DISCUSSION PAPER SERIES

No. 9410

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

Henrik Kleven, Camille Landais, Emmanuel Saez and Esben Schultz

PUBLIC POLICY

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP9410.asp**

ISSN 0265-8003

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

Henrik Kleven, London School of Economics and CEPR
Camille Landais, London School of Economics
Emmanuel Saez, UC Berkeley, and CEPR
Esben Schultz, Kraka Copenhagen

Discussion Paper No. 9410
March 2013

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 9410

March 2013

ABSTRACT

Migration and Wage Effects of Taxing Top Earners: Evidence from the Foreigners' Tax Scheme in Denmark*

This paper 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 three main results. First, the scheme has 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 on foreigners, between 1.5 and 2. Hence, preferential tax schemes for highly paid foreign workers could create severe tax competition between countries. Second, we find compelling evidence of a negative effect of scheme-induced increases in the net-of-tax rate on pre-tax earnings at the individual level. This finding cannot be explained by the standard labor supply model where pay equals marginal productivity, but it can be rationalized by a matching frictions model with wage bargaining where there is a gap between pay and marginal productivity. Third, we find no evidence of positive or negative spillovers of the scheme-induced influx of high-skilled foreigners on the earnings of highly paid natives.

JEL Classification: H22, H31 and J61
Keywords: international migration, taxation and wage bargaining

Henrik Kleven
Department of Economics
London School of Economics
Houghton Street
London WC2A 2AE

Email: H.J.Kleven@lse.ac.uk

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=159550

Camille Landais
Department of Economics
London School of Economics
Houghton Street
London WC2A 2AE

Email: c.landais@lse.ac.uk

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=159730

Emmanuel Saez
University of California, Berkeley
549 Evans Hall, 3880
Berkeley
CA 94720
USA

Email: saez@econ.berkeley.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=145017

Esben Schultz
Kraka
Kompagnistræde 20A
1208 Copenhagen K
DENMARK

Email: ess@kraka.org

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=177027

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

Submitted 20 March 2013

1 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 limiting the redistributive power of governments and potentially creating tax competition. The introduction in many European countries of preferential tax schemes to high-skilled foreigners represents *prima facie* evidence of such tax competition.[1] This debate raises two 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 obtain part of the benefit? Both questions are crucial for evaluating international tax design for top earners, while the second question is also key for understanding the functioning of the labor market for top earners. This paper breaks new ground on these questions using sharp quasi-experimental variation created by a Danish preferential tax scheme for high-earning immigrants.

While an enormous empirical literature has studied labor supply and taxable income responses to taxation (as surveyed by Blundell and MaCurdy 1999 and Saez, Slemrod, and Giertz 2012), there is very little empirical work on the effect of taxation on the spatial mobility of individuals,[2] 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 cast new light on the wage setting process and

[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.

[2] A small literature has considered the mobility of people across local jurisdictions *within* countries. See Kirchgassner and Pommerehne (1996) and Liebig et al. (2007) on mobility across Swiss Cantons in response to Canton taxes, Young and Varner (2011) and Bakija and Slemrod (2004) on mobility across U.S. states in response to state income and inheritance taxes.

[3] A recent exception is Kleven, Landais, and Saez (2013), who analyze the labor market for professional football players across 14 European Union countries and find compelling evidence of tax-induced mobility responses. However, a concern is that football players might be substantially more mobile than other high-skilled workers, because football players 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-*

provide strong empirical evidence in favor of the widely used labor market model with matching frictions and wage bargaining.[5]

In 1992, Denmark enacted a preferential tax regime 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 3 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.[6] 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). Absent the special tax scheme, workers with earnings above the scheme threshold would face average income tax rates of around 55%, about twice as high as the scheme rate. When the 3 years of preferential tax treatment have been used up, the taxpayer becomes subject to the ordinary tax schedule on subsequent earnings.

This unusual piece of tax policy creates large discontinuities in tax liability depending on the contract start date (before and after June 1, 1991), duration of stay in Denmark (3-year rule), and earnings level (earnings eligibility threshold). Hence, the reform generates sharp quasi-experimental variation along several different dimensions, and provides a very powerful way of identifying the effect of taxation on migration and earnings. In this paper, we exploit the different aspects of the tax scheme using quasi-experimental techniques such as bunching approaches and differences-in-differences methods. For this analysis, we have gained access to matched employer-employee administrative data for the full population of Danish residents (Danish citizens and foreigners) since 1980. The data includes detailed information on citizenship, immigration history, income and tax variables, labor market variables, and socio-demographic variables. It also contains specific information for all scheme beneficiaries.[7]

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

level changes in wages, which are interesting in their own right and can be identify non-parametrically.

[5]Another recent example of using sharp tax variation to uncover non-standard wage determination effects is Saez, Matsaganis, Tsakloglou (2012). Using cohort-based payroll tax variation in Greece, they show that employers bear the incidence of employer-side payroll taxes while employees bear the incidence of employee-side payroll taxes even in the long run. This suggests that wage determination is anchored to the posted wage, which is not consistent with the standard model of tax incidence.

[6]The scheme also applies to Danish citizens, who have been abroad with tax residence outside Denmark for a period of at least 3 years.

[7]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 and foremost, 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 afterwards. As a result, the fraction of foreigners in the top one-half percent 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 among foreigners, 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).[8] It can therefore be desirable from a single-country revenue perspective to adopt such preferential schemes for highly paid foreigners. At the same time, those schemes impose negative fiscal externalities by potentially reducing revenue in other countries. This tension between country welfare and global welfare in the design of individual income tax policy has loomed large in the public debate for a long time, but our paper provides the first compelling evidence that this is indeed a major empirical issue. Absent coordination, it is conceivable that such schemes could unravel tax progressivity in Europe.

Second, we find compelling evidence of a negative effect of scheme-induced increases in the net-of-tax rate on pre-tax earnings at the individual level. We show in a differences-in-differences setting that eligible foreigners experience a 5 to 10% decline in their pre-tax earnings relative to non-eligible foreigners after the introduction of the scheme, even after controlling for individual characteristics and differential time trends. Most importantly, we find that migrants who stay in Denmark beyond the 3-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 3-year discontinuity, we ensure that this result is not driven by selection into Denmark or by non-tax aspects of wage-tenure profiles. In the standard labor supply 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

[8]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.

evidence of a hole below the threshold. This suggests that bunchers are coming from the low-tax side of threshold, which is inconsistent with the standard labor supply model where bunching is driven by individuals on the high-tax side increasing their labor supply to qualify for the scheme. While those two empirical findings contradict the standard labor supply model where pay equals marginal productivity, they 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.[9]

Third, we find no evidence of positive or negative spillovers of the scheme-induced influx of high-skilled foreigners on the earnings of natives either at the industry or firm level. The absence of spillover effects might be due to the relatively short stays of the highly paid foreigners and/or to the relatively small share of high-skilled foreigners in the total labor force of high-skilled workers in Denmark, even after the tax scheme was put in place. It should be noted, however, that identifying such spillover effects is intrinsically difficult even with our tax variation, and so the absence of spillover effects is only suggestive.

Finally, while our study is based on a single country characterized by a certain size, culture, quality of life, immigration tradition, etc., 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 non-trivial welfare costs for large countries (Keen and Konrad, 2013 provide a recent survey of the tax competition literature).

The paper is organized as follows. Section 2 describes the Danish foreigners' tax scheme

[9]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).

and the administrative data we use. Section 3 lays out a simple theoretical framework with matching fractions and wage bargaining and contrasts its implications with the standard labor supply model. Section 4 presents the empirical analysis. Section 5 concludes by discussing policy implications.

2 Institutional Background and Data

2.1 The Foreigners' 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 paper, 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 1994-1997, leaving the total scheme rate roughly unchanged around 30%.[10]

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 any number of spells over an unlimited period of time.[11] 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 3-year cutoff and hence a *kink* in the lifetime budget set as a function of duration.

There are two key requirements to become eligible for the preferential tax scheme. The first requirement is that the taxpayer has been recruited abroad and not been tax liable in Denmark in the 3 years prior to going on the scheme.[12] In our data period, citizenship plays no formal role

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

[11]After 2008 (outside our data period), an additional scheme option was introduced whereby eligible workers can choose between the standard scheme rate of 25% for 36 months and a higher scheme rate of 33% for 60 months (the payroll tax contribution comes in addition to both of those rates). Reports suggest that the take-up of the newly introduced 60-month scheme option has been very low.

[12]For taxpayers who split scheme take-up into several spells, the 3-year eligibility 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.

in determining eligibility, and therefore Danish citizens who have been living and paying taxes abroad (i.e., were not residents of Denmark for tax purposes) for at least 3 years can also apply for the scheme.[13] The second requirement is that, unless the worker qualifies as a researcher, annual wage earnings must be above an eligibility threshold. The threshold grows roughly at the rate of average earnings,[14] and it always lies between the 99.2th and 99.4th percentile of the Danish wage earnings distribution. It is equal to 765,600 Danish kroner (about 103,000 Euros) in 2009. As the preferential scheme rate applies to *all* 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.

In terms of administration, the scheme treatment has to be requested by the employer. Hence, the employer has to show the tax authorities that the level of earnings is above the eligibility threshold and that other qualifying requirements are met. Importantly, the threshold for eligibility applies only to earnings with the specific employer requesting the scheme. Having other sources of income or earnings do not help qualify. The threshold of eligibility must be met on an annualized basis. Hence for a contract of 6 months, the eligibility threshold is that the 6 months of pay must be at least half of the annual threshold. Perquisites such as free 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 scheme-qualifying earnings, that income is taxed according to the standard progressive income tax schedule independently of scheme earnings. In other words, scheme earnings are effectively taxed at a flat rate completely independently of the other circumstances of the individual. In particular, when the 36-months scheme duration ends during a given calendar year, and the individual stays in Denmark, any post-scheme earnings will be taxed according to the regular schedule. There is no pro-rating for non-scheme earnings (or unearned income) taxed according to the regular tax schedule. As the personal income tax in Denmark is individually based (and not family based), spouses of scheme recipients are taxed according to the regular tax (except if they themselves qualify for

[13]This rule was changed in 2011 such that Danish citizens must now be foreign tax residents for at least 10 years to become eligible for the scheme, but this reform lies outside our data period.

[14]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, virtually all other methods for imputing the threshold before 1990 deliver identical results.

the scheme). Scheme earners residents of Denmark have access to the same public goods as other residents, including public health insurance and schooling for children.[15]

To summarize, the special Danish tax scheme creates the following quasi-experimental variation. First, the scheme introduced much lower tax rates on a specific sample of workers (high-earning immigrants; not tax liable in Denmark 3 years prior) arriving in Denmark after June 1, 1991. This variation provides an ideal setting for a differences-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 3-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 3-year scheme duration.

2.2 Data and Summary Statistics

Administrative tax data. The data we use in this paper come from an administrative dataset including the universe of tax and payroll records for all resident individuals in Denmark, including both Danish citizens and foreigners since 1980. The data includes detailed information about earnings, tax variables, labor market variables, and socio-demographic variables at an annual frequency. Most importantly, the data contains detailed citizenship and migration information such as daily dates of entry and exit. Each individual working in Denmark must receive a personal identification number (CPR) in order to pay withholding taxes, rent an apartment, register with health insurance, etc. The application for a CPR number 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[16]. The data also contain detailed information specifically for scheme beneficiaries on for example 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 1991-1994. The data have been linked to employers both at the firm level and the establishment level, with information on the 2-digit level industrial sector of the

[15]Scheme earners do not qualify for unemployment and disability insurance and scheme earnings do not count for retirement benefits computation purposes. Scheme earners typically receive additional private and employer provided health insurance.

[16]This update can nevertheless take a few years.

employer.

Summary statistics. Table 1 presents summary statistics. First, among all scheme beneficiaries, 95% have registered for a CPR number implying that they have become actual residents of Denmark. This shows that only a tiny minority of scheme recipients take advantage of the scheme without actually becoming residents. We exclude such non-resident scheme beneficiaries throughout the analysis. Second, the total duration of stay (including both time under the scheme and time after the scheme elapses) for people who ever benefitted from the scheme averages 2.35 years with about a quarter of scheme beneficiaries staying in Denmark more than 36 months, the maximum duration of the scheme. Third, scheme earnings average 153% of the threshold, or about 150,000 Euros as of 2009. The average tax rate of scheme earners (including the scheme flat rate and the payroll tax contribution starting in 1994) is around 30%. Fourth, 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 average earnings of 36,500 Euros for all employees in Denmark. Those firms employ relatively few scheme workers (1.8 on average) even though they employ on average 14 workers paid above the threshold.

