

Taxation and International Migration of Top Earners: Evidence from the Foreigner Tax Scheme in Denmark*

Henrik Jacobsen Kleven, London School of Economics

Camille Landais, SIEPR Stanford University

Emmanuel Saez, UC Berkeley

Esben Schultz, Danish Economic Council

Preliminary Draft: November 2011

Abstract

This paper analyzes the effects of income taxation on the international migration of top earners using the Danish preferential foreigner tax scheme. This scheme, introduced in 1991, allows immigrants with high earnings (above 103,000 Euros per year as of 2009) to be taxed at a flat rate of 25% for a duration of three years instead of the regular progressive schedule with a top marginal tax rate of 59%. Using population wide Danish administrative tax data, we show 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 in excess of one. There is bunching of earnings at the scheme eligibility threshold, evidence of a significant but quantitatively very small response along the intensive earnings margin (work effort or earnings manipulation through tax avoidance). There is also evidence of sharp bunching of durations of stay at the three year duration limit which translates into a significant but quantitatively small intensive duration response. In the end, the migration elasticity is much more larger than the conventional within country elasticity of earnings with respect to the net-of-tax rate. Hence, preferential tax schemes for highly paid workers could generate very harmful tax competition across European countries and severely limit the ability of European governments to use progressive taxation.

*We would like to thank Raj Chetty, Michael Devereux, Rick Hornbeck, Caroline Hoxby, Wojciech Kopczuk, Claus Kreiner, and numerous seminar participants for helpful comments and discussions. Financial support from the Center for Equitable Growth at UC Berkeley and the European Tax Policy Forum is gratefully acknowledged.

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 tax rates on highly paid workers may induce such workers to migrate to countries where the tax burden for top earners is lower. This can limit the ability of governments to raise taxes using progressive taxation by creating harmful tax competition. This issue is further exacerbated by the development of special favorable tax treatment reserved to highly skilled or highly compensated foreigners in a number of European countries.¹ As a result, the mobility response to tax rates looms even larger in the European policy debate on tax progressivity than traditional within-country labor supply responses.

While an enormous empirical literature has studied labor supply and taxable income responses to taxation (as surveyed by, e.g., Blundell and MaCurdy 1999 for labor supply, and Saez, Slemrod, and Giertz 2011 for taxable income), there is very little empirical work on the effect of taxation on the spatial mobility of individuals, and especially mobility among high-skilled workers. A small literature has considered the mobility of people across local jurisdictions *within* countries, including Kirchgassner and Pommerehne (1996) and Liebig et al. (2007) on mobility across Swiss Cantons in response to Canton taxes, Feldstein and Wrobel (1998) and Bakija and Slemrod (2004) on mobility across U.S. states in response to state income and inheritance taxes, and Meyer (2000) on mobility across U.S. states in response to state welfare programs. However, empirical work on the effect of taxation on international mobility appears to be virtually non-existent² partly due to lack of micro data with citizenship information and challenges in identifying causal tax effects on migration. A recent exception is Kleven, Landais, and Saez (2010) 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 highly skilled workers. In that case, the significant migration responses of Kleven, Landais, and Saez (2010)

¹Besides Denmark which we will analyze in this paper, Netherlands, Portugal, Spain, Switzerland also have preferential tax schemes for foreign workers that are significant.

²There is a very large literature on the effects of capital taxation on multinational corporations and international capital mobility (surveyed by, e.g., Gordon and Hines 2002). There is also an enormous literature on wage differentials and international migration (surveyed by, e.g., Borjas 1999), and some work on how international migration is affected by the generosity of social insurance and social welfare programs (e.g. Borjas 1999b).

would not be informative about the migration responses of high-skilled workers in general.³ This paper breaks new ground on this issue by analyzing the migration effects of a special tax scheme that applies to highly compensated foreign workers in Denmark regardless of occupation or industry.

In 1992, Denmark enacted a preferential tax regime for foreign researchers and high-income foreigners in all other professions, who sign contracts for work in Denmark after June 1, 1991. Under this scheme, the income tax rate is reduced to a flat rate of 25% (30% before 1996) for a total period of up to 3 years. Except for researchers, eligibility for this tax scheme requires an annual income above a certain threshold, which was set to 511,200 Danish kroner (about 69,000 Euros) when the scheme was introduced in 1992 and indexed to average earnings growth in subsequent years. As of 2009, the earnings threshold is about 103,000 Euros corresponding roughly to the 99th percentile of the distribution of individual earnings in Denmark.⁴ This scheme is much more generous than the regular tax system, which imposes a top marginal tax rate of 59% above a relatively modest annual income of 350,000 kroner (47,000 Euros). Absent the special tax scheme, workers with earnings above the scheme threshold would face average income tax rates around 55%, more than twice as high as the scheme rate. When the 3 years of preferential tax treatment have been used up, the taxpayer will be subject to the ordinary tax scheme imposing very high tax rates.⁵

This unusual piece of tax policy implies large discontinuities in tax liability depending on the contract start date (before and after June 1, 1991), duration of stay in Denmark (3-year and 7-year rules), earnings level (earnings eligibility threshold), and occupation (researchers versus other professions). Hence, the reform creates very large quasi-experimental variation along several different dimensions, and provides a very powerful way of identifying the effect of taxation on migration. It is very unusual to have large discontinuities in tax liability (as opposed to discontinuities in marginal tax rates) in tax systems. In this paper, we explore

³In particular, top footballers are young workers who earn most of their lifetime income over a short time period, and their profession involves relatively little country-specific capital. Furthermore, Kleven, Landais, and Saez (2010) do not have earnings level data on footballers, making it difficult to control for incidence effects of taxes on wages.

⁴The scheme also applies to Danish citizens, who have been abroad with no Danish source earnings for a period of at least 3 years.

⁵In the original law, after 7 years of residence in Denmark, a worker who had benefited from the preferential tax scheme would be subject to a claw-back whereby ordinary tax rates applied retroactively to the entire income earned during the period of preferential tax treatment. For a high-income worker, this rule implied a very large retroactive tax bill after 7 years of residence. The rule was eliminated for researchers in 1998 and substantially relaxed for all other professions in 2002, so that today the retro-active tax applies to very few workers.

the different aspects of the tax scheme using quasi-experimental techniques such as bunching approaches and difference-in-differences. For this analysis, we have access to administrative data including the universe of tax and payroll records for the entire population of Danish residents (Danish citizens and foreigners) since 1980. The data includes detailed information about citizenship, immigration history, income and tax variables, labor market variables, and socio-demographic information. It also contains specific information for all scheme beneficiaries. 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).

Our analysis of the Danish tax scheme for foreigners yields two main results. First and most important, 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 almost 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. This overall effect implies that the overall elasticity of migration with respect to the net-of-tax rate is very large, in excess of one. The resulting tax revenue maximizing rate for highly paid foreigners is therefore relatively small and around 40%. Taking into account all other taxes, including the VAT and payroll taxes, this implies that the very low 25% flat rate in Denmark is actually very close to the revenue-maximizing tax rate on high-income foreigners. Hence, it can be desirable from a single country perspective to adopt such preferential schemes for highly paid foreigners. At the same time, those schemes impose negative fiscal externalities on other countries and are detrimental to global welfare. This tension between country welfare and global welfare in tax policy making has loomed large in the public debate for a long time, but our paper provides for the first time compelling evidence that this is indeed a major tax policy issue. Absent coordination, it is conceivable that many countries would start adopting and extending such schemes which could unravel tax progressivity in Europe.

Second, using a simple theoretical framework, we decompose the global elasticity described above into the three underlying margins: the extensive migration margin, the intensive duration-of-stay margin, and the intensive earnings margin which includes labor supply adjustment and earnings manipulation through tax avoidance to qualify for the scheme threshold. Overall, the extensive migration margin accounts for over 80% of the total elasticity. The notch created by the earnings eligibility threshold does create a highly visible piling up in the density of earnings

for foreigners at the eligibility threshold (relative to controls) but this piling up is quantitatively extremely small relative the very large notch in the budget set. Using the methodology developed by Kleven and Waseem (2011), we find that the implied elasticity of earnings with respect to the marginal net-of-tax rate is less than .01. The kink in the life-time budget created by 3-year duration limit of the scheme does create sharp bunching in the density of duration of highly paid foreigners at the 3-year cutoff (relative to controls), but this bunching is also quantitatively small. Using the methodology developed by Saez (2010), we find that this bunching translates into an elasticity of duration of stay (conditional on being in the scheme) with respect the net-of-tax rate around .1. Hence, this decomposition analysis reinforces our earlier finding that the migration elasticity is an order of magnitude larger than the elasticity of reported earnings, which has been the primary focus in the empirical literature on behavioral responses to individual taxation. This also makes it particularly important for advanced economies and particularly European countries to start creating and sharing data on labor migration, especially among highly skilled workers, to monitor the extent of tax induced mobility.

Finally, it is important to note that our results are based on a single small country, Denmark, with a specific language, culture, location and climate. It is a country where almost everybody speaks English and it provides a high quality of life and government services. Furthermore, it is a small and relatively homogeneous country with no strong tradition of immigration. Hence, relative to other countries with stronger traditions of immigration, such as the United States, Denmark starts from a very small base of highly paid foreigners. As a result, in terms of external validity, the extremely large elasticity estimated for Denmark might not carry over to larger countries starting with a higher base of foreign workers. This is closely related to the important point made by Kanbur and Keen (2003) in the context of capital mobility that tax heavens tend to be small countries because small countries start with small bases and hence can large elasticities by granting favorable tax treatment to foreigners. We hope that our pioneering work can be applied to the many other countries, which have adopted such schemes if/when data become available in the future.

The paper is organized as follows. Section 2 presents key features of the Danish tax scheme, proposes a simple theoretical framework, and describes the administrative data we use. Section 3 presents the empirical analysis, while section 4 concludes by discussing policy implications and future work on analyzing spillovers.

2 Context, Conceptual Framework, and Data

2.1 The Danish Tax Scheme

In 1992, Denmark enacted a preferential tax scheme for foreign researchers and high-income foreigners in all other professions, who sign contracts for employment in Denmark after June 1st, 1991. The scheme is commonly known in Denmark as the *Researchers' Tax Scheme*. 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 more than 60% and an average tax rate on high-income workers around 55%. The scheme rate was reduced to 25% in 1995, but at the same time a labor market contribution was added, leaving the total scheme rate roughly unchanged around 30%.

