kleven and Waseem (2013) is positive (see Online Appendix A.2 for details).

(continued)

B After introduction of the scheme, 1991-2006

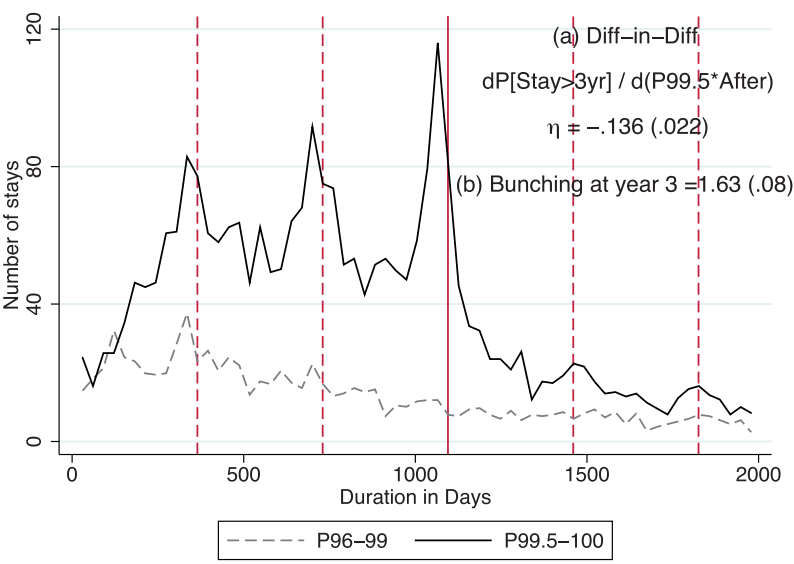


FIGURE V
Continued

and 2, which is evidence of an intensive response along the duration margin (conditional on migration).

4. Tax Policy Implications. Using the elasticities estimated in Table II, the revenue-maximizing tax rate (Laffer rate) can be computed using the classic inverse elasticity formula, $\tau = \frac{1}{(1+\epsilon)}$.²⁵ Note that this is the revenue-maximizing tax rate for a special tax scheme applying solely to foreign immigrants. Using an elasticity of 1.5, the revenue-maximizing tax rate equals $\tau = \frac{1}{(1+1.5)} = 40\%$, which is not very far above the current total tax rate of about 30% when including scheme and payroll taxes. Assuming an elasticity of 2 (as suggested by specifications that control for incomplete take-up), the revenue-maximizing tax rate is only

25. This formula applies in the standard model with no matching frictions. With matching frictions, employers benefit from the tax scheme and obtain higher profits. The same inverse elasticity formula applies in this case as well if we assume that firms' marginal profits are taxed at rate τ as well.

$\tau = \frac{1}{(1+2)} = 33\%$, about the level of the current total tax rate under the scheme. Foreigners pay additional taxes in Denmark when they consume their income in Denmark through the value-added tax (VAT) and various other commodity taxes.²⁶ Taking into account consumption taxes, the Laffer rate on foreigners could be slightly *lower* than the current total tax rate under the scheme.

Importantly, the revenue-maximizing tax rate on natives would be much higher, because the elasticity of the number of natives with respect to the net-of-tax rate is much lower. As shown in Table II, Panel A7, the elasticity for Danish expatriates is close to 0. This implies that the migration effect of changing the top tax rate for all Danish residents would be small.²⁷

IV.B. Wage Effects

We now turn to the effect of the preferential scheme tax for foreigners on the wage earnings of beneficiaries. Recall that the competitive labor supply model and the matching friction model make opposite predictions on the effect of the tax scheme on earnings.

1. Repeated Cross-Section Evidence. Figure VI depicts the average *real* annual earnings for foreigners in their first two full calendar years in Denmark for the sample of foreigners with earnings between 80% and 95% of the scheme eligibility threshold (dashed line) and above 105% of the scheme eligibility threshold (solid line). We exclude those earning between 95% and 105% of the scheme eligibility threshold as bunching at the eligibility threshold naturally biases downward average earnings. For year t , the sample includes foreigners who have arrived in Denmark in year $t - 1$ or $t - 2$, were not residents in the three

26. The VAT normal rate in Denmark is equal to 25%, and on top of that come substantial excises on certain goods. According to Immervoll et al. (2007), the total average consumption tax rate in Denmark is 36%. To the extent that top foreign earners consume a fraction of their Danish income outside Denmark, this rate would have to be scaled accordingly.