Figure 1 reports the composition of beneficiaries of the tax scheme (excluding researchers) by country of citizenship (Panel A) and industrial sector (Panel B) from 1991-2006. Unsurprisingly, the vast majority of scheme workers come from advanced economies: 25% come from Nordic countries (except Denmark), 10% are Danish citizens (who qualify by not being Danish tax residents for at least 3 years as explained above), 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 number coming from emerging countries or Eastern Europe and Russia is modest. The composition of beneficiaries of the tax scheme (excluding researchers) by industrial sector (panel B) reveals that all sectors made use of the scheme, but with a clear overrepresentation of industries such as the financial sector and the sports and entertainment industry.

We can compute a take-up rate for the scheme as the fraction of foreigners arriving in Denmark with (annualized) earnings above the eligibility threshold, who have not paid taxes in Denmark over the last 3 years, and who take advantage of the scheme. The take-up rate is high at 81%, but still significantly below 100% for a variety of reasons. First, companies have

to file an application for each employee eligible for the scheme, and it is conceivable that not all companies knew about the scheme or were willing to bear the administrative burden. Second, it is possible that the non-scheme foreigners were not truly eligible, because they did have one labor contract specifying earnings above the threshold *ex ante* even if they ended up having total earnings above the threshold *ex post* (hence, the take-up rate of 80% is a lower bound). Third, some individuals may not have been willing to take up the scheme (perhaps because of the original claw-back rule after 7 years), or they may not have been fully aware of the existence of the scheme.

3 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 standard labor supply model conventionally used in empirical labor supply and tax studies. Since some of our empirical results cannot be rationalized within the standard labor supply 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. In section 4.3 however we investigate whether we can detect the presence of such incidence effects on high-skilled native workers.

3.1 Standard 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 in order to maximize a utility function that depends on net-of-tax earnings, labor supply, and location. The scheme affects behavior along three dimensions: (1) the migration decision, (2) the duration of stay conditional on migrating, and (3) labor supply and earnings among migrants. Let us review each dimension in turn.

Migration. The scheme reduces the average tax rate on high-earning immigrants to Denmark and therefore makes Denmark more attractive to such workers. Hence, the scheme should

increase immigration of high-earning foreigners into Denmark. Note also that, in principle, their could be a tax avoidance response whereby multinational companies with presence in Denmark try to re-characterize some of their workers abroad with some presence in Denmark as Danish tax residents. As discussed in Table 1, our data show that 95% of scheme eligible workers have real presence in Denmark showing that such tax avoidance responses are small. Furthermore, we exclude non-resident scheme earners from our analysis to concentrate on real migration effects.

Duration. When scheme eligibility elapses after 3 years, marginal migrant workers may no longer find it attractive to stay in Denmark and thus move back to their home country. As illustrated in Figure 2, Panel A, the discrete jump in the tax rate at the 3-year duration threshold D^* produces a *kink* in the *life-time* budget constraint of duration of stay (x-axis) vs. disposable life-time income (y-axis). This should create excess bunching around 36 months in the density of duration of highly paid foreigners in Denmark.[17] Importantly, in the standard model, the wage rate of highly paid immigrants should not change when the scheme elapses because their marginal productivity does not change.[18]

Labor supply. 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 pre-tax earnings (x-axis) vs. disposable post-tax earnings (y-axis) at the scheme eligibility threshold $\bar{z}$ as depicted in Figure 2, Panel B. According to the standard labor supply model, such a notch induces foreign workers just below the threshold to increase earnings to a point just above the threshold, thereby producing a hole in the earnings distribution on the high-tax side and excess bunching in the earnings distribution on the low-tax side of the notch point. As analyzed by Kleven and Waseem (2013), this 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 bunching on the low-tax side should be accompanied by an equal amount of missing mass on the high-tax side. Furthermore, for foreign workers above the threshold, the

[17] If the cost of being in Denmark (relative to one's own country) is a constant flow cost, then all scheme-induced migrants will leave after 3 years. More realistically, if the cost of being in Denmark declines as a function of time spent in the country—as for example when part of the cost is an up-front moving cost—then only a fraction of scheme-induced migrants will leave Denmark when scheme eligibility elapses.

[18] Note also that the standard wage incidence channel should affect workers just below and just above the 3-year duration threshold in the same way (as such workers are perfect substitutes). Hence, there should be no discontinuity in the wage rate due to standard wage incidence effects when the scheme elapses.

scheme creates an uncompensated net-of-tax wage increase, which should increase labor supply and therefore earnings *under the reasonable assumption that the uncompensated labor supply elasticity is positive*. Conversely, when the scheme elapses after the 3-year maximum duration, the net-of-tax wage rate decreases and we should observe a symmetric decrease in labor supply and earnings of immigrants who stay in Denmark beyond 3 years. Bunching at the eligibility threshold $\bar{z}$ should also disappear when the scheme elapses.

It is possible that 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). For example, employees can give up some non-taxable benefits or on-the-job amenities (such as better offices, flexibility in the work schedule, etc.) to get higher taxable pay and qualify for the scheme.[19] Such avoidance effects simply magnify the behavioral responses of the standard labor supply model we have described above.

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 3-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.

3.2 Matching Frictions and Wage Bargaining Model

As we shall see, our empirical findings contradict some of the predictions of the standard labor supply model presented above. We therefore consider an alternative simple model with matching frictions and wage bargaining that can rationalize all of our empirical findings and which nests the standard competitive model. To simplify, we focus on the case of inelastic labor supply

[19] As mentioned above, 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, e.g., colluding with the employee to fake earnings to meet the eligibility threshold.

[20] Such inter-temporal substitution could also be due to real labor supply changes, e.g., 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.

(conditional on migration).

A worker contemplating migration to Denmark has marginal product y for a prospective employer in this country. In the standard competitive model used above, 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 (see e.g., Pissarides, 2000). The worker has a pre-tax 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.[21] 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).

Migration. A worker migrates if and only if $y_0 \leq y$, i.e. when there exists a band of wages for which the move is mutually beneficial to the worker and the prospective 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 pre-tax reservation wage y_0 must satisfy $y_0 \cdot (1 - \tau) = z_h \cdot (1 - \tau_h) + \nu$, i.e.

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

where $y_0^{\tau=0} \equiv z_h \cdot (1 - \tau_h) + \nu$ is the reservation wage if there were 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 $\bar{z}$. Hence, there are two reservation wages

[21] If $y_0 > y$, then no wage can satisfy both the employee and the employer and the match cannot proceed.

y_0^S, y_0^D based on whether the worker is eligible for the scheme or not:

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

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

Duration. The effect of the scheme on duration of stay is qualitatively similar to the standard labor supply model. The 3-year duration of the scheme creates an incentive for stays lasting at most 3 years, thereby producing excess mass in the duration density below 3 years as well as sharp bunching at the 3-year threshold. If the net cost of living in Denmark ν and potential net-of-tax earnings at home $z_h(1-\tau_h)$ were constant over time, then all scheme-induced migrants would leave exactly at the 3-year threshold. With idiosyncratic time variation in those variables, some of the scheme-induced migrants would stay after 3 years and some would leave before 3 years.

Wage determination. Within the band $[y_0, y]$ of acceptable wages, how is the equilibrium wage determined? While there are many potential models of wage determination in this type of setting, the most widely used model assumes that the pre-tax 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 firm). As we shall see, the Nash bargaining framework is particularly useful to solve the model in the presence of *discontinuous* incentive schemes (which arise here because of the scheme eligibility threshold).

Given the tax rate τ, pay z is set to maximize $W = (y-z)^{1-\beta}\left((1-\tau)\,z - y_0^{\tau=0}\right)^{\beta}$ where $y-z$ is the firm's surplus and $(1-\tau)z - y_0^{\tau=0}$ is the worker's surplus. To begin with, we ignore the notch (discrete jump in τ) at the threshold $\bar{z}$ in which case the solution to the bargaining problem is characterized by

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

Note that this model nests the standard frictionless case when $\beta = 1$.[22] From condition (1), earnings under the scheme tax rate τ^S are equal to $z^S = (1-\beta)y_0^S + \beta y$, while earnings under

[22] More generally, any wage z in the acceptable band $[y_0, y]$ implicitly defines a bargaining weight β such that equation (1) holds.

the regular Danish tax rate τ^D are equal to $z^D = (1-\beta)y_0^D + \beta y$. Hence, the scheme *reduces* pre-tax 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 $\bar{z}$. There will be bunching at the threshold as in the standard labor supply model, but a conceptual difference is that the matching model predicts that bunchers may be coming from both the low-tax and the high-tax side of the threshold. This can be understood as follows. First, consider workers with earnings above the threshold $\bar{z}$ in the absence of the scheme, i.e. workers for whom $z^D > \bar{z}$. As shown above, the introduction of the preferential scheme rate reduces pre-tax earnings inside the eligible range when firms have positive bargaining power, and so we will have $z^S < z^D$ for these workers. However, if earnings were sufficiently close to the threshold to begin with, a situation arises where $z^S < \bar{z} < z^D$ 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 $\bar{z}$ in the absence of the scheme, i.e. workers for whom $z^D < \bar{z}$. Even though the tax rate has not changed in this range, the introduction of the notch at $\bar{z}$ 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, note that extensive migration responses also contribute to bunching as marginal entrants have an excess tendency to locate at $\bar{z}$. Again, this reflects bunching either from above (those for whom $\bar{z} < z^D < y_0^D$ without the scheme and $z^S = \bar{z} > y_0^S$ with the scheme) or from below (those for whom $z^D < \bar{z}, y_0^D$ without the scheme and $z^S = \bar{z} > y_0^S$ with the scheme).

Importantly, bunching from below (created by $\beta > 0$) is associated with a hole in the earnings distribution below the threshold, whereas bunching from above (created by $1-\beta > 0$) is created by a shift in the entire distribution above the threshold 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 3 illustrates these theoretical results in a density distribution diagram. Panel A shows how the bunch and the hole is created by bargaining responses (conditional on migration) from below and above depending on the bargaining power of workers relative to firms, while 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 $\bar{z} \leq y$ migrate into the country because of the scheme. This lifts up the density of foreign migrants above $\bar{z}$. A fraction of those migrants will bunch at $\bar{z}$.

Prediction 2: Duration. The 3-year duration of the scheme produces excess mass in the duration density below 3 years and sharp bunching at the 3-year threshold.

Prediction 3: Wages. Among migrant workers paid above the eligibility threshold $\bar{z}$, the scheme *reduces* pre-tax earnings.

Prediction 4: Bunching & Hole. There is bunching at $\bar{z}$ from above when employers have bargaining power $(1 - \beta > 0)$. There is bunching at $\bar{z}$ from below when workers have bargaining power $(\beta > 0)$. There is a hole below $\bar{z}$ only when workers have bargaining power $(\beta > 0)$.

4 Empirical Evidence

In the following, we estimate the empirical effects of the scheme on migration and duration (Section 4.1), wages (Section 4.2) and spillovers on natives (Section 4.3), using the conceptual framework set out above. We show that each of the four predictions of the matching frictions model are borne out by the data.

4.1 Migration Effects

differences-in-differences approach. As a first step in testing whether the Danish tax scheme had an impact on high-skilled migration, we consider the evolution of the number of foreigners

with earnings above the scheme threshold between 1980-2005 in Figure 4, Panel A.[23] This series (labelled 'treatment' in the figure) shows that the number of highly paid foreigners was fairly stable around 800 in the pre-scheme 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 2000 by 1997 and is close to 3000 by 2005. It is of course conceivable that the number of highly paid foreigners would 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 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 (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 below potential confounders to this simple differences-in-differences strategy, two lessons emerge from the use of those 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 5 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 but more slowly than 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 4, 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

[23]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.

earnings between 90% and 99.5% of the threshold (control 2).[24] 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. Naturally the series for arrivals in Panel B are noisier than the series for stocks in Panel A.

Table 2 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): (1) the elasticity of the total number of foreigners (as in Panel A of Figure 4), (2) the elasticity of the number of arrivals of foreigners (as in Panel B of Figure 4), and (3) the elasticity of the number of foreigners with less than 3 years of presence in Denmark. These elasticities are estimated using a 2SLS regression specification of the form

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

where $i = 0, 1$ denotes control and treatment group, t denotes year, N_{it} is the number of foreigners in group i and year t (corresponding to each of the outcomes in columns 1-3), τ_{it} is the average tax rate in group i and year t, $\mathbb{1}[i = 1]$ is the treatment group dummy, and γ_t are year fixed effects. The key variable of interest $\log(1 - \tau_{it})$ is instrumented by the interaction $\mathbb{1}[i = 1] \cdot \mathbb{1}[t > 1991]$. As a baseline, we compute τ_{it} 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, while 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 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 post-reform period, whereas the short-term elasticity includes only 1992-1996 as the post-reform period. All specifications include 1980-1990 as the pre-reform period.