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.⁶ 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 at the 3-year cutoff—a *kink* in the lifetime budget set as a function of duration. Besides this form of duration dependence, the initial scheme design stipulated that a worker who stayed in Denmark for another 48 months after having benefitted from the special tax scheme would face a claw-back equal to the entire tax savings during the period of preferential tax treatment. For a worker who had benefitted from the scheme for the maximum 3-year period, this rule implied a very large retroactive tax bill after 7 years of residence. This form of duration dependence creates a discrete jump in *total* lifetime tax liability at the 7-year cutoff—a *notch* in the lifetime budget set as a function of duration. The duration notch was eliminated for researchers in year 2000 and substantially relaxed for all other professions in year 2002, so that today it applies to very few workers.

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.⁷ Until very recently, citizenship played no formal role

⁶After 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 (a labor market 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.

⁷For 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

in determining eligibility, and therefore Danish citizens who had been living and paying taxes abroad for at least 3 years could also apply for the scheme. Those rules have now been changed such that Danish citizens must be foreign tax residents for 10 years in order to become eligible for the scheme, but this reform lies outside our current data period. The second requirement is that, unless the worker qualifies as a researcher, annual employment earnings must be at least 765,600 Danish kroner (about 103,000 Euros) in 2009 prices. The threshold grows roughly at the same rate as the average earnings level in Denmark. This cutoff lies between the 99th and 99.5th percentile of the Danish employment earnings distribution. As the preferential scheme rate applies to all units of earnings conditional on eligibility, the earnings requirement creates a discrete jump in *total* annual tax liability at the cutoff—a *notch* in the annual budget set as a function of earnings.

In terms of administration, the scheme treatment had 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 does 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, i.e., 51,150 Euros (as of 2009). If the scheme beneficiary has other income besides scheme earnings, that income is taxed according to the standard progressive income tax schedule and completely 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.

To summarize, the special Danish tax scheme creates the following tax variation. First, the scheme introduced much lower tax rates on a specific sample of people (high-income foreigners; not tax liable in Denmark 3 years prior) at a given point in time. This variation provides an ideal setting for a difference-in-differences analysis of migration effects. Second, 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. Third, the scheme originally introduced a 7-year duration notch, which in principle offer more quasi-experimental variation to study duration. However, since this rule was eliminated

spell, the time spent in Denmark under prior scheme spells are not counted.

for most taxpayers not long after the first scheme participants could potentially hit the 7-year notch, we do not have enough observations around the notch to study this precisely. Fourth, the scheme introduced an earnings notch creating very strong incentives for foreigners to have earnings above the income cutoff. This creates compelling quasi-experimental variation to study earnings responses to taxation among foreign migrants.

2.2 Conceptual Framework

We consider a two-country model consisting of a destination country for potential migrants (Denmark) denoted by d and a home country of migrants (rest of the world) denoted by h . The destination country is small and its tax policy therefore does not impact the equilibrium in the rest of the world.

Consider a migrant who spends Y years in Denmark out of a total working life of \bar{Y} years. The utility obtained while in Denmark is given by

$$u^d = \int_0^Y \left\{ z_y^d - T_f^d(z_y^d, y) - nh_z \left(\frac{z_y^d}{n} \right) \right\} dy - \left\{ \mu + \nu h_Y \left(\frac{Y}{\nu} \right) \right\} \cdot 1(Y > 0), \quad (1)$$

where z_y^d is before-tax earnings in Denmark in year y , $T_f^d(\cdot)$ is the tax schedule for foreigners in Denmark (which may depend on both earnings and the year of Danish residence), $h_z(\cdot)$ is disutility of hours worked, n is labor market ability, μ is a moving cost, $h_Y(\cdot)$ is a disutility of spending Y years in Denmark, and the parameter ν captures preferences for Denmark. The disutility functions h_z, h_Y are increasing and convex, so that larger values of the parameters n, ν are associated with lower marginal disutility (of working and being in Denmark, respectively).

The utility outside of Denmark is given by

$$u^h = \int_Y^{\bar{Y}} \left\{ z_y^h - T^h(z_y^h) - nh_z \left(\frac{z_y^h}{n} \right) \right\} dy = (\bar{Y} - Y) \left(z^h - T^h(z^h) - nh_z \left(\frac{z^h}{n} \right) \right), \quad (2)$$

where z_y^h is before-tax earnings at home in year y and $T^h(\cdot)$ is the annual tax schedule in the home country (which does not depend on the number of years spent there). The second equality uses that, because annual tax liability at home is not a function of the year of residence y , the earnings choice is the same in every period ($z_y^h = z^h$).

Total lifetime utility is given by $u^d + u^h$, which is maximized with respect to time spent in Denmark Y , earnings in Denmark in each period z_y^d , and earnings at home z^h . Choosing time

spent in Denmark consists of an extensive margin ($Y = 0$ vs. $Y > 0$) and an intensive margin (size of Y conditional on $Y > 0$).

There is heterogeneity in three dimensions: moving costs μ , preferences for Denmark ν , and labor market ability n . There is a smooth distribution of potential migrants characterized by the density function $f(\mu, \nu, n)$ on the domain $D = (0, \infty) \times (0, \infty) \times (0, \infty)$.

As a benchmark, we start from a smooth (continuously differentiable) tax schedule on foreigners in Denmark. Denoting the marginal tax rate on earnings by $\tau_f^d(z_y^d, y) \equiv \frac{\partial T_f^d(z_y^d, y)}{\partial z_y^d}$, foreigner earnings in year y in Denmark are characterized by

$$h'_z \left(\frac{z_y^d}{n} \right) = 1 - \tau_f^d(z_y^d, y). \quad (3)$$

Duration in Denmark Y (conditional on $Y > 0$) is characterized by

$$h'_Y \left(\frac{Y}{\nu} \right) = \left[z_Y^d - T_f^d(z_Y^d, Y) - nh_z \left(\frac{z_Y^d}{n} \right) \right] - \left[z^h - T^h(z^h) - nh_z \left(\frac{z^h}{n} \right) \right]. \quad (4)$$

The left-hand side is the marginal disutility of time spent in Denmark, while the right-hand side is the marginal consumption increase net of disutility working from time spent in Denmark instead of at home. If the tax schedule were independent of year of residence (in which case $T_f^d(z_Y^d, Y)$ and hence z_Y^d would be independent of Y), then the right-hand side would be constant in Y . For the intensive-margin choice of Y to be positive, the right-hand side must be positive at zero duration, which depends on tax liability at zero duration ($T_f^d(z_0^d, 0)$ compared to $T^h(z^h)$) and on earnings—and therefore marginal tax rates—at zero duration ($\tau_f^d(z_0^d, 0)$ compared to $\tau^h(z^h)$).

Let us now consider the extensive-margin choice of whether to move to Denmark at all. Moving to Denmark requires that lifetime utility $u^d + u^h$ is higher under $Y > 0$ (given optimal choices of Y and z_y^d as characterized above) than under $Y = 0$, which implies

$$\mu \leq \int_0^Y \left\{ z_y^d - T_f^d(z_y^d, y) - nh_z \left(\frac{z_y^d}{n} \right) \right\} dy - \nu h_Y \left(\frac{Y}{\nu} \right) - Y \left(z^h - T^h(z^h) - nh_z \left(\frac{z^h}{n} \right) \right), \quad (5)$$

where $z_y^d = z_y^d(n)$ is determined by (3) and $Y = Y(n, \nu)$ is determined by (4). We denote the right-hand side of (5) by $\mu^* = \mu^*(n, \nu)$. The total number of people migrating to Denmark is then given by

$$N = \int_0^\infty \int_0^\infty \int_0^{\mu^*(n, \nu)} f(\mu, \nu, n) d\mu d\nu dn. \quad (6)$$

The total amount of time spent by migrants in Denmark is given by

$$S = N \cdot E[Y] = \int_0^\infty \int_0^\infty \int_0^{\mu^*(n,\nu)} f(\mu, \nu, n) Y(n, \nu) d\mu d\nu dn, \quad (7)$$

where $E[Y]$ is the average duration among those who migrate to Denmark. The total amount of earnings created by migrants in Denmark is given by

$$N \cdot E[Y \cdot z^d] = \int_0^\infty \int_0^\infty \int_0^{\mu^*(n,\nu)} \int_0^{Y(n,\nu)} f(\mu, \nu, n) z_y^d(n) dy d\mu d\nu dn, \quad (8)$$

where $z^d \equiv \left[\int_0^Y z_y^d dy \right] / Y$ is average earnings per year in Denmark and $E[Y \cdot z^d]$ is average earnings over the entire stay in Denmark among all migrants.

We define by

$$e = \frac{1 - t^d}{S} \frac{\partial S}{\partial(1 - t^d)}$$

the global elasticity of migrants S with respect to the net-of-tax rate $1 - t^d$ where t^d is the average tax rate in Denmark. This elasticity is defined for a given maximum duration y^* and given tax rates in the home country. Naturally, e is likely to increase with y^* as it is likely to be small for small y^* (and actually zero when $y^* = 0$). For very large y^* , the elasticity converges to \bar{e} which is the conventional elasticity of migration with respect to the net-of-tax rate using in optimal tax models with migration such as Mirrlees (1982).

We define by

$$e^A = \frac{1 - t^d}{N} \frac{\partial N}{\partial(1 - t^d)}$$

the elasticity of arrivals N with respect to the net-of-tax rate $1 - t^d$ where t^d is the average tax rate in Denmark. This capture the extensive margin of arrivals.

We will estimate the effect of taxation on all three dimensions of behavior in eqs. (6)-(8): the number of migrants, average duration of stay, and average earnings among those who migrate to Denmark. We do this using a reform that introduced sharp changes in the tax schedule on foreigners $T_f^d(z_y^d, y)$. Before the reform, the tax schedule $T_f^d(z_y^d, y)$ was the same as for domestic residents, a standard piecewise linear function of earnings with no duration dependence. After the reform, the tax schedule $T_f^d(z_y^d, y)$ had much lower rates than the domestic schedule and a very different structure. The new schedule offers a low flat tax rate under two eligibility requirements. First, the schedule is dependent on duration of stay in Denmark with low tax rates applying to years $y \leq y^*$ and high tax rates (standard domestic rates) applying to years $y > y^*$. As the higher rates apply only to the period after y^* , this policy creates a discrete jump

in the *marginal tax* with respect to duration Y (which is given by $T_f^d(z_Y^d, Y)$) and therefore a *kink* in the budget set with respect to duration Y . Second, eligibility for the low flat tax in a given year requires earnings z_y^d above a cutoff z^* , with standard domestic tax rates applying to immigrants with earnings below z^* . As the low flat tax applies to all units of income for those with earnings above the cutoff z^* , this piece of policy creates a discrete jump in *tax liability* with respect to earnings and therefore a *notch* in the budget set.