27. See Piketty and Saez (2013) for a detailed exposition. Our scheme experiment allows us to measure how many Danish expatriates come back because of the scheme which is only one side of the natives behavioral response. The other side, which is how many Danish people would stay in Denmark (instead of migrating away) if the Danish top tax rate were lowered cannot be directly estimated with our scheme-induced tax variation.

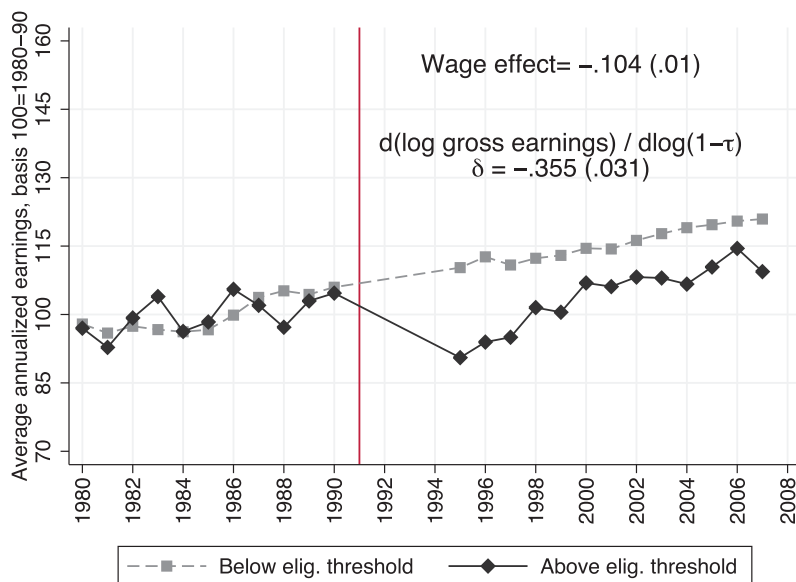


FIGURE VI

Effects of the Tax Scheme on Pretax Earnings: Repeated Cross-Section Evidence

The figure depicts the average real annual earnings for foreigners in their first two full calendar year in Denmark for the sample of foreigners with earnings between 80% and below 95% of the scheme eligibility threshold (dashed line) and above 105% of the scheme eligibility threshold (solid line). Both series are normalized to 100 on average for years 1980 to 1990, before the scheme starts. We exclude those with earnings between 95% and 105% of the eligibility threshold to abstract from bunching effects. For year t , the sample includes foreigners who arrive during year $t-1$ or $t-2$ were not Danish residents in the three years before arrival, and stay the full calendar year t in Denmark. Data for years 1991–1994 are not available. The graph shows that the pretax earnings of foreigners above the scheme eligibility threshold decrease after the scheme is in place, suggesting that the scheme reduces pretax earnings. This is consistent with the matching friction model.

years before arrival, and stay the full calendar year t in Denmark so that they would be eligible for the scheme based on duration requirements. Both series are normalized to 100 on average for years 1980 to 1990, before the introduction of the scheme. The graph shows that the earnings of foreigners below and above the eligibility threshold follow a parallel trend from 1980 to 1991, before the scheme introduction. The earnings of foreigners above the scheme eligibility threshold decrease sharply relative to foreigners below the eligibility threshold after the scheme is in

place. Hence, paralleling our identification strategy for migration effects presented in Section IV.A, a differences-in-differences estimate based on Figure VI would imply that the scheme reduces pretax earnings. This is consistent with the matching friction model. Online Appendix A1 presents systematic differences-in-differences estimates of the effect of the scheme on log annual pretax earnings, as well as the corresponding Wald elasticity estimates of the effects of the net-of-tax rate on earnings. In our basic specification, we obtain a significantly negative effect of the scheme on earnings of -10.4 log points, which translates into an elasticity $\delta = \frac{d \log z}{d \log(1-\tau)}$ of pretax earnings with respect to the net-of-tax rate of -0.36 (0.03).

A potential confounder for the findings from Figure VI is that the set of foreigners who arrive after the scheme is in place might be different because the scheme induces a very large migration response as we showed in Section IV.A. For example, if scheme-induced immigrants have lower skills and earnings than immigrants who would have come without the scheme (conditional on having earnings above 105% of the eligibility threshold), then the estimates reported in Table A1 would be biased downward.