Our baseline estimates in Panel A1 of Table 2 show that elasticities are large and precisely estimated, between 1.5 and 2 across the different elasticity definitions. The large magnitude of

[24] Again, both control series are normalized so that they match the treatment series in 1990 the year before the scheme was first implemented.

elasticities can be understood directly from Figure 4: 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 $\log(2.2)/\log(.7/.4) \simeq 1.5$. 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 3 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 come back later.

The following panels consider various robustness checks. Panel A2 controls for a potential difference in pre-existing trends between the treatment and control groups. We first regress $\log N_{it}$ for all years prior to the reform on group fixed effects and two group specific time trends (i.e., $\mathbb{1}[i = 0] \cdot t$ and $\mathbb{1}[i = 1] \cdot t$). And then we use the difference between $\log N_{it}$ and the predicted values of the first regression $\widehat{\log N_{it}}$ 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 group is defined as foreigners with income between 90% and 99% percent of the eligibility threshold and the control group is defined as foreigners with income between 80% and 90% of the threshold (with the net-of-tax rate variable—the denominator of the elasticity—being the same as in Panel A1). This specification also tests for the potentially confounding implications of shifting around the eligibility threshold (via earnings or avoidance responses) in order to qualify for the scheme, since 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 (.1 or less) and insignificant, which confirms the graphical evidence that the two control groups follow extremely similar trends and shows that shifting around the threshold is second order (we come back to this below).

Panel A4 controls for imperfect take-up by taking into account that some individuals in the treatment group (those with earnings above the eligibility threshold) are not on the scheme. Specifically, we compute actual tax rates (given actual take-up) and instrument 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 much bigger effect on the estimate for the total number of foreigners in column (1) as foreigners who stay beyond 3 years are no longer eligible for the scheme.[25]

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, UK, US) 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, in Panel A7, we report estimates of the effect of the scheme on expatriates, i.e., Danish citizens returning to Denmark after a spell of at least three years abroad (so that they qualify for the scheme upon 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 is very important for the policy debate on taxes and mobility as we discuss below.

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 just below the threshold. There are two potential confounders: (1) a fanning-out of the earnings distribution after 1990, (2) an endogenous earnings response to the scheme threshold (notch) as analyzed in the conceptual framework of Section 3. 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 absent the scheme. However, this concern is unlikely to be very important in the case of Denmark as Kleven and Schultz (2012) have shown that top earnings shares have been remarkably stable in Denmark since 1980. To address this issue more directly and control for any change in the earnings distribution, Figure 5 plots 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

[25]This elasticity can be used to predict the effect on the total number of foreigners if the scheme were of infinite duration, extrapolating from the estimated effects of the actual 3-year duration.

with positive earnings, Figure 5 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). Note that the figure features a gap for the treatment group between 1991-1994, because the scheme data does not provide scheme earnings for those years. Two important findings emerge from the figure. First, the fraction of foreigners in each percentile group is extremely stable before 1991.[26] Second, 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 4, there is a doubling of the fraction foreigners above the 99.5th percentile relative to percentile groups just below the scheme eligibility threshold. This finding confirms that changes in the earnings distribution do not pose a threat to our findings. This graphical result is confirmed in Table 2, Panel B where we define the treatment group as individuals with earnings above the 99.5th percentile and the control group as individuals with earnings between the 95th and 99th percentile. The elasticity estimates are slightly attenuated relative to our baseline specification of panel A1, but remain very large around 1.2 and precisely estimated.

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 4 (those between 90-99% of the threshold and those between 80-90% of the threshold) along with the placebo estimates in Panel A3 of Table 2 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 6 plots such densities before the scheme was introduced (1980-1990 in dashed grey) and after the scheme was 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 pre-scheme period. The

[26]The fraction of foreigners is higher in levels, around 3%, above the 99.5th percentile, than in percentiles 96 to 99.5 where it is around 2%.

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 4 and showing that this response occur 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 figure reports that the excess mass due to bunching is statistically significant while the missing mass on the left of the threshold is not statistically significant. 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.[27] The presence of bunching is consistent with both the standard labor supply model and the matching friction model presented above. 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 2, Panel B). Using the method developed by Kleven and Waseem (2013), the

[27] Formally, in the absence of directly observable counterfactual for the distribution in the absence of the notch, one would estimate models of the form: $c_j = \sum_{i=0}^{p} \alpha_i^- \cdot (z_j)^i + \sum_{i=0}^{p} \alpha_i^+ \cdot (z_j)^i \cdot \mathbb{1}[z > \bar{z}] + \sum_{i=l}^{u} \gamma_i \cdot \mathbb{1}[z_j = i] + \nu_j$, where $[l, u]$ is the excluded range around the notch point, c_j is the log number of observation in each bin of earnings z_j, and $(z_j)^i$ are non-parametric polynomial fits. The counterfactual distribution is then, $\hat{c_j} = \sum_{i=0}^{p} \widehat{\alpha_i^-} \cdot (z_j)^i + \sum_{i=0}^{p} \widehat{\alpha_i^+} \cdot (z_j)^i \cdot \mathbb{1}[z > \bar{z}]$. From this counterfactual distribution, missing mass M and bunching B can easily be estimated as $\widehat{M} = \frac{\sum_{i=l}^{\bar{z}} (c_j - \hat{c_j})}{\widehat{c_{\bar{z}}}}$ and $\widehat{B} = \frac{\sum_{i=\bar{z}}^{u} (\hat{c_j} - c_j)}{\widehat{c_{\bar{z}}}}$. The difference here is that we take advantage of the existence of a counterfactual distribution prior to 1991 in the absence of a notch to enrich the quality of the counterfactual estimate of the distribution in the absence of a notch after 1991. We do so by estimating a model of the form

$$c_{j,t} = \overbrace{\sum_{i=0}^{p} \alpha_{i,t_1} \cdot (z_{j,t_1})^i}^{\text{baseline density before 1991}} + \overbrace{\alpha_{0,t_2}}^{\text{shift in the distribution over time}} + \sum_{i=0}^{p} \eta_{i,t_2} \cdot (z_{j,t_2})^i \cdot \mathbb{1}[z_{j,t_2} > \bar{z}] + \sum_{i=l}^{u} \gamma_i \cdot \mathbb{1}[z_{j,t_2} = i] + \nu_{jt}$$

where $t = t_1$ is before 1991, and $t = t_2$ is after the scheme was introduced in 1991 and $[l, u]$ is the excluded range around the notch point. In practice we chose to exclude the range between .95 and 1.05 of the threshold and we used polynomial specifications of order 6.

implied labor supply elasticity in the standard model of Section 3.1 would be extremely small, less than .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 standard 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 standard labor supply model).

Duration. We now analyze 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 7. Panel A focuses on the pre-reform 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 3-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 post-reform 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 then doubly normalized using both the Panel A normalization for P99.5-100 and the Panel B normalization 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). Two clear changes emerge after the introduction of the scheme. First, there is a big jump in the duration density for the treatment group in the interval below 3 years compared to the interval above 3 years, confirming that the scheme favors durations of at most three years. Using a differences-in-differences specification, the scheme has reduced the probability of staying more than 3 years in Denmark (conditional on migration) by about 15 percentage points. Second, large and sharp bunching emerges in the duration density precisely at the 3-year threshold consistent with the

conceptual framework. Interestingly, bunching also emerges at the 1-year and 2-year thresholds (and to a small extent at the 4-year and 5-year thresholds), which shows that scheme foreigners tend to negotiate work contracts in full years. Excess bunching at the 3-year threshold is larger than excess bunching at all the other year thresholds as one would expect.

Tax policy implications. Using the elasticities estimated in Table 2, the revenue-maximizing tax rate (Laffer rate) can be computed using the classic inverse elasticity formula, $\tau = 1/(1+e)$. 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 $\tau = 1/(1+1.5) = 40\%$, 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 $\tau = 1/(1+2) = 33\%$, about the level of the current total tax rate under the scheme.

Note also that foreigners pay additional taxes in Denmark when they consume their income in Denmark through the value-added-tax (VAT) and various other commodity taxes.[28] Accounting for such consumption taxes would imply that 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 2, Panel A7, the elasticity for Danish expatriates is much smaller, and at most 0.1. This implies that the migration effect of changing the top tax rate for all Danish residents would be small (see Piketty and Saez, 2013 for a detailed exposition).[29]

4.2 Wage Effects

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

[28] 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.

[29] Our scheme experiment allows us to measure how many Danish expatriate 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.

Repeated cross-section evidence. Figure 8 depicts the average *nominal* annual earnings for foreigners in their first full calendar year in Denmark (denoted by year 1) for the sample of foreigners with year 1 earnings between 70% and 95% of the scheme eligibility threshold (dashed line) and between 105% and 400% 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.[30] For year t, the sample includes foreigners who have arrived in Denmark in year $t-1$, and stay the full calendar year t in Denmark. Both series are normalized to 100 in year 1990, the year 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 4.1, a differences-in-differences estimate based on Figure 8 would imply that the scheme reduces pre-tax earnings. This is consistent with the matching friction model (and inconsistent with the standard labor supply model). We discuss the plausibility of the identification assumption below.

Table 3 presents systematic differences-in-differences estimates of the effect of the scheme on log annual pretax earnings using the same strategy as the one depicted on Figure 8. For year t, the sample again includes foreigners who arrived in Denmark in year $t-1$, and stayed the full calendar year t in Denmark, so that they would be eligible for the scheme based on duration requirements. Hence, the sample is a set of repeated cross-sections with non-overlapping sets of foreigners in each year. We denote by z_{it} the earnings of individual i in year t. We always exclude years t =1991–1994 from the sample (as we have no scheme earnings information for those years). In all columns of Table 3, except (5), we exclude potential bunchers by removing all individuals with earnings between 95% and 105% of the threshold.

We consider first the following reduced form specification:

$$\log z_{it} = \alpha + \beta \cdot \mathbb{1}[z_{it} \geq \bar{z}] + \gamma_t + \eta \cdot \mathbb{1}[z_{it} \geq \bar{z}] \cdot \mathbb{1}[t > 1991] + \nu_{it}, , \tag{3}$$

where $\mathbb{1}[z_{it} \geq \bar{z}]$ is a dummy variable for having earnings above the eligibility threshold $\bar{z}$, γ_t are year fixed effects, and $\mathbb{1}[z_{it} \geq \bar{z}] \cdot \mathbb{1}[t > 1991]$ is the interaction term for being above the

[30] We exclude those with earnings above 400% of the threshold to reduce noise due to outliers. We include outliers in Table 3 below where we consider log-earnings elasticity specifications that naturally dampen the influence of outliers.

eligibility threshold and arriving after the scheme is in place. Hence, η is the coefficient of interest and is reported in the first row of Table 3. In the traditional labor supply model, η is positive as the lower tax rate from the scheme should increase labor supply and hence earnings (if the uncompensated labor supply elasticity is positive). In the matching friction model, η is negative as the lower tax rate from the scheme allows the employer to reduce the pre-tax earnings paid to the employee.

Next, we present elasticity estimates based on the following 2SLS specification

$$\log z_{it} = \alpha + \beta \cdot \mathbb{1}[z_{it} \geq \bar{z}] + \gamma_t + \delta \cdot \log(1 - \tau_{it}) + \nu_{it}, , \tag{4}$$

where τ_{it} is the average tax rate for individual i in year t. The key variable of interest $\log(1-\tau_{it})$ is instrumented by the interaction $\mathbb{1}[z_{it} \geq \bar{z}] \cdot \mathbb{1}[t > 1991]$. We again compute τ_{it} assuming a 100% take-up rate in which case the results should be interpreted as intent-to-treat effects. The coefficient δ can be interpreted as the elasticity of pre-tax wage earnings with respect to the individual net-of tax rate: $\delta = \frac{d \log z}{d \log(1-\tau)}$. It is reported in the second row of Table 3. Again, in the traditional labor supply model, this elasticity is positive while it is negative in the bargaining model.

Consistent with Figure 8, column (1) of Table 3 shows a significantly negative effect of the scheme on earnings of -4.8 log-points which translates into an elasticity $\delta = \frac{d \log z}{d \log(1-\tau)}$ of pre-tax earnings with respect to the net-of-tax rate of -.18. Column (1) clusters standard errors at the group×year level. Because with only 46 clusters, robust standard errors might not be fully accurate, we use a grouped estimator in column (2) where we collapse all observations at the group×year level to obtain more conservative standard errors based on this aggregated sample of 46 observations. The point estimates are naturally the same as in column (1) but the standard errors are slightly larger. The estimates in column (2) however remain highly significant with a t-statistics in excess of 3.