Figure 1 illustrates behavioral responses to the duration kink (Panel A) and the earnings notch (Panel B). Panel A considers individuals at a given ability n (high enough for earnings to be above the earnings notch z^*), but with different preferences ν for staying in Denmark. Indifference curves represent preferences over duration Y and "net consumption" defined as lifetime consumption net of disutility of working. The slope of the indifference curves is given by $h_Y'(\frac{Y}{\nu})$, corresponding to the left-hand side of equation (4). The slope of the budget line (defined with respect to net consumption) is given by the right-hand side of equation (4). Starting from a duration-independent tax schedule (so that $T_f^d(z_Y^d, Y)$ does not change with Y and therefore z_Y^d does not change with Y), the slope of the budget line is constant in Y . In this case, the durations among immigrants is smoothly distributed reflecting the smooth distribution of ν . Suppose now that tax liability $T_f^d(z_y^d, y)$ is increased after year y^* (at a given level of earnings). This creates a kink in the budget line at $Y = y^*$ as the slope on the right-hand side of equation (4) jumps. This kink produces bunching of individuals whose durations were falling into a segment $[y^*, y^* + \Delta y^*]$ before the kink was introduced. Individual L has the shortest pre-kink duration (smallest ν) among those who bunch at the kink; this individual chooses duration y^* both before and after the reform. Individual H has the longest pre-kink duration (largest ν) among those who bunch at the kink; this individual chooses duration $y^* + \Delta y^*$ before the tax reform and duration y^* after the reform. Every individual between L and H locates at the kink point, which creates total bunching $B = \int_{y^*}^{y^* + \Delta y^*} g_0(y) dy$ where $g_0(\cdot)$ denotes the counterfactual (pre-kink) duration distribution. Based on estimates of $g_0(y)$ and B , this relationship yields the duration response Δy^* which can be used to back out the compensated elasticity of duration with respect to the tax rate following the method developed in Saez 2010. Saez (2010) shows that the intensive elasticity e^I can be approximated by the following equation:

$$\frac{\Delta y^*}{y^*} \simeq e^I \frac{\Delta \tau}{1 - \tau}$$

where t^* is the location of the kink equal to 3 years here and $\Delta \tau$ is the change in marginal

incentive (in the life-time budget) and τ the current marginal tax rate (in the life-time budget).

Panel B considers the implications of the earnings notch. This diagram considers earnings choices z_y^d at a point in time $y < y^*$ for individuals with different abilities n . Indifference curves represent preferences over labor supply z_y^d/n and consumption c_y^d at a point in time in Denmark (the slope of which is $h'_z(z_y^d/n)$) and the budget set is given by $c_y^d = z_y^d - T_f^d(z_y^d, y)$. For simplicity of exposition, we start from a linear tax schedule before the reform (although in practice the pre-existing schedule was piecewise linear). Under the baseline tax system, there is a smooth distribution of earnings reflecting the smooth distribution of n . Now a preferential flat tax is introduced for foreigners, conditional on having earnings above a cutoff z^* . This creates an upward notch in the budget set at z^* , which produces bunching at z^* combined with a hole in the density distribution in an interval $[z^* - \Delta z^*, z^*]$. Individual L has the lowest pre-notch income (lowest n) among those who locate at the notch point; this individual chooses earnings $z^* - \Delta z^*$ before the reform and is exactly indifferent between $z^* - \Delta z^*$ and the notch point z^* after the reform. Individual H has the highest pre-notch income (highest n) among those who locate at the notch point; this individual chooses earnings z' before the reform and earnings z^* after the reform. Individuals who were initially between z' and z^* move to the upper bracket (strictly above z^*) such that no individuals locate between $z^* - \Delta z^*$ and z^* under the notched schedule. The total amount of bunching is given by $B = \int_{z^* - \Delta z^*}^{z'} h_0(z) dz$ where $h_0(\cdot)$ denotes the counterfactual (pre-notch) earnings distribution. A technical complication is that estimates of B and $h_0(z)$ are not sufficient to reveal Δz^* as we do not know z' .⁸ A lower bound on the size of the earnings response can be obtained by assuming $z' \simeq z^*$ (corresponding to an assumption that the uncompensated elasticity is not too large). Based on an estimate of Δz^* obtained in this way, it is possible to back out the compensated elasticity of earnings with respect to the marginal tax rate using the method set out by Kleven and Waseem (2011).⁹

⁸This is an issue specific to *upward* notches and does not arise in the case of downward notches (Kleven and Waseem 2011).

⁹When implementing this strategy empirically, the above bunching formula should be modified to account for the fact that the observed below notch density will not be completely empty on the interval $[z^* - \Delta z^*, z^*]$ due to the presence of frictions. To deal with this, Kleven and Waseem (2011) propose a bunching formula of the form $B = \int_{z^* - \Delta z^*}^{z^*} [h_0(z) - h(z)] dz$ where $h(z)$ is the observed below notch earnings density (which captures those who are stuck on the segment $[z^* - \Delta z^*, z^*]$ due to frictions).

2.3 Administrative Tax Data

The data we use in this paper comes 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 include detailed information about income and tax variables, labor market variables, and socio-demographic information. Most importantly, the data contains very specific citizenship and migration information. Each individual working in Denmark must receive a personal identification number (CPR) in order to pay withholding taxes, rent an apartment, etc. The application for a CPR number contains detailed questions about citizenship and date of entry in Denmark, as well as about the country of in-migration. The registry administration updates this information in case an individual leaves the country, but this update can take a few years, and therefore we only have precise information about in and out migration until 2005. The data also contains very detailed information for all scheme beneficiaries with precise information on starting and ending dates of labor contracts. Unfortunately, this information was not computerized for the first years of the scheme and therefore we do not have individual wage information available for scheme beneficiaries for the years 1991 to 1994 included.

2.4 Summary Statistics

Number and Composition of Scheme Workers.

Figure 2 reports the composition of beneficiaries of the tax scheme (excluding researchers) by country of citizenship (Panel A) across all years 1991 to 2006. Unsurprisingly, the vast majority of scheme workers come from advanced economies, 25% come from Nordic countries, 10% are Danish citizens (who qualify as others as long as they have not paid taxes in Denmark for 3 years or more), 19% from the United Kingdom or Ireland, 10% from Northern-America, and about 20% from Germany, France, and Benelux combined. The number coming from emerging countries or Eastern Europe and Russia is extremely modest.

Take-up Rate.

Figure 3 reports the take-up rate for the scheme among foreigners arriving in Denmark with (annualized) earnings above the eligibility threshold and who have not paid taxes in Denmark over the last 3 years. The take-up rate is high, around 70%, but still significantly below 100%

for a variety of reasons. First, companies have to file out an application form for each employee eligible for the scheme. It is conceivable that not all companies knew about the scheme or were willing to do the administrative paperwork. Second, some individuals may not have been willing to take-up the scheme perhaps because of the claw-back rule after 7 years. Third, individuals may not be fully aware of the existence of the scheme, or might come to Denmark as political refugees. Following the Dayton Peace Agreements in December 1995 for instance, there was an important influx of migrants from Bosnia-Herzegovina, who were for the most part political refugees and some of them being high-skilled workers, creating a drop in the take-up rate of the scheme in 1995-1996.

3 Empirical Evidence on Behavioral Responses

In this section, we first estimate in Section 3.1 the overall effect of the reform and summarize the effect with a global elasticity. In the following subsections and using the conceptual framework we developed earlier, we decompose the global effect into a migration extensive effect, intensive labor supply effect on the job, and intensive effect along the duration of stay margin.

3.1 Global Reduced-Form Elasticity

The most transparent way to test whether the Danish tax scheme had an impact is to plot the number of foreigners with earnings above the threshold overtime as we do in Figure 4. The series (denoted as treatment in the figure) shows that the number of highly paid foreigners was fairly stable around 800 from 1980 to 1990 before the scheme was implemented. After the scheme takes place in 1991, denoted as a vertical line in the figure, there is a steady increase in the number of highly paid foreigners. The number highly paid foreigners reaches over 3000 in 2006. Naturally, it is conceivable that the number of highly paid foreigners could have increased absent the scheme because of European Union labor market integration following the Single European Act signed in 1986 and implemented in 1992. The simplest way to control for such trends is to also plot series of the number of highly paid foreigners slightly below the threshold of eligibility for the scheme. Figure 4 therefore 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 series are normalized so that they match the treatment series in 1991 the year before the scheme was first implemented. Three lessons

emerge from those controls.

First, the control series follow very closely the treatment series before the scheme is implemented. All three series are remarkably stable providing credibility to our assumption that those groups of foreigners below the threshold are good control groups for the treatment group 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 much more slowly than the treatment series. Indeed, after 1995, the treatment series are consistently about twice as high as the control series. Third, there might be a concern that highly paid foreigners above the threshold are displacing highly paid foreigners slightly below the threshold through intensive earnings responses as we described in the theory section. For example, a company that was planning to hire a foreigner just below the threshold might adjust the contract to require slightly more work for a slightly higher pay qualifying for the contract. It might also shift the form of compensation to increase cash compensation at the expense of fringe benefits or on the job perks that do not qualify toward the threshold. Such shifting should produce a dip in the number of foreigners just below the threshold relative to foreigners further down below the threshold. The very close and parallel pattern of evolution of both control groups (the one just below the threshold 90% to 99.5% and the one further down the threshold 80 to 90%) shows that this dip effect was not significant. We will come back to this issue in more detail when we analyze earnings densities for foreigners around the threshold in Section 3.2.