A simple way to assess the potential selection bias due to migration effects is to consider a group for which migration effects (and hence potential selection bias) are much smaller. As we saw in Section III.A, Danish expatriates display zero scheme-induced migration effects. Therefore, in Online Appendix Figure A3, Panel A, we repeat Figure VI for Danish expatriates. The sample in year t includes Danish citizens who came back to Denmark in year $t - 1$ or $t - 2$, were not Danish residents in the three years before arrival, and stay the full calendar year t in Denmark. This group would be eligible for the scheme based on duration requirements. Note that Figure VI included solely foreign immigrants and not Danish expatriates. Online Appendix Figure A3, Panel B duplicates Figure VI for comparison. The figure for expatriates is naturally noisier than for foreigners due to a smaller sample size, but the pattern is very similar. Expatriate average earnings above the eligibility threshold fall noticeably relative to expatriates earnings below the threshold after the scheme is introduced. Online Appendix Table A1 presents the corresponding estimates and shows that the elasticity of earnings with respect to the net-of-tax rate is -0.30 (0.09) for expatriates, not statistically different from the elasticity of

−0.36 (0.03) for foreigners. This suggests that composition effects may not bias the wage effects estimated using repeated cross-section evidence.

We next turn to an alternative estimation method using a balanced panel of migrants who stay five or more years in Denmark that allows us to fully control for individual fixed effects.

2. Panel Evidence. The scheme elapses after three years, producing a large tax increase among scheme beneficiaries who stay in Denmark more than three years. This allows us to estimate the effects of the scheme on earnings while controlling for individual fixed effects.

To control for the effects of selection into staying in Denmark, we restrict the sample to a balanced panel following exactly the same set of individuals over their first five full calendar years of stay in Denmark. Figure VII depicts the average real annual earnings for foreigners arriving in Denmark after the scheme is in place (from 1995 to 2002) in Panel A and before the scheme is in place (from 1980 to 1991) in Panel B. Year 0 is the year of arrival, year 1 is the first full calendar year in Denmark, and so on. Earnings are normalized by year 2 earnings. The sample includes all foreigners who stay five or more full calendar years in Denmark and have gross earnings in year 1 between 80% and 100% of the scheme eligibility threshold in dashed line (control group not eligible for the scheme) and between 100% and 130% of the scheme eligibility threshold in solid line (treatment group eligible for the scheme after enactment).

Panel A shows that earnings increase in years 3 to 5 (relative to years 1 and 2) for those eligible for the scheme. This implies that the end of the scheme leads to an *increase* in earnings, which is consistent with the matching friction model. Note that year 3 is a transition year when individuals are eligible part-year for the scheme. The spike in earnings in year 3 (relative to years 4 to 5) could be due to retiming of compensation to maximize scheme benefits as discussed in the theory section. Therefore, to eliminate potential inter-temporal shifting from years 4–5 to year 3, the legitimate comparison is between years 1 and 2 versus years 3 to 5. Panel B is a placebo comparison (for entrants before the scheme was introduced) showing that no such differential increase takes place for immigrants arriving before the scheme is in place, confirming that the results from Panel A are due to the scheme.

Formally, we can estimate the differences-in-differences effects of the scheme on earnings using the following reduced-form specification.

$$(3) \quad \log z_{id} = \alpha_i + \beta \cdot 1[z_{i,d=1} \geq z^*] + \gamma_d + \eta \cdot 1[z_{i,d=1} \geq z^*] \cdot 1[d \geq 3] + v_{id},$$