In column (3), we add individual controls for age, citizenship, and 27-digit industry codes. This hardly affects the estimates. In column (4), we further control for potential differential trends in log earnings before the reform between the control and treatment group.[31] This leads to a larger but less precise estimate (significant only at the 5% level). Column (5) adds to

[31]More precisely, we begin by regressing individual log-earnings on two linear time trends $t \cdot \mathbb{1}[z_{it} \geq \bar{z}]$ and $t \cdot \mathbb{1}[z_{it} < \bar{z}]$ for all observations before 1991, and then estimate our regression specifications (3) and (4) using as an outcome the difference between actual earnings and predicted earnings from the initial regression.

the sample individuals with earnings between 95% and 105% of the threshold. The estimate is larger than in column (1) and remains highly significant, with an implied elasticity of -.30.

The last column of Table 3 presents a robustness check by estimating a triple-difference model where foreigners with more than 3 years of presence are used as a control.[32] This again translates into a significant estimate of -9.4 log-points and an implied elasticity of -.30.

A potential confounder for the findings from Figure 8 and Table 3 is that the set of foreigners who arrive after the scheme is in place might be different as the scheme induces a very large migration response as we showed in Section 4.1. For example, if scheme induced immigrants have lower skills and earnings than immigrants who would have come absent the scheme (conditional on having earnings above 105% of the eligibility threshold), then the estimates reported in Table 3 would be biased downward. Therefore, to control for such compositional effects, we next turn to an alternative test using a balanced panel of migrants who stay five or more years in Denmark.

Panel evidence. The scheme elapses after 3 years, producing a large tax increase among scheme beneficiaries who stay in Denmark more than 3 years. This allows us to estimate the effects of the scheme on earnings while controlling for individual fixed effects.

Figure 9 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, etc. Earnings are normalized by year 2 earnings. The sample includes all foreigners who stay 5 or more full calendar years in Denmark and have gross earnings in year 1 between 70% and below 95% of the scheme eligibility threshold in dashed line (this is the control group never eligible for the scheme) and between 105% and 150% of the scheme eligibility threshold in solid line (this is the treatment group eligible for the scheme after enactment). We exclude those earning between 95% and 105% of the scheme eligibility threshold as bunching at the eligibility threshold naturally biases downward average earnings.

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. If we were not controlling for selective attrition, the effects of the

[32] More precisely, in year t, we include foreigners who arrived in Denmark in year $t-3$ or before (and therefore are not eligible for the scheme in year t) with earnings in year t in the relevant control and treatment ranges.

variation in net-of-tax rates at year 3 on observed average earnings would be mixing two effects: the potential direct effects of the net-of-tax rates on the distribution of earnings, and a selection effect through the effect of the net-of-tax rate on the reservation wage. Because we focus on a fixed sample of stayers, we get rid of the selection effect, and variations in the earnings distribution in year 3 cannot be attributed to variations in the distribution of unobserved fixed-effect characteristics correlated with the net-of-tax rate. The estimated treatment effect of the change in tax rate at year 3 remains a local average treatment effect though, in the sense that it is estimated on the population of stayers, an important external validity limitation to keep in mind.

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 (and inconsistent with the standard labor supply 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 re-timing 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 vs. 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. This confirms 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.

$$\log z_{id} = \alpha_i + \beta \cdot 1[z_{i,d=1} \geq \bar{z}] + \gamma_d + \eta \cdot 1[z_{i,d=1} \geq \bar{z}] \cdot 1[d \geq 3] + \nu_{id}, \tag{5}$$

where z_{id} are earnings of individual i in years $d = 1, .., 5$, α_i is an individual fixed effect, γ_d is a year fixed effect, and $1[z_{i,d=1} \geq \bar{z}] \cdot 1[d \geq 3]$ 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:

$$\log z_{id} = \alpha_i + \beta \cdot 1[z_{i,d=1} \geq \bar{z}] + \gamma_d + \delta \cdot \log(1 - \tau_{id}) + \nu_{id}, \tag{6}$$

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

The estimates corresponding to specifications (5) and (6) are presented in the first and second rows of Table 4 respectively. Consistent with Figure 9, Panel A, column (1) of Table 4 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 comparable in magnitude to the coefficients estimated using repeated cross-sections in Table 3. Consistent with Figure 9, Panel B, column (3) of Table 4 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).[33]

While the panel analysis controls for individual fixed-effects, it is conceivable that individuals earnings differ not only in level but also in profile across individuals. Such profile effects cannot be controlled solely with individual fixed effects. If selection into staying in Denmark was based on earnings *profiles* (and not only on earnings *levels*), then treated and control groups in our sample of stayers might 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 pre-tax earnings raises when the scheme elapses. Such a scenario seems implausible to us because pay for such high top earners is set by contract in advance of realized productivity. Furthermore, Figure 9 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.

Nevertheless, to alleviate this concern, we estimate in Table 4 two-step Heckman models where the second stage is the same fixed effect model as in (6), but where we use citizenship

[33]For the placebo elasticity estimate, we again assume that the group above the eligibility threshold would have benefitted from the scheme in their first 3 years of stay when computing the average tax rates.

and imputed average tax rates in the home country at the time scheme elapses[34] as exclusion restrictions in the selection equation. Those estimates are reported in columns (2) and (4) of Table 4 for post-scheme entrants and pre-scheme entrants respectively. Even though results suggest that we cannot reject that selection is partially correlated with wage profiles, this correlation does not affect at all the estimates of the effect of scheme lapse, which confirms that the effect of the three-year discontinuity on pre-tax earnings cannot be attributed to selection effects only.

4.3 Spillover Effects

An important rationale put forward by Danish policy makers as well as other European governments which have adopted similar preferential tax schemes for highly paid foreigners (see OECD 2011) is that highly skilled workers generate positive externalities on their co-workers and the economy at large, over and above the fiscal effects we have described. At the same time, people often worry that these highly paid foreigners might displace or reduce the earnings of native workers. The Danish scheme creates an exogenous influx of highly skilled workers in the Danish economy, concentrated among specific industries and firms and it is worth investigating whether we can find evidence of negative or positive spillovers of the scheme on highly paid native workers, at the industry level and/or at the firm level. Unfortunately, we do not have reliable measures of individual productivity, nor do we have data on firms' profits before 1998. We therefore restrict our attention to the effect of the scheme on the distribution of wage earnings of high-skilled natives, defined as all native workers paid at least 75% of the scheme eligibility threshold.[35] This group represents approximately the top 3% of the native earnings distribution.

Spillover effects on the earnings distribution of high-skilled natives might arise from two broad classes of phenomena: productivity spillovers and incidence effects. The analysis of spillovers from human capital through knowledge diffusion has a long tradition in economics dating back to Marshall (1890) and has been used extensively in theoretical endogenous growth

[34]We use average tax rates computed using the *OECD Taxing Wages* calculator for a single individual with taxable income equivalent to 167% of average earnings in the home country.

[35]Native workers paid slightly below the scheme eligibility threshold and native workers paid slightly above are close substitutes, and there is no reason for scheme workers to have a spillover impact solely on natives paid above the scheme eligibility threshold. Our findings that the scheme had no spillover effects are however not sensitive to how exactly we define the group of high-skilled natives.

models. There is also substantial work in the urban economics literature trying to uncover such productivity spillover effects empirically (see e.g., Moretti, 2004). There is also a large literature on the effects of immigrants on native wages (see e.g., Borjas, 1999 for a survey). While productivity spillovers are expected to be positive, incidence effects would generally play in the opposite direction. In a standard model with downward-sloping labor demand due to diminishing returns to labor, the influx of high-skilled workers could reduce the demand for native high-skilled workers and depress high skill wages (for the benefit of other complementary factors in production).

Importantly, even if the migration elasticities we have estimated are large, because the fraction of foreigners at the top of the Danish distribution is relatively modest (see Figure 5), the influx of foreigners is small relative to the total stock of highly paid workers in Denmark. Therefore, we cannot expect to find very large spillover effects, whether the channel for spillovers is the standard equilibrium incidence effect or complementarities in the production function.

Industry level spillovers. At the industry level, we test whether industrial sectors which used scheme workers intensively experienced more or less growth in the average earnings of the highly paid native individuals, and in the number of highly paid native individuals, where highly paid for natives is defined as having earnings above 75% of the scheme eligibility threshold. To be concise, we refer to this group as *top natives* in what follows.

Panel A of Figure 10 plots the average growth of the real gross average earnings of top natives from the period 1980-1990 (on the y-axis) to the period 1991-2000 (on the x-axis) against the scheme use across industrial sectors. Our measure of scheme use is the growth of the number of top foreign employees (defined as foreign employees with less than 3 years of presence in Denmark and annualized earnings above the scheme eligibility threshold) between 1980-1990 and 1991-2000 divided by the total number of top native employees in the period 1980-1990. This scaling is useful to assess the relative size of the influx of scheme workers relative to the stock of top natives. The average across industries is around 2%, with substantial variation across industries. The size of the dots is proportional to top native employment in the period 1991-2000. As expected, some sectors such as telecommunications, or sports have used the scheme more intensively.[36] Most importantly, the graph shows that there is almost no correlation between

[36]Kleven, Landais, and Saez (2013) have shown that the Danish scheme increased dramatically the fraction of foreign players in the Danish football first league relative to comparable countries.

scheme use and the wage growth of top natives at the industry level, implying that the scheme has produced no visible spillover effects, either positive or negative, on the wages of top natives.

To formalize this test, Panel A of Table 5 presents the corresponding industry level regressions. It estimates the effects of the scheme use on the log of the average real gross earnings of top natives (cols. 1-2) and on the number of top natives (cols. 3-4). Column (1) presents the OLS regression of the difference in log average real gross earnings of top natives between 1980-1990 and 1991-2000 for each industry on the difference in the log number of top foreign employees (defined as foreigners with annualized earnings above the scheme eligibility threshold) by industry between 1980-1990 and 1991-2000. Consistent with Panel A of Figure 10, the coefficient is small and insignificant. Column (2) instruments for the difference in log number of top foreign employees using the log initial number of top foreign employees by industry. There is a strong first stage but the IV coefficient remains small and insignificant. Both regressions control for the log initial number of top natives in the industry and the log initial size of the industry (measured by total number of workers).

Column (3) and (4) repeat the same specifications using the log number of top natives as an outcome. In that case, the coefficients are strongly positive and significant: Industries which experienced a surge in the number of highly paid foreigners also experienced a surge in the number of top natives. Obviously, this correlation cannot be interpreted as causal, as growing industries (and particularly growing industries that employ high-skilled workers) will naturally hire both more top natives and top foreign earners. The instrument is unlikely to eliminate bias in this case because having many top foreigners before the scheme is enacted could be a good predictor of the dynamism of the industry and hence its subsequent high-skill employment growth for natives.

Firm level spillovers. At the firm level, we compare firms who use scheme workers to firms who never use scheme workers and assess whether they differentially change the average pay of their top native workers or the number of their high-skill native workers. As above, top natives are defined as Danish workers with earnings above 75% of the scheme eligibility threshold.

Panel B of Figure 10 plots the evolution of average real gross earnings of top natives in a balanced panel of firms active in all years between 1986 and 1996, broken down according to scheme use. Scheme participating firms are firms with at least one scheme employee between 1991 and 1996. The figure shows that average earnings of top natives follow almost exactly the

same pattern both before and after the scheme introduction among firms which use the scheme and among firms which never use the scheme. This is consistent with the scheme having no spillover effects on the pay of high-skill natives, consistent with our industry level analysis.

Panel B of Table 5 presents the corresponding firm-level regressions. The first row of estimates are reduced form estimates where we regress the difference in log average real gross earnings of top native employees in the firm between 1986-1990 and 1991-1996 (cols. 1-3) or the difference in the log number of top native employees in the firm between 1986-1990 and 1991-1996 (cols. 4-6) on an indicator for scheme participation (having at least one scheme employee between 1991 and 1996). For the second row of estimates, we regress the difference in log average real gross earnings of top native employees (or log number of top natives) in the firm on the difference in the log number of top foreign employees between 1980-1986 and 1991-1996. This estimates the elasticity $d\log y/d\log(\text{Top foreigners})$ with y the outcome of interest. All specifications control for firm size, industry, initial average earnings in the firm for all non-scheme employees, and initial number of top earners.

For each of the two outcome variables (log average earnings of top native workers, and log number of top native workers), the first column is a simple OLS regression, the second column presents a matching estimator using the Mahalanobis distance based on the same controls as in the OLS regression. The third column is an IV regression where we instrument the right-hand-side variable using the log initial number of top foreign employees in the firm (as shown in Table 5, the partial R-squared from the first stage are reasonably large which proves that the first stage is strong).

For the outcome log average earnings of top natives in columns (1)-(3), consistent with Panel B of Figure 10, we find small and insignificant effects in all three specifications.[37] This suggests again that the scheme had no spillover effects on the wages of top natives.