Figure 5 shows heterogeneity in responses by country of origin of foreign immigrants. It repeats Figure 4 but restricting foreigners to sub-groups of countries: Nordic Countries (Finland, Iceland, Norway, Sweden) in Panel A, English Speaking Countries (Australia, Canada, Ireland, New Zealand, UK, US) in Panel B. Continental Europe (all other European countries excluding Eastern Europe) in Panel C, and all other countries in Panel D. The figure shows that each country group experienced a large migration effect. However, there is heterogeneity across country groups: the response is somewhat smaller for Nordic countries and continental Europe and significantly larger for English speaking countries. Other countries (primarily Eastern Europe) show at first no response and then a very large response after 1996 when migration from Eastern Europe to Denmark was liberalized.

An additional concern with our analysis is that there might have been a widening of earnings inequality in Denmark since the 1990s. Such widening would have increased the number of workers paid above the threshold both among Danish natives and foreigners and hence led to a divergence in the number of foreigners in the treatment group and control groups even absent any scheme effect. The threshold for eligibility always lies between the 99th and the 99.5th percentile of the full earnings distribution among Danish adults with positive earnings. This concern is unlikely to be an issue in the case of Denmark as Kleven and Schultz (2011) have shown that top income shares have remained remarkably stable in Denmark since 1980. We address this issue with our data in two ways.

First, Figure 6 plots average earnings in various upper percentile groups both below and above the threshold among Danish citizens (normalized to 1 in 1990). It shows quite parallel trends for all groups confirming the findings of Kleven and Schultz (2011) that the upper tail of the earnings distribution did not widen in Denmark in recent decades.

Second, Figure 8, Panel A then plots the fraction of foreigners in various percentiles of the earnings distribution. We have a gap in 1991-1994 because the scheme related data does not provide scheme earnings for those years. Two important findings should be noted from the figure. First the fraction of foreigners in each percentile is very stable before 1991.¹⁰ Second, after 1991, the fraction of foreigners increases much more rapidly—both in absolute and percentage terms—in the the two top groups above percentile 99.5th where the scheme applies. Normalizing the fraction of foreigners to one in 1991, in Figure 8, Panel B, we see that the fraction of foreigners doubles from 1990 to 1995 above percentile 99th while it remains pretty stable below percentile 99th.

In conclusion, estimates based on the fraction of foreigners in various fixed percentiles of the distribution, which can fully control for changes in earnings dispersion deliver roughly the same doubling effect as our earlier estimates based on the increase in foreigners above the fixed (in real terms) threshold. Hence, if we use as identification assumption that, absent the scheme, the relative number of foreigners above the threshold would have increased at the same rate as the number of foreigners slightly below the threshold, we find an extremely large doubling effect of the scheme.

Table 1 summarizes the graphical evidence described above and gives more formal estimates

¹⁰The fraction of foreigners is higher, around 5%, above the 99.9th percentile, than in percentiles 96 to 99.9 where it is around 2 to 3%.

of the effect of the tax scheme on the total number of foreigners in Denmark. We estimate difference-in-difference models of the following form:

$$Y_{it} = \alpha_0 + \alpha_{tr} \cdot \mathbf{1}[i = 1] + \alpha_{rf} \cdot \mathbf{1}[t > 1991] + \alpha_{tr*rf} \cdot \mathbf{1}[i = 1] \cdot \mathbf{1}[t > 1991] + \nu_{it}$$

where Y_{it} is the total number of foreigners (weighted by duration) in income group $i = 0, 1$ present in Denmark in year t . Assuming that, absent the reform, the number of foreigners would have followed the same trend in the control ($i = 0$) and treatment ($i = 1$) group, the effect of the reform on the total number is identified by the coefficient α_{tr*rf} . We also compute the percentage increase in the total number of foreigners as $\frac{\alpha_{tr*rf}}{\alpha_0 + \alpha_{rf} + \alpha_{tr}}$, where $\alpha_0 + \alpha_{rf} + \alpha_{tr}$ is the counterfactual number of foreigners that would have prevailed in the treatment group after 1991 in the absence of a reform.

In column (1) we define control as the income group lying between 80% and 100% of the scheme threshold level, and treatment as the the income group above the threshold level. Pre-reform period is years 1980 to 1990 and post-reform is defined as years 1995-2000, in order to estimate the longer term effect of the reform. The estimated percentage increase in the number of foreigners is 90% and precisely estimated. In column (2), we focus on the short term effect by comparing years 1991-1995 to 1980-1990. The effect is still large and significant, with an estimated percentage increase of .65. In column (3), we estimate a placebo specification, where the treatment group is constituted of individuals with income between .9 and 1 times the threshold level, while the control group is constituted of individuals with income between .8 and .9 times the threshold level. This specification is also a test for the presence of shifting around the threshold in case of strong behavioral responses in terms of earnings or avoidance, since shifting should produce a dip in the number of foreigners just below the threshold relative to foreigners further down below the threshold. The coefficient α_{tr*rf} is negative but very small and not significant, which confirms the graphical evidence that shifting behaviors are second-order. In column (4), we use individuals with income above the 99.5-th percentile of the income distribution as treatment, and individuals with income between the 96-th and the 99-th percentile of the income distribution as control. The results are in line with column (1) and the estimated percentage increase in high-skilled foreigners implied by the reform is 111%. We then decompose the effect by regions of origin in column (5) to (8). As suggested by figure 5, the effect of the reform is stronger among high-skilled individuals from English-speaking countries,

with an estimated percentage increase of 112%.

Elasticity Estimate.

Based on the graphical evidence we have presented, we see that the number N of foreigners above the threshold increases by about 100% relative to the number of foreigners slightly to somewhat below the threshold. Our identification assumption is that all of this extra-increase is due to the introduction of the tax scheme. As discussed above, this assumption is validated by the pre-scheme data from 1980 to 1990 and with the post-scheme data when comparing various control groups below the threshold.

The average tax rate τ in the scheme is a flat 25% while the tax rate outside scheme is around 55%. This implies that the scheme produces an increase in the net-of-tax rate from 45% to 75%, i.e., a percent increase in the net-of-tax rate of $75/45-1=67\%$. Hence, the global elasticity e is estimated as:

$$e = \frac{1 - \tau}{N} \frac{\partial N}{\partial(1 - \tau)} \simeq \frac{1}{0.67} = 1.5.$$

This is a very large elasticity. This elasticity applies for the existing scheme with its three year duration. Conceivably, the elasticity for a scheme with indefinite duration could be even higher.

This large elasticity implies that the tax revenue maximizing rate is $\tau^* = 1/(1 + e) = 40\%$ using the conventional inverse elasticity formula. When taking into account the 20% VAT tax (that applies to only to the fraction of consumption subject to the VAT rate) and the 8% payroll tax, this shows that the 25% scheme tax is not very far from the tax revenue maximizing rate.

As we discussed in the theory section, the elasticity we have estimated can be due to (a) the extensive margin of migration to Denmark of highly skilled foreigners due to lower average tax rate of the scheme, (b) the intensive margin along the earnings level margin to take advantage of the scheme that creates a notch in the individual budget set at the threshold of eligibility, (c) the intensive margin along the duration of stay margin as the scheme lasts only three years.

3.2 Decomposing the Global Elasticity

Extensive Elasticity: Arrivals in Denmark.

Next, we analyze the response along the extensive margin, namely the migration to Denmark of highly paid foreigners. Figure 9 reports the number of foreigners with earnings above the scheme eligibility threshold (treatment series) arriving each year in Denmark from 1980 to 2006. As

control groups, it reports the number of foreigners arriving 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. This figure is quantitatively very similar to our initial graph with the absolute number of foreigners. It shows a slightly more than doubling of arrivals of foreigners above the threshold relative to foreigners below the threshold. This shows that the extensive arrival elasticity is even slightly higher than the global elasticity we estimated initially. The difference could be due to composition effects as people attracted to Denmark because of the scheme might have lower durations of stay than people coming to Denmark absent the scheme.

Intensive Elasticity along the Earnings Margin

In order to analyze the intensive elasticity along the earnings level margin, we exploit the fact that the scheme creates a notch in the individual budget set as eligibility kicks in above a fixed threshold. As we described in the theory section, behavioral responses to take advantage of the notch could be along the labor supply margin (working more to earn more and qualify for the scheme) or the tax avoidance margin (adjusting the form of compensation to increase cash compensation to qualify for the scheme at the expense of other compensation). Such behavioral responses should create a piling up in the earnings density of foreigners at the earnings threshold along with a gap just below the threshold.

To examine this, Figure 10 plots such densities before the scheme was introduced (1988-1990 in light grey) and after the scheme was introduced (1995-1997 in black). The vertical line denotes the threshold and the x-axis is normalized so that zero corresponds to the threshold. The density is smooth around the threshold before the scheme was introduced. After the scheme was 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 consistent with a strong migration response at all earnings levels above the threshold. Second, there is clear bunching at the threshold, but with only a very small dip to the left of threshold. Even though the bunching is clearly visible in the figure and therefore provides very compelling evidence of a behavioral response, it is in reality very modest relative to the extremely large notch in budget set created by the scheme.

Using the methodology developed by Kleven and Waseem (2011), we can estimate that the spike at the notch point translates into an elasticity of earnings with respect to the marginal net-of-tax rate that is extremely small, less than .01. This estimation assumes iso-elastic utility with no income effects in which case the intensive earnings elasticity e with respect to the marginal net-of-tax rate is determined by

$$\frac{z^* - \Delta z^*}{z^*} = (1 + e) \frac{1 - \tau^s}{1 - \tau^d} - e \left(\frac{z^*}{z^* - \Delta z^*} \right)^{\frac{1}{e}},$$

where τ^s is the average tax rate under the scheme, τ^d is the average tax rate in the regular tax schedule, z^* is the earnings cutoff, and Δz^* is the earnings response by the marginal bunching individual as described in the theory section. The earnings response Δz^* can be determined as described in section 2.2 based on the amount of excess bunching at the notch relative to the local counterfactual density, absent the notch.

Using figure 10, we have $\Delta z^* \simeq 10,000$ kroner as the excess mass at the notch point corresponds roughly to the density mass between the threshold and the threshold + 10,000 kroner. Moreover, we have $z^* \simeq 550,000$ kroner (scheme eligibility threshold in the mid-1990s shown in the figure) along with average income tax rates $\tau^d = 45\%$ below the cutoff and $\tau^s = 25\%$ above the cutoff. Using those numbers in the formula above, we obtain $e \simeq .004$. This elasticity is extremely tiny as the amount of bunching, although significant, is very modest relative to the enormous change in average tax rate of 20 points at the notch.¹¹ For example, even a modest elasticity of 0.1 would translate into a Δz^* of around 100,000 kroner, i.e., a massively large amount of bunching.