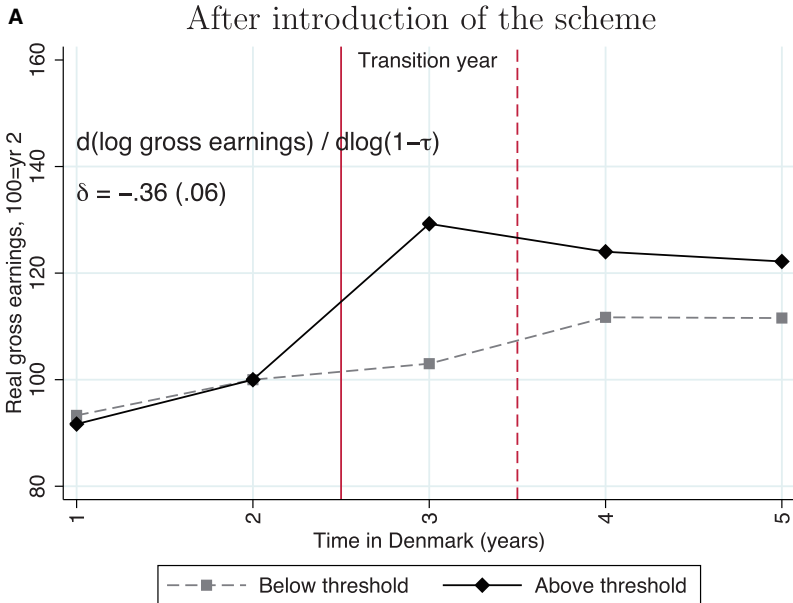


FIGURE VII

Effects of the Tax Scheme on Pretax Earnings: Panel Evidence

The figure depicts the average real annual earnings for foreigners arriving in Denmark after the scheme is in place (from 1995 to 2002) in Panel A and before the scheme is in place (from 1980 to 1991). Year 0 is the year of arrival, year 1 is the first full calendar year in Denmark, and so on. Earnings are normalized by year 2 earnings. The sample includes all foreigners who stay five or more full calendar years in Denmark and have gross earnings in year 1 between 80% and below 100% of the scheme eligibility threshold (dashed line) and between 100% and 130% of the scheme eligibility threshold (solid line). Hence, the sample is a balanced panel. Panel A shows that earnings increase in years 3 to 5 (relative to years 1 and 2) for those eligible for the scheme. This implies that the end of the scheme leads to an increase in earnings, which is consistent with the matching friction model. Year 3 is a transition year when individuals are eligible part-year for the scheme. The spike in earnings in year 3 (relative to year 4) could be due to retiming to maximize scheme benefits. Panel B is a placebo comparison (for entrants before the scheme was introduced).

(continued)

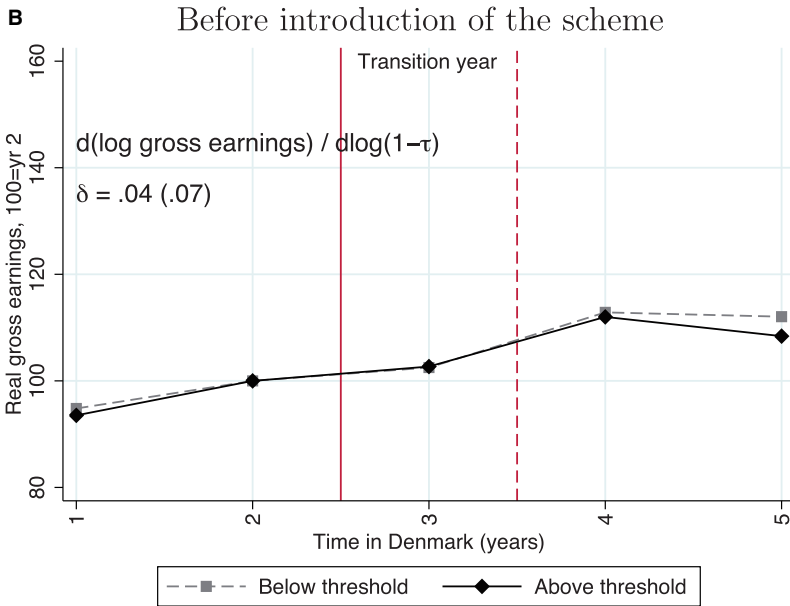


FIGURE VII
Continued

where z_{id} are earnings of individual i in years $d = 1, \dots, 5$, α_i is an individual fixed effect, γ_d is a year fixed effect, and $1[z_{i,d=1} \geq z^*] \cdot 1[d \geq 3]$ is the interaction of having year 1 earnings above the eligibility threshold and the scheme having elapsed. Next, to obtain an elasticity estimate, we consider the following 2SLS fixed-effects specification:

$$(4) \quad \log z_{id} = \alpha_i + \beta \cdot 1[z_{i,d=1} \geq z^*] + \gamma_d + \delta \cdot \log(1 - \tau_{id}) + v_{id},$$

where τ_{id} is the average tax rate. The variable $\log(1 - \tau_{id})$ is instrumented with the interaction $1[z_{i,d=1} \geq z^*] \cdot 1[d \geq 3]$.