For the outcome log number of top natives, both the OLS and the matching estimators are large and significant, implying that firms using the scheme experience a growth in the number of top natives they employ. As in the case of Panel A, we cannot interpret this correlation as causal as growing firms will naturally start to employ both more native and foreign highly paid employees. Interestingly however, the instrumented estimates in column (6) are not significant.

[37]Note that the Durbin-Wu-Hausman test of exogeneity is not rejected in the case of log average earnings, which suggests that the endogeneity problem is not very important, contrary to the case of log number of top natives in column (6), where the test supports the presence of endogeneity.

Firms with more top foreign employees before the scheme are more likely to employ scheme workers (first stage) but they are not more likely to employ more top native workers. In that case, the exogeneity assumption for the instrument is more plausible at the firm level than at the industry level in Panel A. Those insignificant coefficients are therefore our best evidence suggesting that the scheme had no impact on the number of top native employees at the firm level.

To summarize, while the spillover evidence is at best suggestive given the lack of fully compelling identification, our findings suggests that scheme workers did not have strong spillover effects (either negative or positive) on the number of top native employees, and perhaps more convincingly on the average earnings of top native employees.

5 Conclusion

Our paper 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 a unique 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 three 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. Hence, preferential tax schemes for highly paid foreign workers could create severe tax competition across European countries. Second, we find compelling evidence of a negative effect of scheme-induced increases in the net-of-tax rate on pre-tax earnings at the individual level. This finding cannot be explained by the standard 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. Third, we find no evidence of positive or negative spillovers of the scheme-induced influx of highly skilled foreigners on the earnings of natives, but the identification of such spillovers is in general very difficult even with the sharp tax variation in our data.

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 workers are likely to generate tax competition across European countries by limiting the ability of European governments to use progressive taxation. This will require international coordination and the design of rules regulating such special schemes.

In future work, we hope to make progress on estimating spillovers of foreign workers on native co-workers using specifically the researcher part of the scheme that we have not analyzed in this paper. We plan to analyze the effect of the influx of researchers in Denmark on the patents, publications, and placement of Ph.D. candidates of Danish research centers vs. comparable European countries such as Sweden. The comparison could be done across academic fields using the pre-scheme fraction of foreign researchers across fields.

References

Bakija, Jonathan and Joel B. Slemrod (2004). "Do the Rich Flee from High State Taxes? Evidence from Federal Estate Tax Returns," NBER Working Paper No. 10645.

Blundell, Richard and Thomas MaCurdy (1999). "Labor Supply: A Review of Alternative Approaches," in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, Volume 3A, North-Holland, Amsterdam.

Borjas, George J. (1999). "The economic analysis of immigration." In: Orley C. Ashenfelter and David Card, Ed(s), *Handbook of Labor Economics*, Elsevier: Amsterdam, Volume 3A.

Chetty, Raj, John Friedman, Tore Olsen, and Luigi Pistaferri (2011). "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records," *Quarterly Journal of Economics* 126(2): 749-804.

Feldstein, Martin. (1999). "Tax Avoidance and the Deadweight Loss of the Income Tax," *Review of Economics and Statistics* 81(4): 674-680.

Fullerton, Don and Gilbert Metcalf (2002). "Tax Incidence," in A. Auerbach and M. Feldstein (eds.), *Handbook of Public Economics*, Volume 4, North-Holland, Amsterdam.

Hall, Robert E. (2005). "Employment Fluctuations with Equilibrium Wage Stickiness," *American Economic Review* 95(1), 50–65.

Howitt, Peter, and R. Preston McAfee. (1987). "Costly Search and Recruiting." *International Economic Review* 28(1): 89–107.

Immervoll, Herwig, Henrik J. Kleven, Claus T. Kreiner, and Emmanuel Saez (2007). "Welfare Reform in European Countries: A Microsimulation Analysis." *Economic Journal* 117, 1-44.

Kanbur, Ravi and Michael Keen (1993). "Tax Competition and Tax Coordination When Countries Differ in Size," *American Economic Review* 83(4), 877-892.

Keen, Michael and Kai Konrad (2013). "The Theory of International Tax Competition and Coordination," in A. Auerbach, R. Chetty, M. Feldstein, and E. Saez (eds.), *Handbook of Public Economics*, Volume 5, North-Holland, Amsterdam.

Kirchgassner, Gebhard and Werner Pommerehne (1996). "Tax harmonization and tax competition in the European Union: Lessons from Switzerland," *Journal of Public Economics* 60, 351-371.

Kleven, Henrik J., Camille Landais, and Emmanuel Saez (2013). "Taxation and International Migration of Superstars: Evidence from the European Football Market," forthcoming *American Economic Review* 103

Kleven, Henrik J. and Esben A. Schultz (2012). "Estimating Taxable Income Responses using Danish Tax Reforms." LSE Working Paper, Revised May 2012.

Kleven, Henrik J. and Mazhar Waseem (2013). "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan." *Quarterly Journal of Economics* 128(2).

Liebig, Thomas, Patrick A. Puhani, and Alfonso Sousa-Poza (2007). "Taxation and Internal Migration: Evidence from the Swiss Census using Community-Level Variation in Income Tax Rates." *Journal of Regional Science* 47(4), 807–836.

Marshall, Alfred. (1890) *Principles of economics,* New York: Macmillan.

Mirrlees, James A. (1982). "Migration and Optimal Income Taxes." *Journal of Public Economics* 18, 319-41.

Moretti, Enrico (2004). "Workers' Education, Spillovers and Productivity: Evidence from Plant-Level Production Functions." *American Economic Review* 94(3).

Mortensen, Dale and Christopher Pissarides (1999). "New Developments in Models of Search in the Labor Market," in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, Volume 3B, North-Holland, Amsterdam.

OECD. (2011). "The Taxation of Mobile High-Skilled Workers," Chapter 4 in *OECD Tax Policy Study No. 21: Taxation and Employment*, OECD, Paris.

Piketty, Thomas and Emmanuel Saez (2013). "Optimal Labor Income Taxation," in A. Auerbach, R. Chetty, M. Feldstein, and E. Saez (eds.), *Handbook of Public Economics*, Volume 5, North-Holland, Amsterdam.

Pissarides, Christopher A (2000). *Equilibrium Unemployment Theory,* 2nd ed., Cambridge, MA: MIT Press.

Saez, Emmanuel (2010). "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy* 2(3), 180-212.

Saez, Emmanuel, Manos Matsaganis, and Panos Tsakloglou (2012). "Earnings Determination and Taxes: Evidence from a Cohort Based Payroll Tax Reform in Greece." *Quarterly Journal of Economics* 127(1), 493–533.

Saez, Emmanuel, Joel Slemrod, and Seth Giertz (2012). "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." *Journal of Economic Literature* 50(1), 3–50.

Young, Cristobal and Charles Varner (2011). "Millionaire Migration and State Taxation of Top Incomes: Evidence from a Natural Experiment." *National Tax Journal* 64, 255–284.

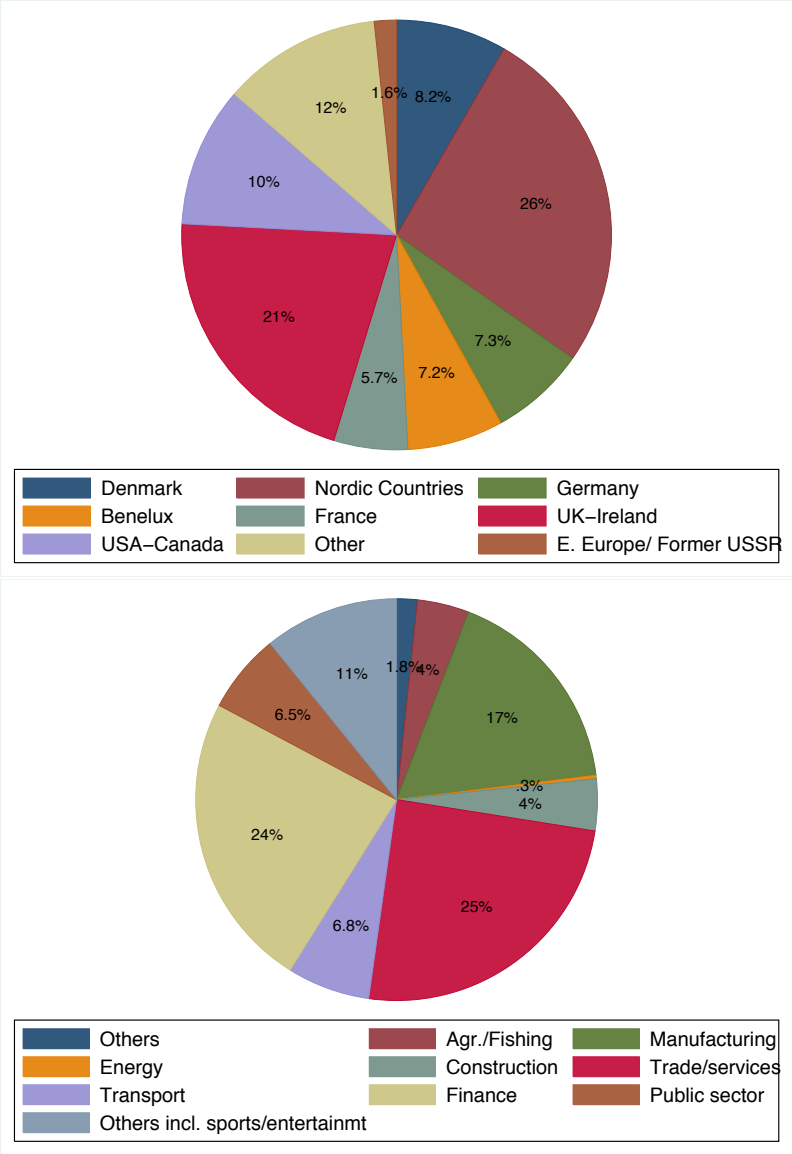


Figure 1: Citizenship and Industry Composition of Scheme Beneficiaries, 1991-2010

Notes: Panel A reports the composition of tax scheme spells (excluding researchers) by country of citizenship of the beneficiaries (at the time of scheme) across all years 1991 to 2010. Panel B reports the composition of tax scheme spells (excluding researchers) by industry across all years 1991 to 2010.

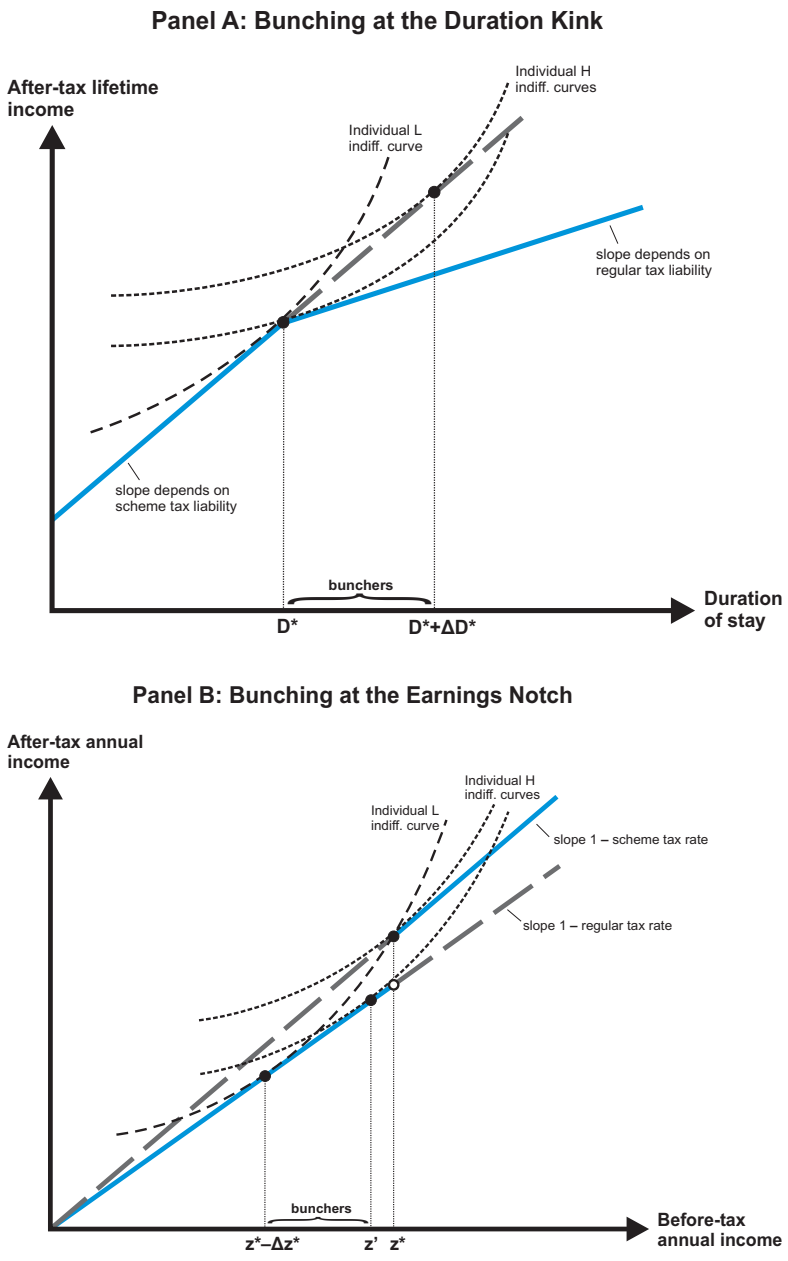


Figure 2: Responses to the Tax Scheme in the Standard Labor Supply Model

Notes: 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 3.1. Panel A depicts the kink created by the scheme in the life-time budget defined over duration D of stay in Denmark and life-time consumption. The solid blue line represents the budget under the scheme with maximum duration D^* (= 3 years). The dashed grey line represents the budget under a scheme with indefinite duration. 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 blue 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.