Naturally, it is possible that this elasticity is so tiny because of large frictions in the ability of highly paid foreigners to adjust their labor supply and earnings as in the theory proposed by Chetty (2011). In the particular case of the scheme, such an adjustment needs to be negotiated between the employee and the employer before hiring takes place and while the worker is still abroad as only new immigrants are eligible for the scheme. Hence, such an agreement might be difficult to set-up unless employers actively cooperate. As scheme workers are so few, employers may not want to systematically change their behavior to accommodate them as in the theory proposed by Chetty et al. (2011) following their kink bunching analysis in Denmark. Therefore,

¹¹Kleven and Waseem (2011) obtain larger but still quantitatively small elasticities of .05-.1 for the self-employed and .01-.04 for wage earners in the case of Pakistan where the jumps in average tax rates are much smaller, around 1%, and yet observable bunching is larger than in the case of the Danish tax scheme.

our result does not imply that the intensive elasticity of earnings with respect to the net-of-tax rate for highly paid foreigners is zero in all circumstances but rather that it is zero in this particular case.¹² In particular, this implies that firms and foreign workers are almost completely unable to game the system by manipulating earnings to take advantage of the scheme. Hence, we can conclude with great confidence that earnings manipulation or labor supply changes around the scheme notch does not affect our core migration effects that are the primary focus of our analysis.

Intensive Elasticity along Duration Margin

In order to analyze the intensive elasticity along the duration of stay margin, we exploit the fact that the scheme creates a kink in the individual life-time budget set as the scheme lasts for only 3 years. As we explained in the theory section, a number of individuals should stay exactly 3 years and hence create bunching in the density of stay durations at 3 years. To examine this, Figure 11 reports the density of duration of foreigners with (annualized) earnings above percentile 99.5th (P99.5-100) and hence above the eligibility threshold for the scheme and earnings between percentile 96 and percentile 99 (P96-99) and hence below the eligibility threshold for the scheme. Panel A is for years 1991-2005 (after the scheme was implemented) while Panel B is a control graph for years 1980-1990 (before the scheme was implemented). There is a clear spike at 3 years only in the upper group eligible for the scheme after the scheme was implemented providing very strong compelling evidence of an intensive response along the duration margin.

However, the spike is small relative to the large change in marginal incentives created the scheme. Figure 12 depicts the density of duration of foreigners with (annualized) earnings above percentile 99.5th for years 1991-2005, zooming around the 3-year cut-off date. The excess bunching at the kink represents about 2 month = 2/12 year of density. Let us denote by Δy^* this excess bunching (expressed in duration density equivalent terms). Using the methodology of Saez (2010), the intensive elasticity e^I can be estimated using the formula

$$e^I \simeq \frac{\log(1 + \Delta y^*/y^*)}{\log((1 - \tau_s)/(1 - \tau_d))} = \frac{\log(1 + 2/36)}{\log((1 - .25)/(1 - .55))} = .106$$

where $\tau_s = .25$ is the average tax rate on the scheme and $\tau_d = .55$ is the average tax rate on the

¹²Recent estimations of intensive elasticities of earnings among all workers in Denmark from Chetty et al. (2011) and Kleven and Schultz (2011) based on very compelling bunching evidence and tax reforms evidence (respectively) find positive but modest elasticities around 0.1 still much smaller than the migration elasticity we have found.

regular schedule. This formula applies for large changes in marginal incentives (as is the case for the scheme) for iso-elastic utilities with no income effects. Therefore, the intensive elasticity along the duration margin is also very small relative to the total elasticity.

4 Conclusion and Future Work

Our paper has analyzed the effects of income taxation on the international migration of top earners using the Danish tax scheme introduced in 1991 that offers preferential tax treatment to highly paid migrants to Denmark. This scheme offers a unique opportunity to credibly estimate elasticities of international mobility with respect to tax rates. Using population wide Danish administrative tax data, 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 around 1.5. We have shown that there is a significant but quantitatively very small response along the intensive earnings margin (work effort or earnings manipulation through tax avoidance). There is also evidence of sharp bunching of durations of stay of highly paid foreigners at the three year duration limit which translates into a significant but quantitatively small intensive elasticity of duration with respect to the net-of-tax rate. In the end, the migration elasticity is much more larger than the conventional within country elasticity of earnings with respect to the net-of-tax rate. As is the case for wealth mobility across countries (Kanbur and Keen, 1993), it is conceivable that elasticities are particularly large from small countries—such as Denmark—and that those small countries have the most to gain from preferential tax schemes for foreigners. Hence, “tax heavens” for highly skilled workers are likely to generate severe tax competition across European countries and severely limit the ability of European governments to use progressive taxation. This will require international coordination and the design of rules regulating such special schemes particularly in the European union.

Another important rationale put forward by Danish tax authorities as well as other European governments which have adopted similar preferential tax scheme for highly paid foreigners is that highly skilled workers generate positive externalities on their co-workers and the economy at large. 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 models. There is also substantial work in the urban economics

literature trying to uncover such spillover effects empirically (see e.g., Moretti, 2004). The Danish scheme creates a large and exogenous influx of highly skilled workers in the Danish economy, concentrated among specific industries. Therefore, in future work, we hope to make progress on estimating spillovers of scheme workers on co-workers and broader economic growth in two ways.

First, we could use a Bartik type instrument measure for the exposure of each industry or firm to the scheme. The instrument would be the fraction of highly paid foreign workers in the industry or individual firm just before the introduction of the scheme. This would allow us to measure the effect of attracting scheme workers on wage dynamics of domestic workers or non eligible foreigners as well as firms' hiring and growth behavior. Second, we plan to study 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, J. M., and J. B. Slemrod.** (2004) “Do the Rich Flee from High State Taxes? Evidence from Federal Estate Tax Returns,” NBER Working Paper No. 10645.
- Blundell, R. and T. 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, G. J.** (1999). “The economic analysis of immigration.” In: Orley C. Ashenfelter and David Card, Ed(s), *Handbook of Labor Economics*, Elsevier: Amsterdam, Volume 3A.
- Borjas, G. J.** (1999b). “Immigration and Welfare Magnets”, *Journal of Labor Economics* 17(4), Part 1, 607-637.
- Chetty, Raj** (2011). “Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply”, forthcoming *Econometrica*.
- 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, M. and M. V. Wrobel** (1998). “Can state taxes redistribute income?” *Journal of Public Economics* 68, 369-396.
- Gordon, R. H. and J. R. Hines** (2002). “International taxation,” in A. Auerbach and M. Feldstein (eds.), *Handbook of Public Economics*, Volume 4, North-Holland, Amsterdam.
- Kanbur, Ravi and Michael Keen** (1993). “Tax Competition and Tax Coordination When Countries Differ in Size,” *American Economic Review* 83(4), 877-892.
- Kirchgassner, G., and W. W. 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** (2010). “Taxation and International Migration of Superstars: Evidence from the European Football Market.” NBER Working Paper No. 16545.
- Kleven, Henrik J. and Esben A. Schultz** (2011). “Estimating Taxable Income Responses using Danish Tax Reforms.” LSE Working Paper, August 2011.
- Kleven, Henrik J. and Mazhar Waseem** (2011). “Behavioral Responses to Notches: Evidence from Administrative Tax Records in Pakistan.” LSE Working Paper, November 2011.
- Liebig, T., P. A. Puhani, and A. 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.
- Meyer, Bruce** (2000). “Do the Poor Move to Receive Higher Welfare Benefits?”, Working Paper.
- 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).

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

Saez, E., J. Slemrod, and S. Giertz (2011). “The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review.” *Journal of Economic Literature*.

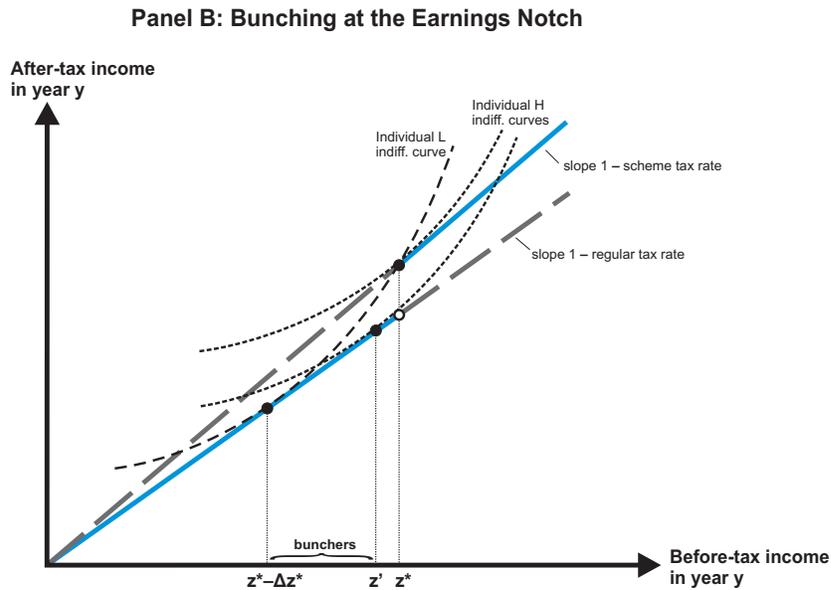
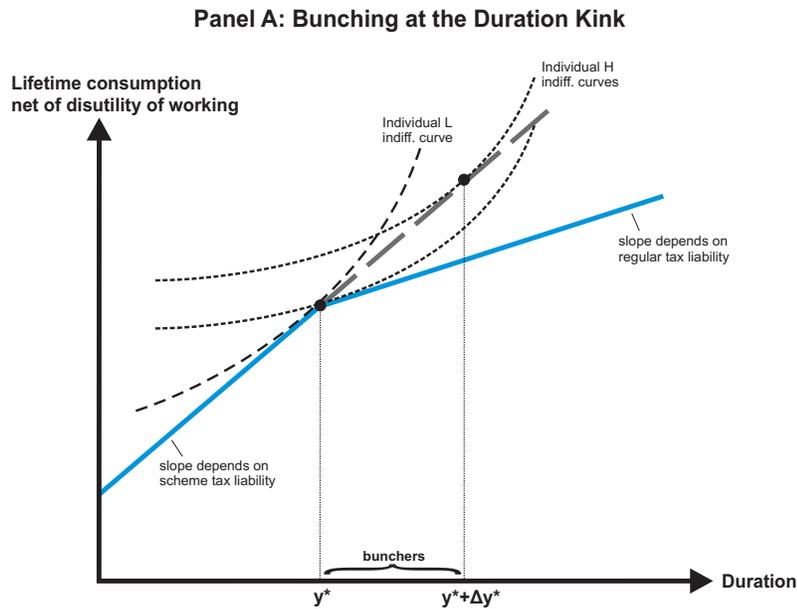


Figure 1: Duration and Earnings Responses to the Foreigner Tax Scheme in Denmark

The Figure illustrates behavioral responses to the duration kink (Panel A) and the earnings notch (Panel B). Panel A depicts the kink created by the scheme in the life-time budget defined over duration y of stay in Denmark and net consumption. The blue solid line budget represents the scheme where the slope is higher up to duration $y^*=3$ years and lower afterwards. The grey dashed line budget represents a scheme with indefinite duration. A scheme a finite duration generates bunching at the $y^*=3$ year duration 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 eligibility earnings threshold z^* in the budget set of the individual (solid blue line). The notch creates piling up of individuals at the eligibility threshold and a corresponding gap in the density distribution just below the notch.