The estimates corresponding to specifications (3) and (4) are presented in the first and second rows of Table III, respectively. Consistent with Figure VII, Panel A, column (1) of Table III shows that both the reduced form coefficient η and the elasticity coefficient δ are negative and significant in the case of entrants after 1991 (when the scheme is in place). Note also that the coefficient δ is very similar in magnitude to the coefficients estimated using repeated cross-sections in Online Appendix Table A1.

TABLE III
PANEL ESTIMATES OF THE EFFECTS OF THE TAX SCHEME ON PRETAX EARNINGS

	(1)	(2)	(3)	(4)
	Treatment: Entry after 1991		Placebo: Entry 1980 to 1990	
	OLS (stayers only)	Heckman 2-step	OLS (stayers only)	Heckman 2-step
Reduced-form estimate	0.0955*** (0.0261)	0.0860*** (0.0248)	-0.0100 (0.0190)	0.0415 (0.0218)
Elasticity $\frac{d \log z}{d \log(1-\tau)}$ estimate	-0.356*** (0.061)	-0.319*** (0.092)	0.037 (0.071)	-0.154 (0.081)
First stage: Prob[stay > 3 years]				
Exclusion restriction:		Child birth in the first 3 years in Denmark		Child birth in the first 3 years in Denmark
Birth		0.582*** (0.0675)		0.496*** (0.0876)
<i>N</i>	2,951	6,138	2,470	3,818
λ		0.187 (0.0158)		0.136 (0.0214)
LR test of independence		$\chi^2 = 106.5$ Prob > $\chi^2 = .00$		$\chi^2 = 34.36$ Prob > $\chi^2 = .00$

Notes. Robust standard errors clustered at the individual level in parentheses, * $p < .05$, ** $p < .01$, *** $p < .001$. The table presents panel estimates of the effect of scheme lapse on pre-tax earnings using the reduced form specification (3) in row 1 and the 2SLS specification (4) in row 2. All specifications are fixed-effects models. The reduced-form estimate is the effect on log earnings of being after year 3 and having been eligible for the scheme in the first three years of presence in Denmark. Column (1) is the OLS regression on a balanced panel of stayers only. Column (2) controls for potential selection on the earnings profile and implements a two-step Heckman estimator using the birth of a child in the first three years of presence in Denmark as an exclusion restriction in the selection equation for staying more than three years. λ is the estimated inverse Mills ratio. Columns (3) and (4) repeat the same specifications on a placebo sample of top foreign earners who entered Denmark between 1980 and 1990 and hence were never eligible for the scheme.

Consistent with Figure VII, Panel B, column (3) of Table III shows that both the reduced-form coefficient η and the elasticity coefficient δ are insignificant for the placebo case of entrants before 1990 (when the scheme is not yet in place).²⁸

While the panel analysis controls for individual fixed effects, it is conceivable that individual earnings differ not only in level

28. For the placebo elasticity estimate, we assume that the group above the eligibility threshold would have benefitted from the scheme in their first three years of stay when computing the average tax rates.

but also in profile. Such profile effects cannot be controlled for solely by individual fixed effects. If selection into staying in Denmark is based on earnings profiles (and not only on earnings levels), then treated and control groups in our sample of stayers may have different earnings profiles, potentially creating a bias. For example, suppose individual earnings fully reflect productivity and that productivity varies across years idiosyncratically. Suppose further that individuals leave Denmark when their net-of-tax earnings fall below their net-of-tax reservation wage (i.e., individuals behave in a completely myopic way and consider only current potential earnings when deciding to leave Denmark). Net-of-tax earnings mechanically fall when the scheme elapses due to the tax rate increase. Therefore, workers for whom the scheme elapses are more likely to leave Denmark (relative to ineligible workers) precisely when they experience a fall in earnings. Hence, in that case, the panel treatment sample would be selecting scheme workers who tend to experience pretax earnings increases when the scheme elapses. Such a scenario seems implausible to us for two reasons. First, pay for such high top earners is set by contract in advance of realized productivity. Second, Figure VII shows that the profiles of earnings are very close across the control and treatment groups before the scheme elapses. If the treatment group were selected based on disproportionate increases in productivity at the time the scheme elapses, it seems unlikely that the productivity profile from year 1 to year 2 would be so close between the treatment and control groups.