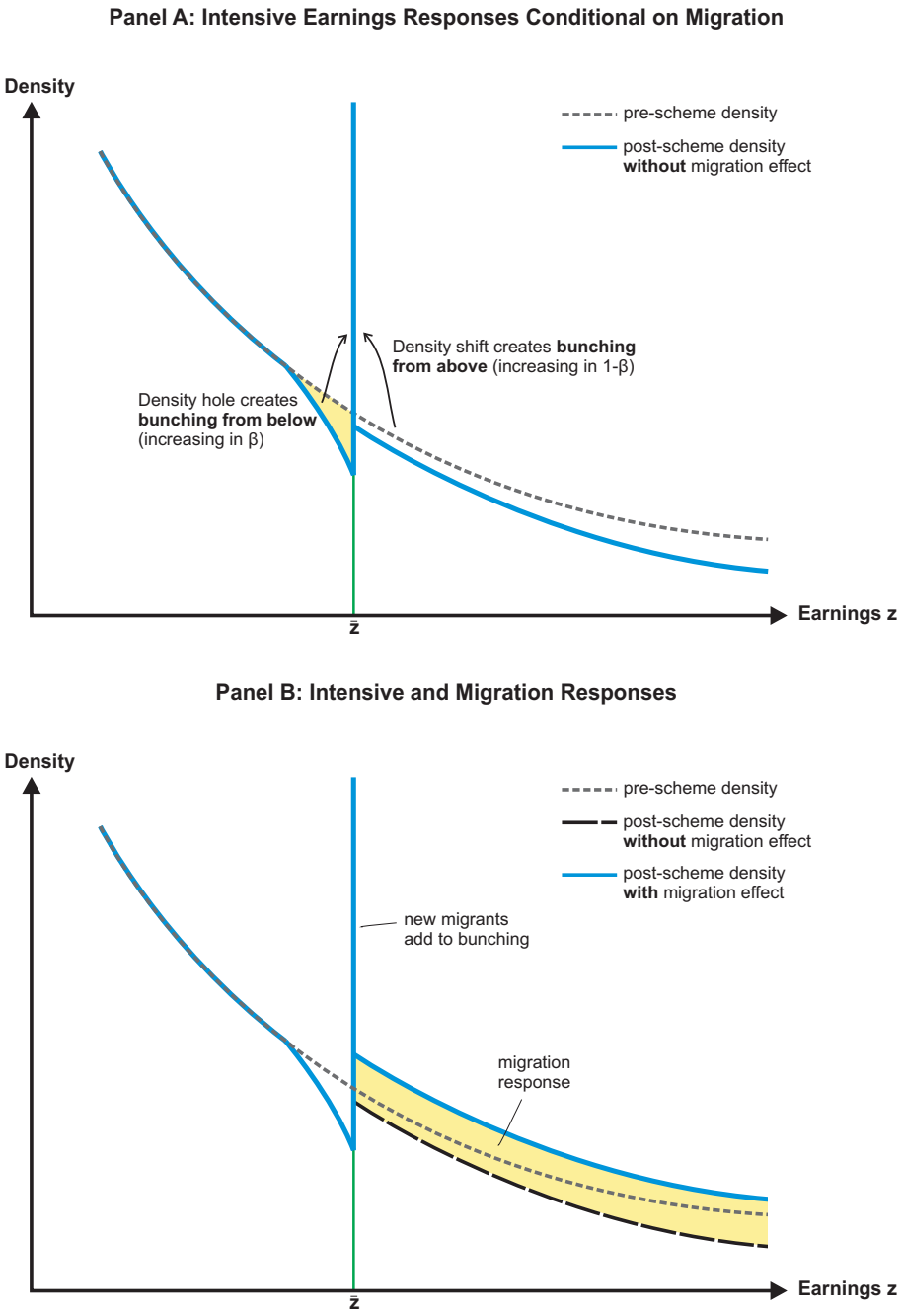
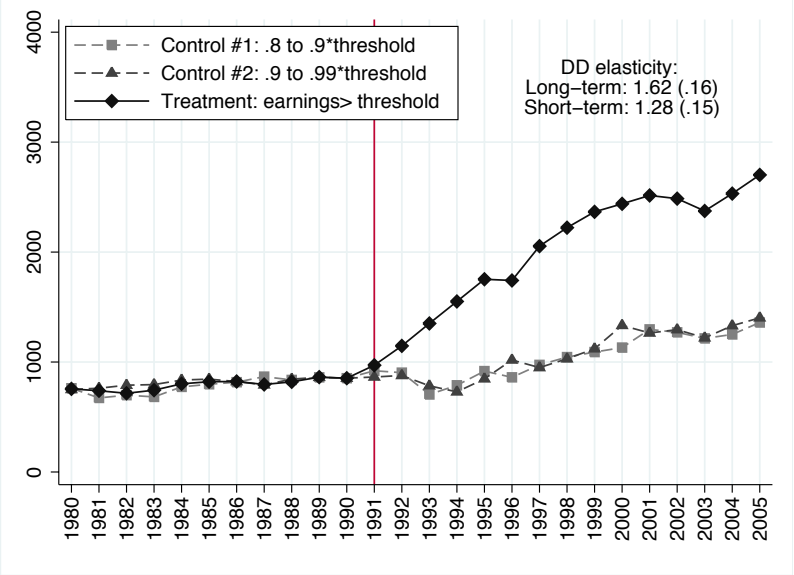


Figure 3: Responses to the Tax Scheme in the Matching Frictions Model

Notes: 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 3.2. Absent the scheme, the earnings density of immigrants is smoothly decreasing (short-dashed line). The top panel shows that the scheme reduces earnings above the eligibility threshold $\bar{z}$ if employers have positive bargaining power ($1-\beta > 0$), which shifts the density left (to the solid blue 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 $\bar{z}$. The bottom panel shows that the scheme induces migration, which shifts the distribution upward above the eligibility threshold $\bar{z}$ and adds to bunching at $\bar{z}$.

A. Number of foreigners in different earnings groups



B. Number of foreigners' arrivals

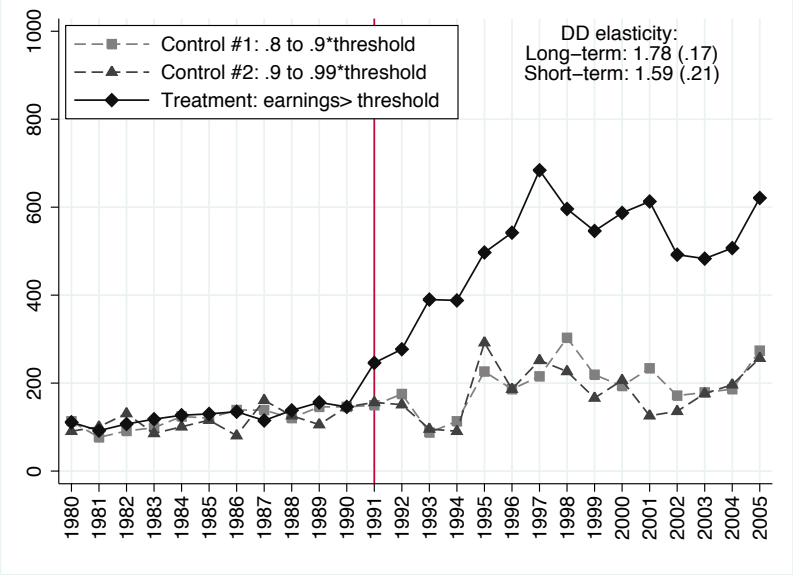


Figure 4: Migration Effects of the Tax Scheme: Diff-in-Diff Evidence

Notes: 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 1st, 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 2).

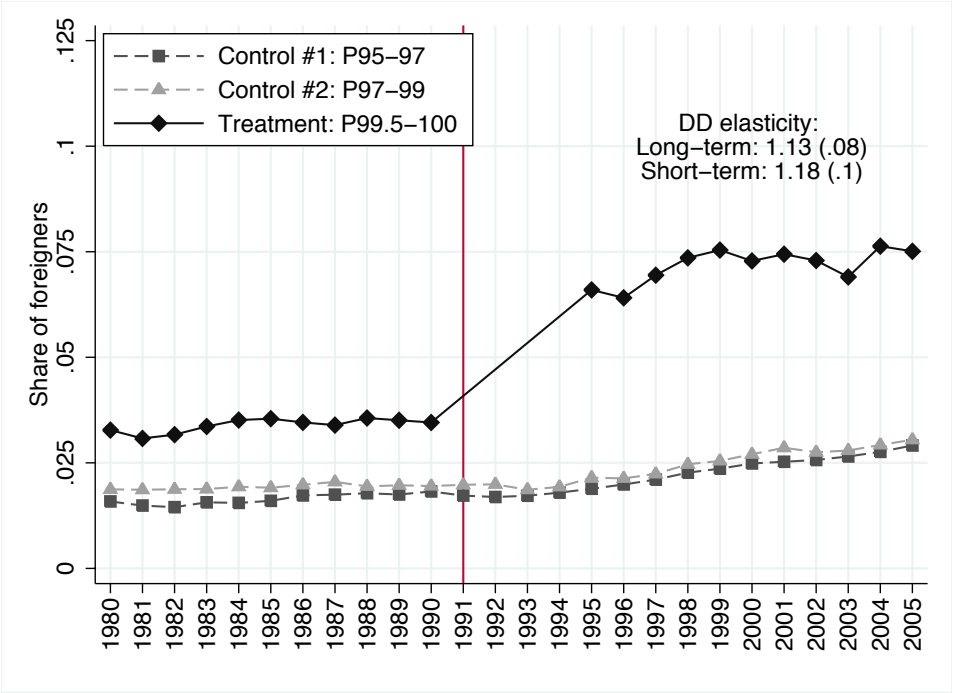


Figure 5: Fraction of Foreigners in Different Earnings Percentiles

Notes: The figure plots the fraction of foreigners in various upper percentile groups of the distribution of earnings (percentiles are defined including solely Danish citizens with positive earnings) from 1980 to 2005. The threshold for eligibility to the scheme is always between the 99.2th and the 99.4th percentile. P95-97 denotes all individuals between the 95th and 97th percentile, etc. Earnings are annualized for part-year residents. There is a gap in 1991-1994 for the top group because the data do not provide scheme earnings for those years. The vertical line at year 1991 denotes the year the scheme was first implemented. The corresponding elasticity estimates are reported (see Table 2).