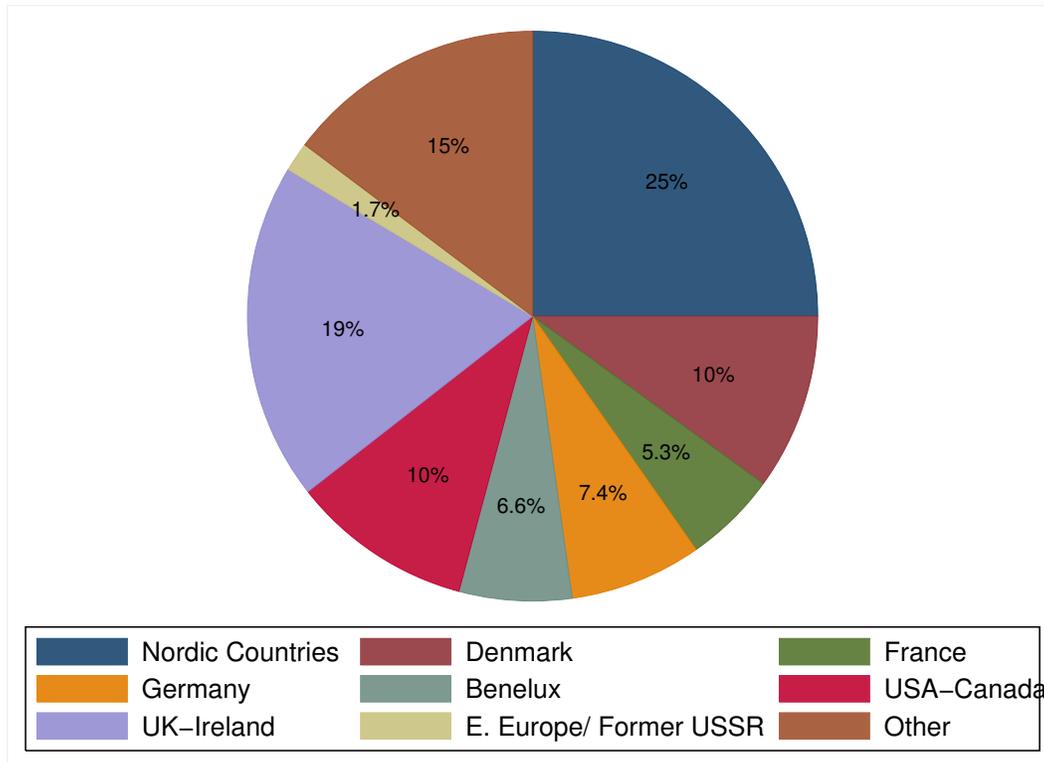


Figure 2: Citizenship of Scheme Beneficiaries, 1991-2006

The Figure reports the composition of beneficiaries of the tax scheme (excluding researchers) by country of citizenship (at the time of scheme) across all years 1991 to 2006.

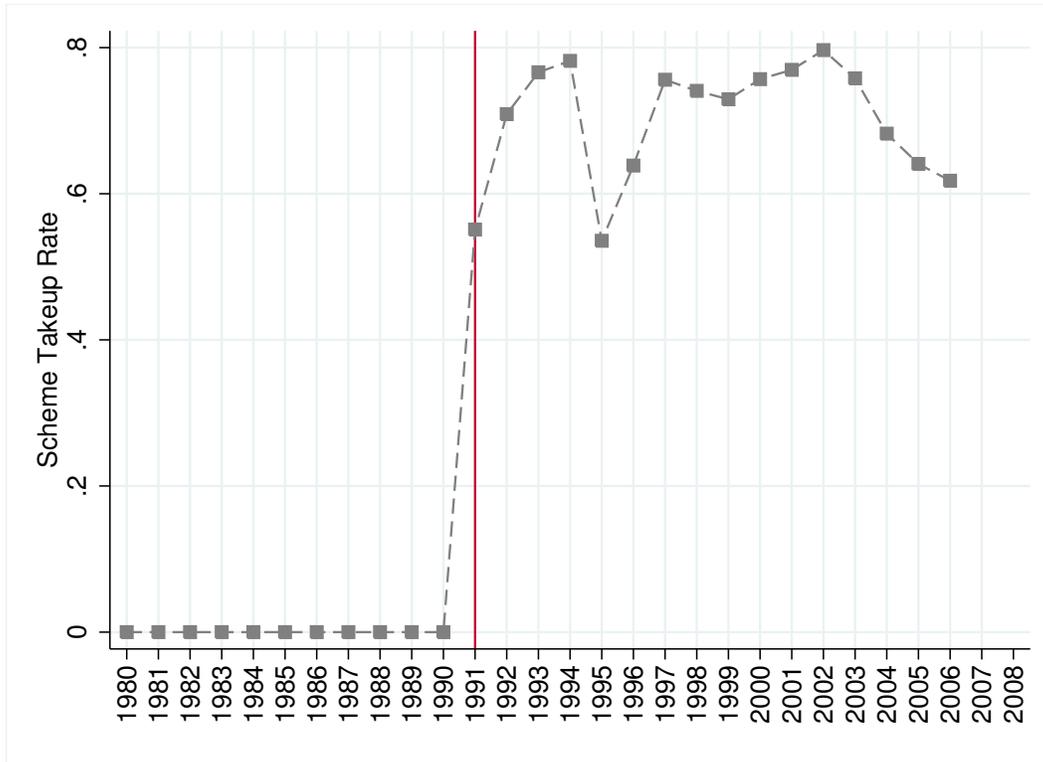


Figure 3: Scheme Take-up rate for eligible foreigners upon arrival

The Figure reports the take-up rate for the scheme among foreigners arriving in Denmark with (annualized) earnings above the eligibility threshold and who have not paid taxes in Denmark over the last 3 years.

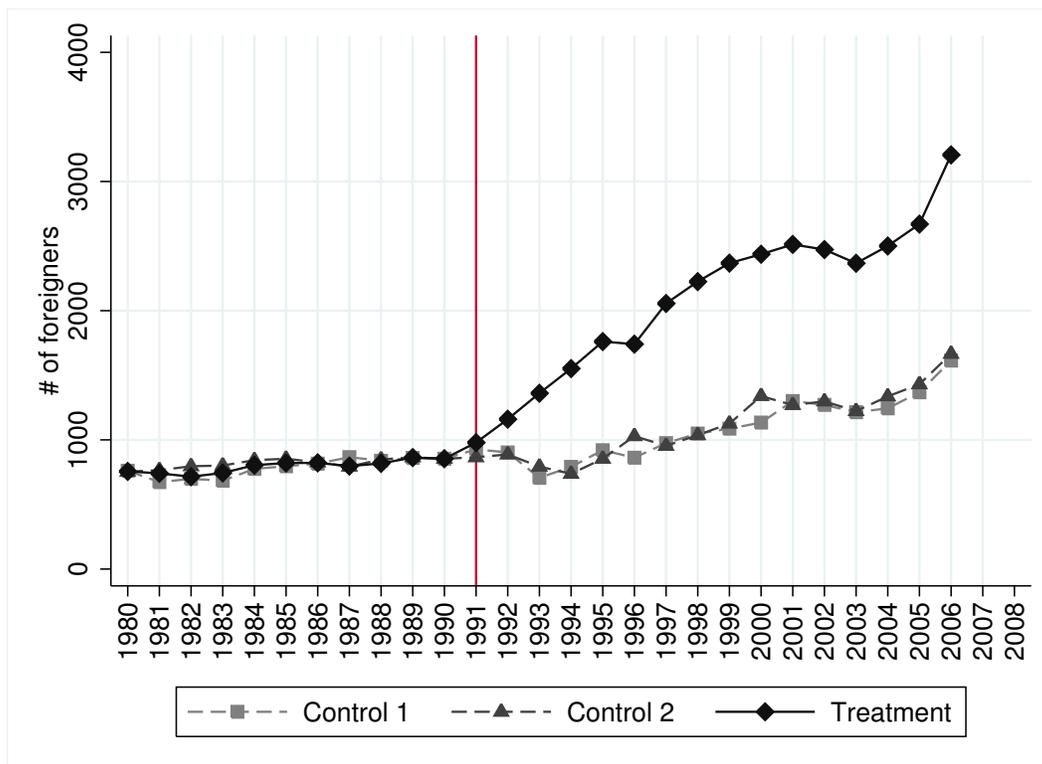


Figure 4: Total number of foreigners in different income groups

The Figure reports the number of foreigners with earnings above the scheme eligibility threshold (treatment series) from 1980 to 2006. 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 1991). All numbers are weighted by duration of stay during the year for part-year foreign residents and earnings are also annualized for part-year residents.