To provide an additional robustness check, we need an instrument that affects the likelihood to stay in Denmark without affecting the wage profile between years 1–2 versus 3–5. Individuals who have a child during the first three years of stay in Denmark are 50% more likely to stay in Denmark after year 3 than those who do not. Hence, we divide our sample into individuals who experience the birth of a child in their first three years of stay in Denmark and individuals who do not. Online Appendix Figure A4 depicts the reduce-form evidence using the model of Figure VII, Panel B. It shows that the wage effects are very similar across the two groups. This suggests that the increase in wages after the scheme elapses is not due to selection on wage profiles but reflects the causal effect of the change in taxes on earnings.

We estimate in Table III two-step Heckman models where the second stage is the same fixed-effect model as in specification (4), but where we use birth of a child as an exclusion restriction in the selection equation. These estimates are reported in columns (2) and (4) of Table III for postscheme entrants and prescheme entrants, respectively. These Heckman estimates are very close to the initial estimates in columns (1) and (2) confirming the findings of Figure A4 that selection is not biasing our panel estimates.

Identification in the Heckman model relies on the exclusion restriction that experiencing a birth in the first three years of presence affects the probability of staying more than three years, but does not directly affect the wage profile in year 3. An argument against our identification assumption is that having a child may have an effect on the pretax earnings profile through either a standard labor supply response or because of reverse causality (e.g., individuals knowing that they will be experiencing a wage increase are more likely to have children). Two points can be made to alleviate this concern. First, a labor supply response to the birth of a child would affect earnings profiles after the birth of the child, but we should not expect the birth group to experience an increase in pretax earnings exactly at year 3, because individuals in the birth group experience a birth in *any* of their first three years in Denmark. Figure A4 in the Online Appendix shows that in both the birth and the no-birth groups, increase in earnings happens exactly in year 3. Second, if the invalidity of the exclusion restriction is driving our result because of correlation between child birth and earnings profiles, we should expect that the placebo specifications of columns (3) and (4) of Table III, where we focus on entrants in years 1980 to 1990 (before the scheme was in place), would generate different results for the birth versus no-birth groups. However, estimates in columns (3) and (4) are both small and insignificant, even though the first stage (the effect of having a child in the first three years on the probability of staying more than three years) is virtually the same as in our postscheme sample. This confirms that the effect of the scheme lapsing on pretax earnings cannot be attributed to selection.

As a caveat, note that the effects of wage earnings we obtain could in principle be explained within the competitive labor market model if each new cohort of foreign migrants is an imperfect substitute with other cohorts and wages adjust immediately.

In that case, a large fraction of the cohort of migrants leaves when the scheme elapses, which leads to an increase in the wage of this cohort relative to younger cohorts still in the scheme. If the intensive labor supply response to the tax increase is small, earnings could rise when the scheme elapses due to the wage effect.

V. CONCLUSION

Our article has analyzed the effects of income taxation on the international migration and earnings of top earners using a Danish preferential tax scheme targeted to highly paid migrants into Denmark. This scheme offers an opportunity to credibly estimate elasticities of international mobility with respect to tax rates as well as the effects of top tax rates on individual earnings. Using population-wide Danish administrative tax data, we have obtained two results. First, we have shown that the scheme doubled the number of highly paid foreigners in Denmark relative to slightly less paid ineligible foreigners, which translates into a very large elasticity of migration with respect to the net-of-tax rate. Second, we find compelling evidence of a negative effect of scheme-induced increases in the net-of-tax rate on pretax earnings at the individual level. This finding cannot be explained by the competitive labor supply model, where pay equals marginal productivity, but can be rationalized within a simple matching friction model of job search and wage bargaining where there is a gap between pay and marginal productivity.

Our findings show that the migration elasticity is much larger than the conventional within-country earnings elasticity with respect to the net-of-tax rate. As in the case of wealth mobility across countries (Kanbur and Keen 1993), it is conceivable that elasticities of worker mobility are particularly large for small countries (with small tax bases relative to the global economy) and that those small countries therefore have the most to gain from preferential tax schemes to foreigners. Such incentives to offer tax havens for highly skilled foreign workers are likely to generate tax competition across European countries. This will require international coordination and the design of rules regulating such special schemes.

LONDON SCHOOL OF ECONOMICS
 LONDON SCHOOL OF ECONOMICS
 UNIVERSITY OF CALIFORNIA BERKELEY AND NATIONAL BUREAU
 OF ECONOMIC RESEARCH
 KRAKA, COPENHAGEN

SUPPLEMENTARY MATERIAL

An Online Appendix for this article can be found at QJE online (qje.oxfordjournals.org).