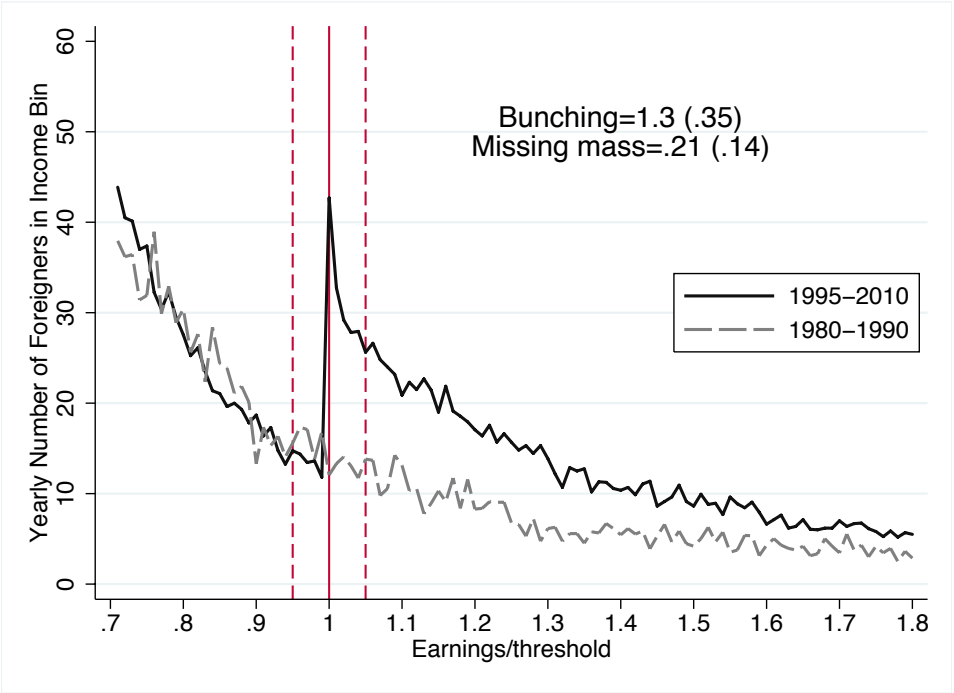


Figure 6: Earnings Density for Foreigners Before and After Scheme Introduction

Notes: The figure reports the density of the earnings distribution of foreigners around the eligibility threshold (denoted by the vertical line) in 1995-2010 (dark line after scheme implementation) and in 1980-1990 (grey line before scheme implementation). 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. Non-parametric 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 the main text for details of the estimation method).

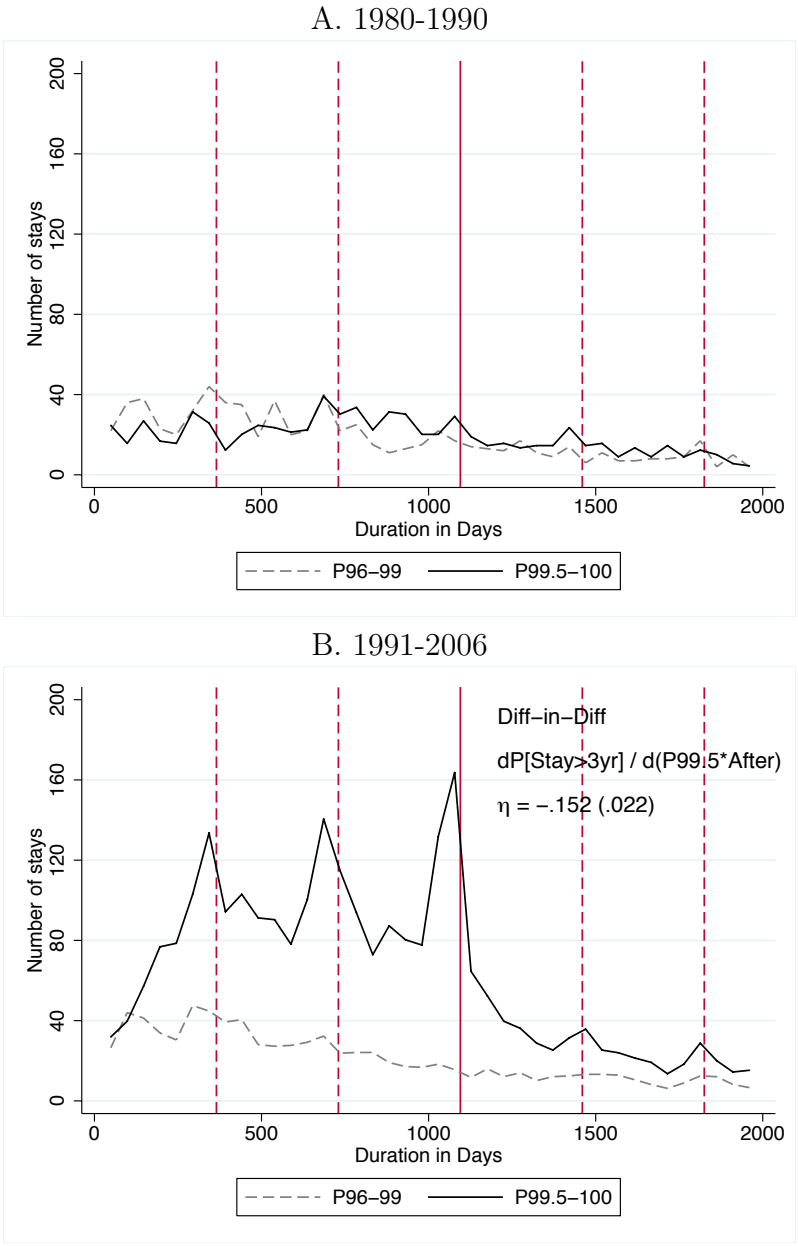


Figure 7: Density of the Duration of Stay of Foreigners

Notes: The figure reports the number of stays by duration of foreigners with (annualized) earnings above percentile 99.5th (P99.5-100) in solid line and earnings between percentile 96 and percentile 99 (P96-99) in dashed line. The P99.5-100 group is always above the eligibility threshold for the scheme while 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) while Panel B is for years 1991-2000 (after the scheme was implemented). Vertical lines demarcate year thresholds, with the solid vertical line representing the 3-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). 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 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 for P99.5-100 and the Panel B normalization 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.

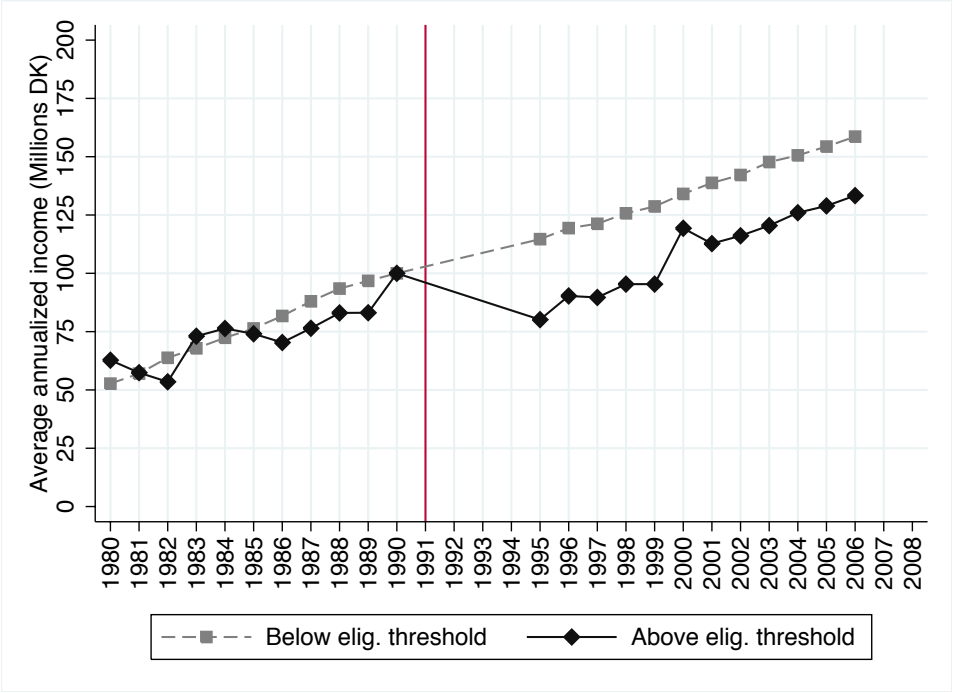
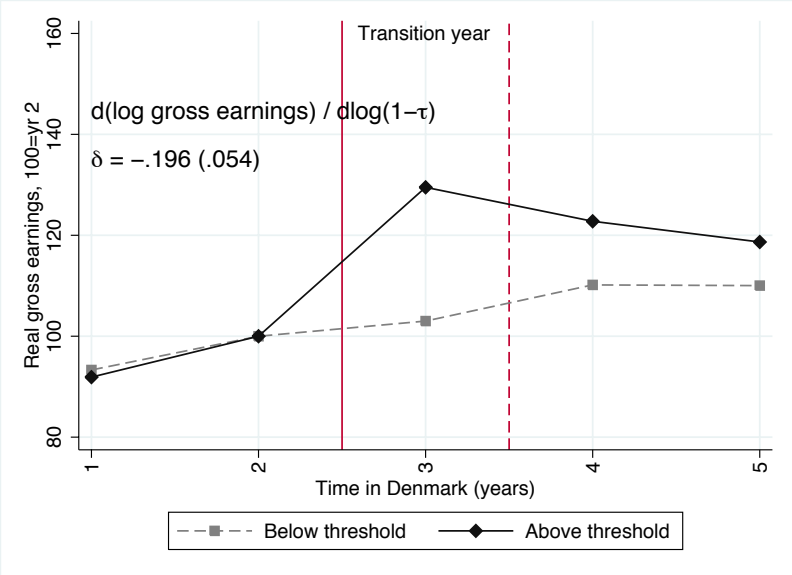


Figure 8: Effects of the Tax Scheme on Pre-tax Earnings: Repeated Cross-Section Evidence

Notes: The figure depicts the average *nominal* annual earnings for foreigners in their first full calendar year in Denmark (denoted by year 1) for the sample of foreigners with year 1 earnings between 70% and below 95% of the scheme eligibility threshold (dashed line) and between 105% and 400% of the scheme eligibility threshold (solid line). Both series are normalize to 100 in year 1990, the year 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 $t-1$, stay the full calendar year t in Denmark. Data for years 1991-1994 are not available. The graph shows that the earnings of foreigners above the scheme eligibility threshold decrease after the scheme is in place, suggesting that the scheme reduces pre-tax earnings. This is consistent with the matching friction model (and inconsistent with the standard labor supply model).

A. After introduction of the scheme



B. Before introduction of the scheme

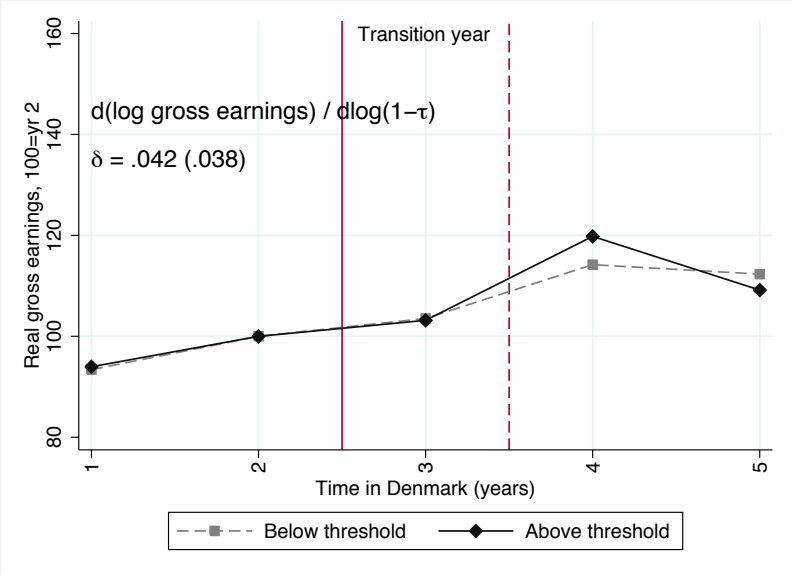
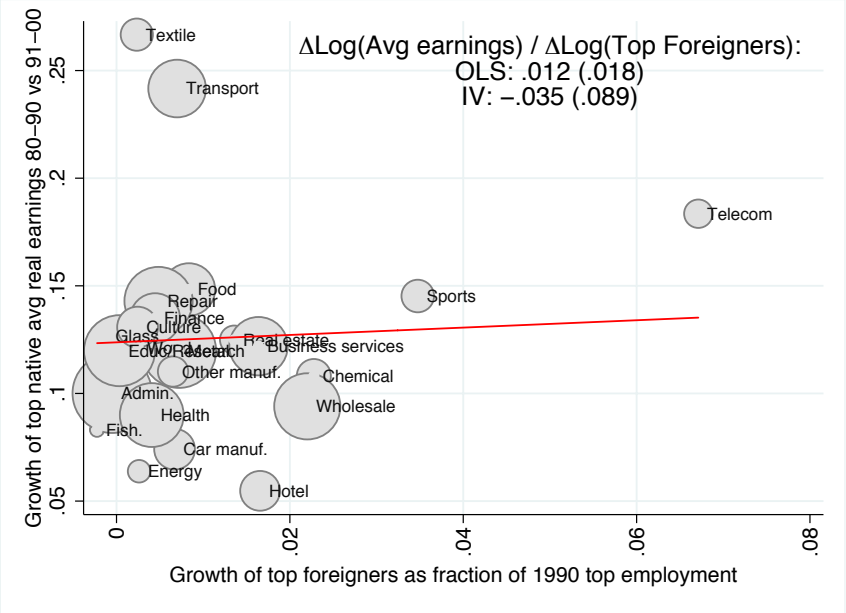


Figure 9: Effects of the Tax Scheme on Pre-tax Earnings: Panel Evidence

Notes: The figure depicts the average 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, etc. 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 70% and below 95% of the scheme eligibility threshold (dashed line) and between 105% and 150% 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 (and inconsistent with the standard labor supply 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 re-timing to maximize scheme benefits. Panel B is a placebo comparison (for entrants before the scheme was introduced).