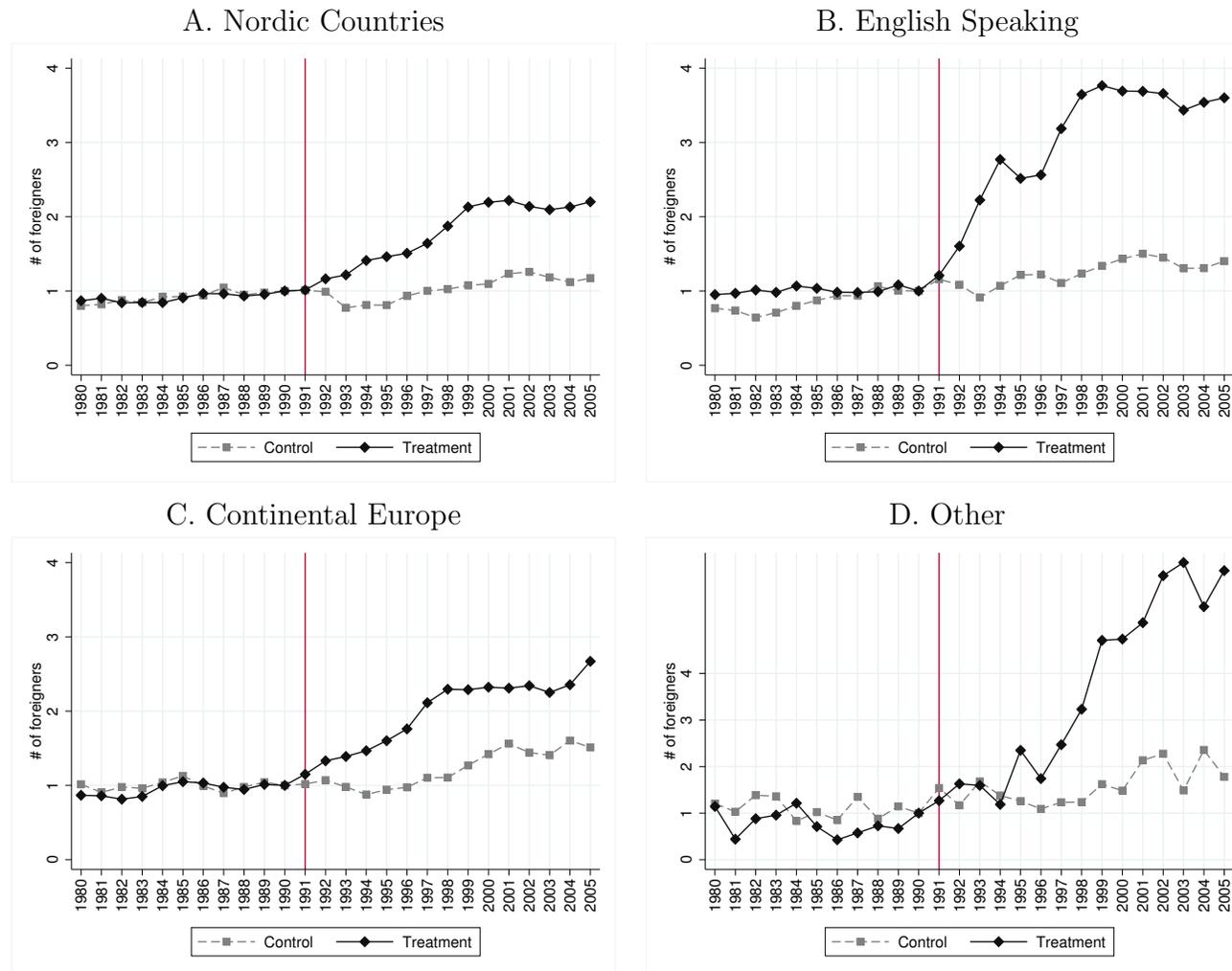


Figure 5: Total Number of foreigners in different income groups by regions of origin

The Figure reports total number of foreigners with earnings above the scheme eligibility threshold (treatment series) from 1980 to 2006 by country group of origin. As control groups, it reports the number of foreigners in Denmark with earnings between 80% and 99.5% of the threshold (control). All series are normalized to 1 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 1991). All numbers are weighted by duration of stay during the year for part-year foreign residents and earnings are also annualized for part-year residents. Nordic Countries=Finland, Iceland, Norway, Sweden. English Speaking= Australia, Canada, Ireland, New Zealand, UK, US. Continental Europe= all other European countries excluding former communist Republics.

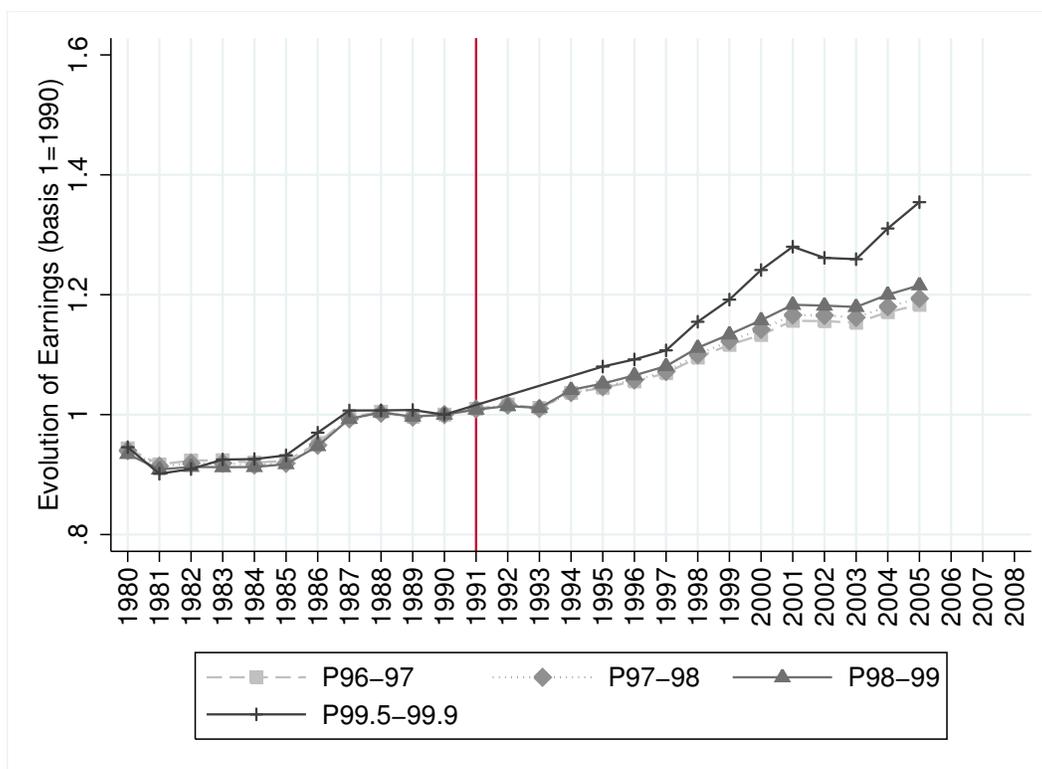


Figure 6: Average Earnings of Danish workers in different fractiles

The Figure plots average real earnings in various upper percentile groups both below and above the threshold among Danish citizens with positive earnings (normalized to 1 in 1990). P96-97 denotes all individuals between the 96th and 97th percentile, etc. The goal of this figure is to show that there were no widening of the upper part of the Danish earnings distribution during this period as all series grow at the same rate. The threshold for eligibility to the scheme is always between the 99th and the 99.5th percentile. The vertical line at year 1991 denotes the year the scheme was first implemented.

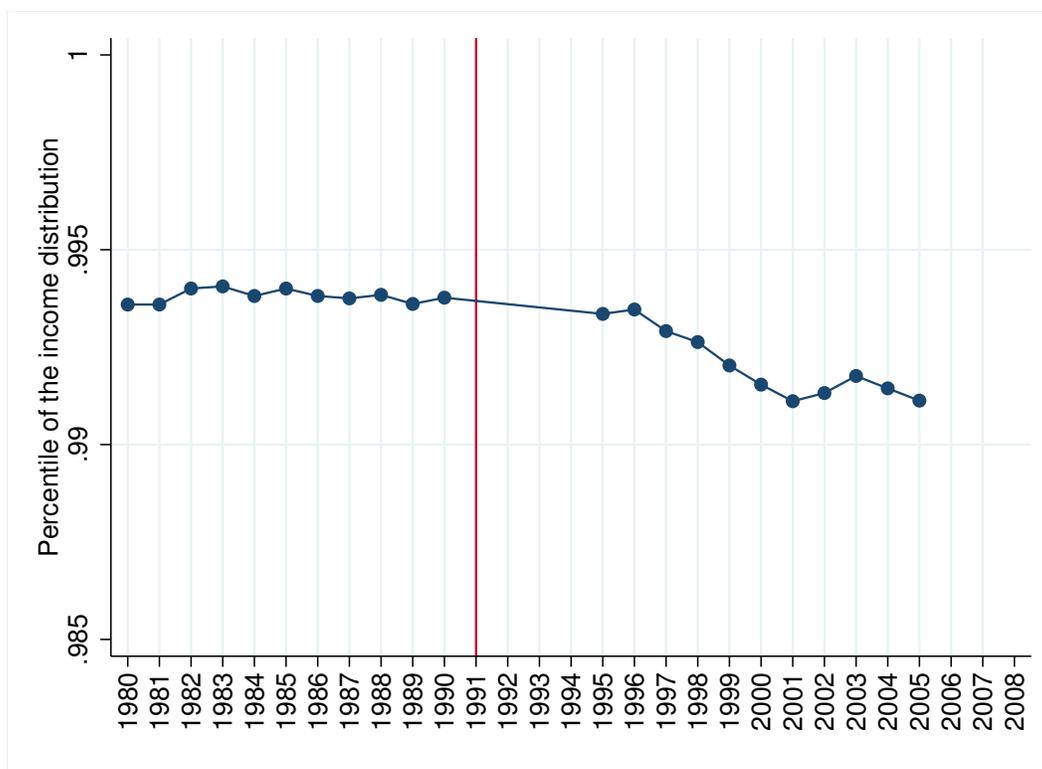
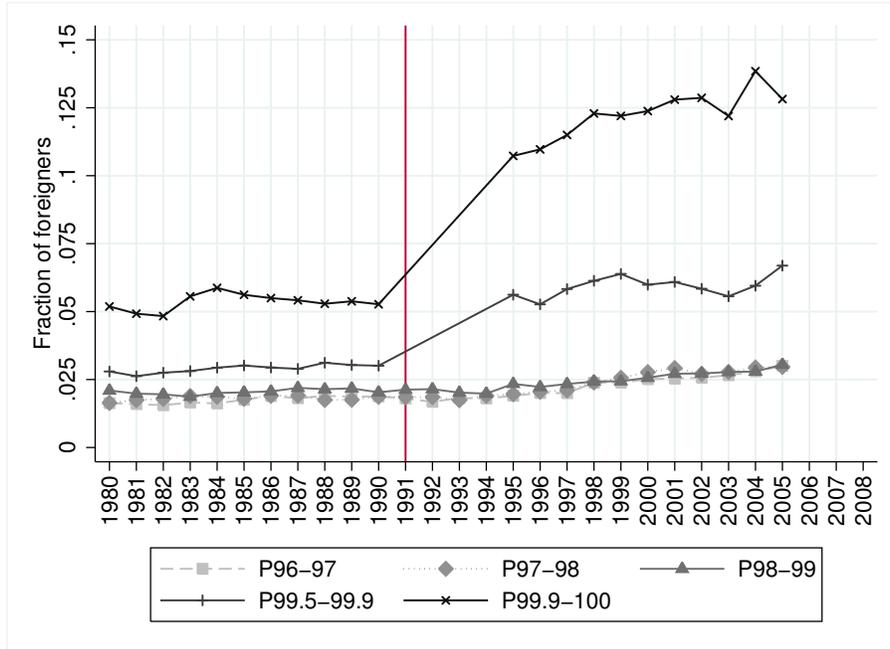