REFERENCES

- Bakija, Jonathan, and Joel B. Slemrod, "Do the Rich Flee from High State Taxes? Evidence from Federal Estate Tax Returns," NBER Working Paper No. 10645, 2004.
- Blundell, Richard, and Thomas MaCurdy, "Labor Supply: A Review of Alternative Approaches," in *Handbook of Labor Economics*, vol. 3A, Orley Ashenfelter, and David Card, eds. (Amsterdam: North-Holland, 1999).
- Chetty, Raj, John Friedman, Tore Olsen, and Luigi Pistaferri, "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records," *Quarterly Journal of Economics*, 126 (2011), 749–804.
- Feldstein, Martin, "Tax Avoidance and the Deadweight Loss of the Income Tax," *Review of Economics and Statistics*, 81 (1999), 674–680.
- Fullerton, Don, and Gilbert Metcalf, "Tax Incidence," in *Handbook of Public Economics*, vol. 4, Alan Auerbach, and Martin Feldstein, eds. (Amsterdam: North-Holland, 2002).
- Hall, Robert E., "Employment Fluctuations with Equilibrium Wage Stickiness," *American Economic Review*, 95 (2005), 50–65.
- Howitt, Peter, and R. Preston McAfee, "Costly Search and Recruiting," *International Economic Review*, 28 (1987), 89–107.
- Immervoll, Herwig, Henrik Kleven, Claus Kreiner, and Emmanuel Saez, "Welfare Reform in European Countries: A Microsimulation Analysis," *Economic Journal*, 117 (2007), 1–44.
- Kanbur, Ravi, and Michael Keen, "Tax Competition and Tax Coordination When Countries Differ in Size," *American Economic Review*, 83 (1993), 877–892.
- Kirchgassner, Gebhard, and Werner Pommerehne, "Tax Harmonization and Tax Competition in the European Union: Lessons from Switzerland," *Journal of Public Economics*, 60 (1996), 351–371.
- Kleven, Henrik J., Camille Landais, and Emmanuel Saez, "Taxation and International Migration of Superstars: Evidence from the European Football Market," *American Economic Review*, 103 (2013), 1892–1924.
- Kleven, Henrik J., Camille Landais, Emmanuel Saez, and Esben Schultz, "Migration and Wage Effects of Taxing Top Earners: Evidence from the Foreigners' Tax Scheme in Denmark," NBER Working Paper No. 18885, 2013.
- Kleven, Henrik J., and Esben A. Schultz, "Estimating Taxable Income Responses Using Danish Tax Reforms," *American Economic Journal: Economic Policy*, 6 (2014).
- Kleven, Henrik J., and Mazhar Waseem, "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan," *Quarterly Journal of Economics*, 128 (2013), 669–723.
- Liebig, Thomas, Patrick A. Puhani, and Alfonso Sousa-Poza, "Taxation and Internal Migration: Evidence from the Swiss Census using

- Community-Level Variation in Income Tax Rates," *Journal of Regional Science*, 47 (2007), 807–836.
- Mortensen, Dale, and Christopher Pissarides, "New Developments in Models of Search in the Labor Market," in *Handbook of Labor Economics*, vol. 3B, Orley Ashenfelter, and David Card, eds. (Amsterdam: North-Holland, 1999).
- OECD (Organisation for Economic Co-operation and Development), "The Taxation of Mobile High-Skilled Workers," chapter 4 in *OECD Tax Policy Study No. 21: Taxation and Employment* (Paris: OECD, 2011).
- , *Revenue Statistics* (Paris: OECD, 2012).
- , *Taxing wages* (Paris: OECD, 2013).
- Piketty, Thomas, and Emmanuel Saez, "Optimal Labor Income Taxation," in *Handbook of Public Economics*, vol. 5, Alan Auerbach, Raj Chetty, Martin Feldstein, and Emmanuel Saez, eds. (Amsterdam: North-Holland, 2013).
- Saez, Emmanuel, "Do Taxpayers Bunch at Kink Points?," *American Economic Journal: Economic Policy*, 2 (2010), 180–212.
- Saez, Emmanuel, Joel Slemrod, and Seth Giertz, "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review," *Journal of Economic Literature*, 50 (2012), 3–50.
- Young, Cristobal, and Charles Varner, "Millionaire Migration and State Taxation of Top Incomes: Evidence from a Natural Experiment," *National Tax Journal*, 64 (2011), 255–284.