A. Industry level



B. Firm level

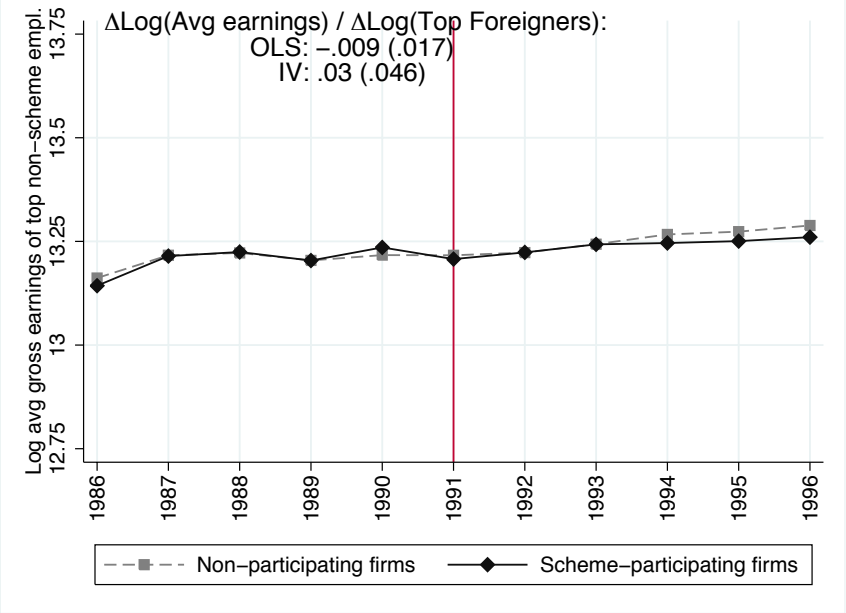


Figure 10: Evidence on Spillovers Effects of the Scheme

Notes: Panel A plots the average growth of the real gross earnings of top native employees from the period 1980-1990 to the period 1991-2000 against the scheme use across industrial sectors. Top native employees are defined as employees with earnings above 75% of the scheme threshold (which corresponds roughly to the top 4% of the distribution of earnings among the natives.) Our measure of scheme use is the growth of the number of top foreign employees (defined as foreign employees with less than 3 years of presence in Denmark and annualized earnings above the scheme eligibility threshold) between 1980-1990 and 1991-2000 divided by the total number of top native employees in the period 1980-1990. The size of the dots is proportional to top native employment in the period 1991-2000. The graph shows that there is almost no correlation between scheme use and the wage growth of top native employees at the industry level. Panel B plots the evolution of average real gross earnings of top native employees (defined as above) in a balanced panel of firms alive between 1986 and 1996, broken down according to scheme use. Scheme participating firms are firms with at least one scheme employee between 1991 and 1996.

Table 1: Descriptive Statistics

	(1) mean	(2) standard deviation
	Scheme employees 1991-2010 Number of scheme spells=11642	
Fraction with proof of residence	.949	.220
Duration of stay (years)	2.345	1.448
fraction > 3 years	*.251*	*.434*
Scheme earnings (2009 DK1000s)	1217.8	2094.7
as a fraction of threshold	*1.53*	*2.67*
Average tax rate	.308	.003
Age	39.97	8.80
	Scheme take-up rate: .81	
	Firms 1991-2010 N=2235	
Number of employees	438.06	2316.47
average gross wage (2009 DK1000s)	437.6	305.0
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 always 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% of spells, we were able to find a proof of residence from the registry files. All other statistics in the table (and in the subsequent analysis of the paper) are restricted to spells with proof of residence. The average tax rate is 30% before 1995 and 25%*(1-AMB)+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 of scheme-participating firms (including scheme employees is 437,000 DK2009.

Table 2: Migration Elasticity Estimates

	(1) Total number of foreigners	(2) Number of arrivals	(3) Number of 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***	1.779***	2.049***
	(0.162)	(0.168)	(0.148)
ε_{st} (short-term)	1.280***	1.590***	1.756***
	(0.151)	(0.228)	(0.170)
A2. Control for pre-existing trends			
ε_{lt} (long-term)	1.756***	1.771***	2.152***
	(0.176)	(0.168)	(0.158)
A3. Placebo			
ε_{lt} (long-term)	-0.0602	-0.0101	0.0796
	(0.0823)	(0.245)	(0.161)
A4. Control for imperfect take-up (IV)			
ε_{lt} (long-term)	2.892***	1.945***	2.392***
	(0.232)	(0.167)	(0.138)
A5. Nordic countries			
ε_{lt} (long-term)	1.442***	1.805***	2.208***
	(0.166)	(0.287)	(0.257)
A6. English-speaking countries			
ε_{lt} (long-term)	1.852***	2.186***	2.281***
	(0.222)	(0.246)	(0.206)
A7. Danish expatriates			
ε_{lt} (long-term)	0.0185	-0.0913	-0.0998
	(0.0280)	(0.0708)	(0.0613)
	Panel B: Treatment: percentile 99.5-100, Control: percentile 95-99		
B1. Baseline			
ε_{lt} (long-term)	1.133***	1.015***	1.257***
	(0.0802)	(0.169)	(0.109)

Notes: Robust standard errors in parentheses, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. 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 3 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 post-reform period. We always 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 pre-existing 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, UK, US. 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 percentile of the earnings distribution of natives, and the treatment group is all foreigners with (annualized) earnings above the 99.5th percentile.

Table 3: Repeated Cross-section Estimates of the Effects of the Tax Scheme on Pre-tax Earnings

	(1)	(2)	(3)	(4)	(5)	(6)
	DD: Less than 3 years of presence					DDD: More than 3 years as control
		Grouped estimator			*With bunchers*	
Reduced form estimate	-0.0476***	-0.0476**	-0.0566***	-0.130*	-0.0951***	-0.0946***
	(0.0103)	(0.0138)	(0.0117)	(0.0535)	(0.0103)	(0.0205)
Elasticity $\frac{d\log z}{d\log(1-\tau)}$ estimate	-0.180***	-0.180**	-0.171***	-0.390*	-0.296***	-0.297***
	(0.0371)	(0.0498)	(0.0342)	(0.156)	(0.0345)	(0.0604)
Industry, Age, Citizenship			×	×	×	×
Differential time trends				×		
N	18518	46	18518	18518	21245	22382

Notes: Robust standard errors clustered at the group×year level in parentheses, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. The table presents differences-in-differences estimate of the effect of the scheme on log annual pretax earnings. For year t, the sample includes foreigners who arrive during year $t-1$, stay the full calendar year t in Denmark, and stay for a duration of less than 3 years in Denmark. The sample includes years 1980-2006 but excluding years 1991-1994 (data for years 1991-1994 are not available). The control group are foreigners with earnings between 70% and 95% of the scheme eligibility threshold while the treatment group are foreigners with earnings above 105% of the scheme eligibility threshold. In all columns, except (5) we exclude potential bunchers by removing all individuals with earnings between 95% and 105% of the threshold. The first row reports the effect of the scheme is given by the interaction between having earnings above the scheme eligibility and having entered Denmark after 1991 (specification (3) in the main text). The second row reports the corresponding elasticity estimate obtained with a 2SLS regression (specification (4) in the main text). We cluster standard errors at the group×year level. Because with 46 clusters, inference can be problematic, we use a grouped estimator in column (2) where we collapse all observations at the group×year level. In column (3) we add controls for age, citizenship and 27-digit industry codes. In column (4) we control for potential differential time trends in log earnings before the reform between the control and treatment group. Column (6) estimates a triple-difference model where foreigners with more than 3 years of presence are used as a control.

Table 4: Panel Estimates of the Effects of the Tax Scheme on Pre-tax Earnings

	(1)	(2)	(3)	(4)
	Treatment		**Placebo**	
	Entry after 1991		**Entry 1980 to 1990**	
	OLS	*Heckman*	*OLS*	*Heckman*
	(stayers only)	*2-step*	*(stayers only)*	*2-step*
Reduced form estimate	0.0925***	0.0829**	-0.0229	0.0314
	(0.0256)	(0.0265)	(0.0228)	(0.0245)
Elasticity $\frac{d\log z}{d\log(1-\tau)}$ estimate	-.196***	-.176**	.049	-.067
	(.054)	(.056)	(.048)	(.052)
Exclusion restrictions:		Average tax rate in home country + citizenship		Average tax rate in home country + citizenship
N	2943	5616	2341	3508
λ		0.187		0.152
		(0.0149)		(0.0162)
ρ		0.824		0.747
LR test of independence		χ^2= 118.5		χ^2= 91.35
		Prob > χ^2= .00		Prob > χ^2= .00

Notes: Robust standard errors clustered at the individual level in parentheses, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. The table presents panel estimates of the effect of scheme lapse on pre-tax earnings using the reduced form specification (5) in row 1 and the 2SLS specification (6) 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 3 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 average tax rates in the home country at the time scheme elapses and citizenship dummies as exclusion restrictions in the selection equation for staying more than 3 years. λ is the estimated inverse Mills ratio and ρ is the estimated correlation between the error terms in the first and second stage equations. Column (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.

Table 5: Regression-Based Estimates of Spillovers Effects of the Scheme

A. Industry Level

	(1)	(2)	(3)	(4)
	Log average earnings of top non-scheme employees		*Log number of top non-scheme employees*	
	OLS	IV	OLS	IV
$\frac{d\log y}{d\log(\text{Top foreigners})}$	.012	-.035	.471	.564
	(.018)	(.089)	(.07)	(.312)
N	27	27	27	27

B. Firm Level

	(1)	(2)	(3)	(4)	(5)	(6)
	Log average earnings of top non-scheme employees			*Log number of top non-scheme employees*		
	OLS	Matching	IV	OLS	Matching	IV
Reduced form estimate	-.002	.018	-.028	.187	.163	-.006
	(.013)	(.017)	(.045)	(.019)	(.062)	(.051)
$\frac{d\log y}{d\log(\text{Top foreigners})}$	-.009	.036	.03	.573	.32	.012
	(.017)	(.033)	(.046)	(.022)	(.122)	(.104)
Post-estimation						
Durbin-Wu-Hausman test of endogeneity		$F(1,2825)=1.0799$ Prob $>F=.299$			$F(1,26938)=31.214$ Prob $>F=.00$	
Shea partial-R-square			.0941			.0446
N	2852	2852	2852	26966	26966	26966

Notes: Robust s.e. in parentheses. Panel A presents industry level regressions. It estimates the effects of the scheme use on the log of the average real gross earnings of top native employees (cols. 1-2) and on the number of top native employees (cols. 3-4). Top native employees are defined as employees with earnings above 75% of the scheme threshold (which corresponds roughly to the top 3% of the distribution of earnings among the natives.) Column (1) regresses the difference in log average real gross earnings of top native employees between 1980-1990 and 1991-2000 for each industry on the difference in the log number of top foreign employees (defined as foreigners with annualized earnings above the scheme eligibility threshold) by industry between 1980-1990 and 1991-2000. Column (2) instruments for the difference in log number of top foreign employees using the log initial number of top foreign employees by industry. Column (3) and (4) repeat the same specifications using the log number of top native employees as an outcome. All specifications control for the log initial size of the industry, and in column (1) and (2) for the log initial number of top employees in the industry.

Panel B repeats the same strategy at the firm level. The estimation sample is a balanced panel of all firms active in Denmark in all years between 1986 and 1996. Reduced form estimates regress the difference in log average real gross earnings of top native employees in the firm between 1986-1990 and 1991-1996 (cols. 1-3) or the difference in log number of top native employees between 1986-1990 and 1991-1996 (cols. 4-6) on an indicator for scheme participation (having at least one scheme employee between 1991 and 1996). $d\log y/d\log(\text{Top foreigners})$ is obtained from the same regression where the indicator for scheme participation is replaced by the difference in log number of top foreign employees. In columns (2) and (5), we present a matching estimator using the Mahalanobis distance based on the same controls as in the OLS regression. In columns (3) and (6), we instrument for scheme participation (reduced form) or for the difference in log number of top foreign employees using the log initial number of top foreign employees in the firm. All specifications control for firm size, industry, initial average earnings in the firm for all non-scheme employees and initial number of top earners. The bottom part of Panel B presents results from the Durbin-Wu-Hausman test of exogeneity of our instrument and the Shea partial-R-square for estimates in columns (3) and (6).