Figure 7: Evolution of the Level of the Scheme Income Threshold in the Income Distribution

The Figure reports the evolution of the level of the income threshold in the distribution of income in Denmark. The threshold for eligibility to the scheme always lies between the 99th and the 99.5th percentile, but as more foreigners enter the country with income above the threshold, the fraction of individuals with income above the eligibility threshold increases over time.

A. Fraction of Foreigners



B. Evolution of Normalized Fraction of Foreigners

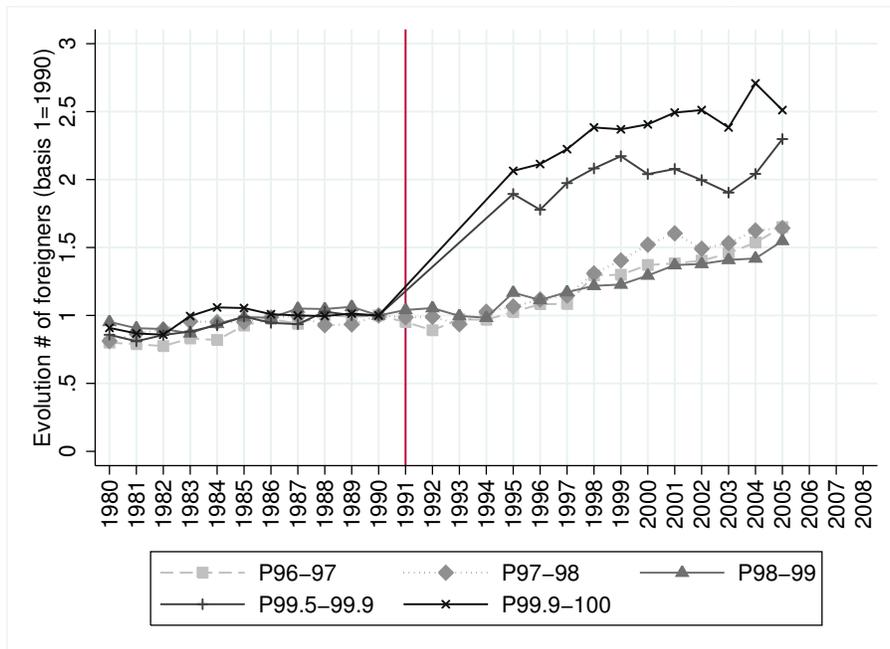


Figure 8: Total number of foreigners in different earnings groups

Panel A of the Figure plots the fraction of foreigners in various upper percentile groups of the distribution of earnings (including both Danish citizens and foreigners) from 1980 to 2005. The threshold for eligibility to the scheme is always between the 99th and the 99.5th percentile. P96-97 denotes all individuals between the 96th and 97th percentile, etc. There is a gap in 1991-1994 for the top two groups because the scheme related data does not provide scheme earnings for those years. Panel B repeats the same graph but normalizing the series to one in 1990 (just before the scheme introduction) to assess relative trends. The vertical line at year 1991 denotes the year the scheme was first implemented. Individuals are weighted by duration of stay for part-year residents.

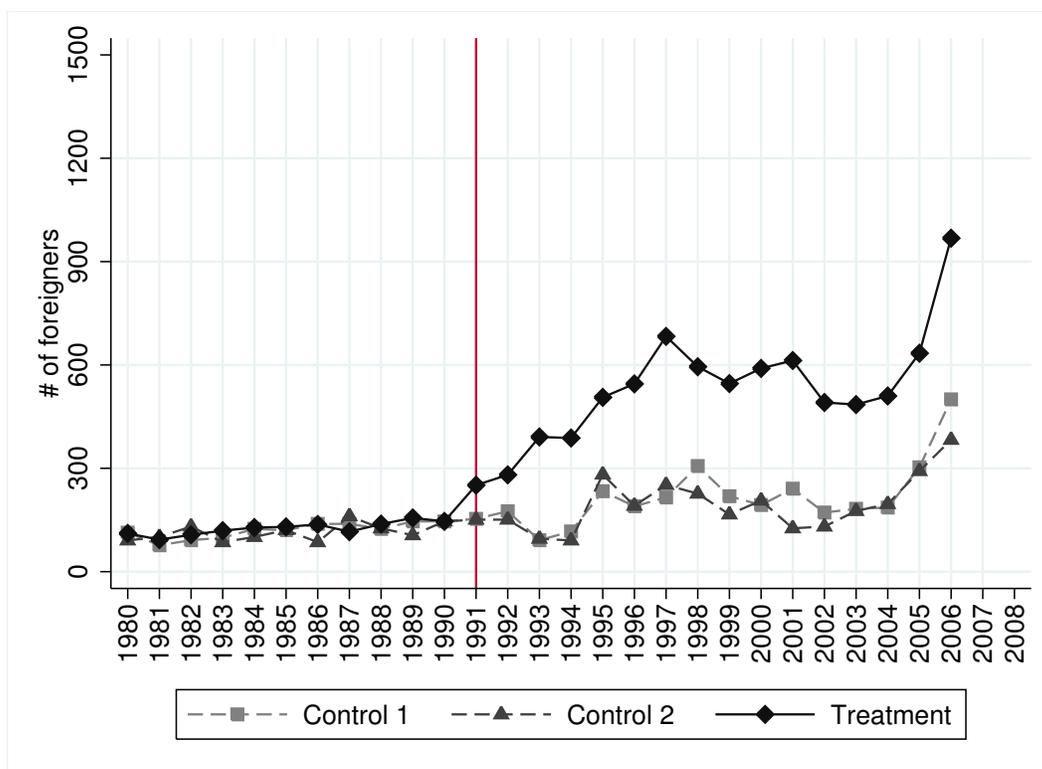


Figure 9: Total number of arrivals in different earnings groups

The Figure reports the number of foreigners with earnings above the scheme eligibility threshold (treatment series) arriving each year in Denmark from 1980 to 2006. As control groups, it reports the number of foreigners arriving 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.

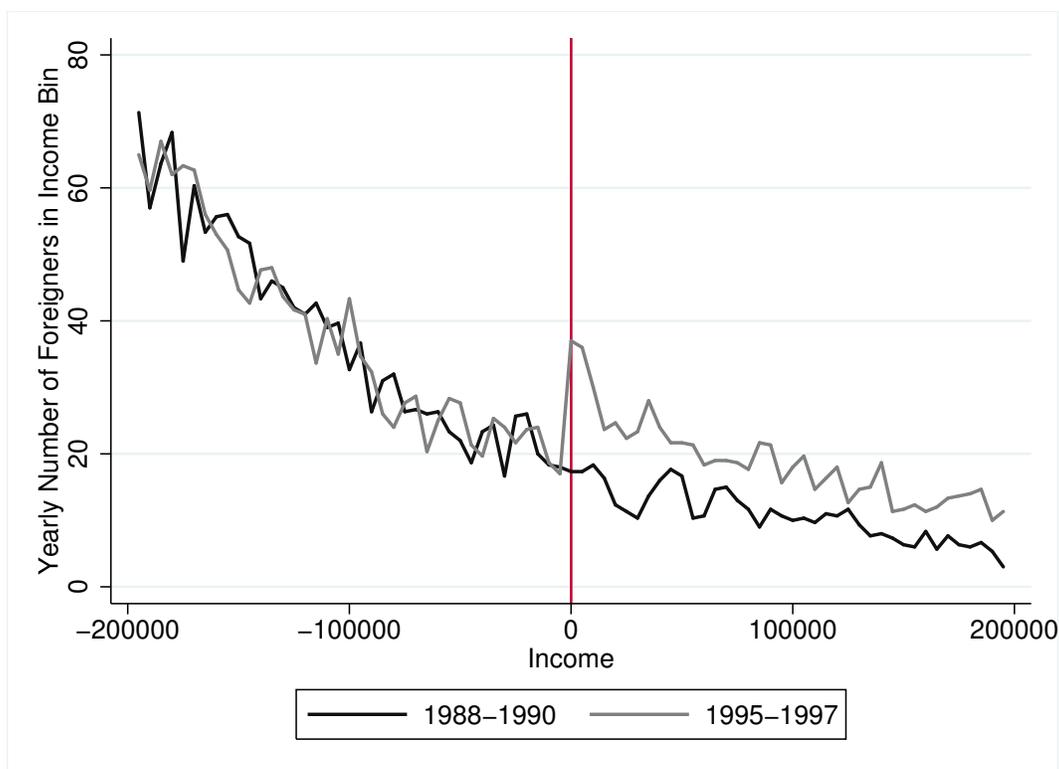


Figure 10: Earnings Density for Foreigners Before and After Scheme Introduction

The Figure reports the density of the earnings distribution of foreigners around the eligibility threshold (denoted by the vertical line) in 1995-7 (solid line after scheme implementation) and in 1988-90 (grey line before scheme implementation). For part year foreigners, earnings are annualized (i.e., divided by fraction of year present as scheme eligibility threshold is based on such annualized earnings). The graph shows that the scheme almost doubled the density above the threshold due to extensive migration responses and also created a small spike just at the eligibility threshold due to an intensive margin earnings response.

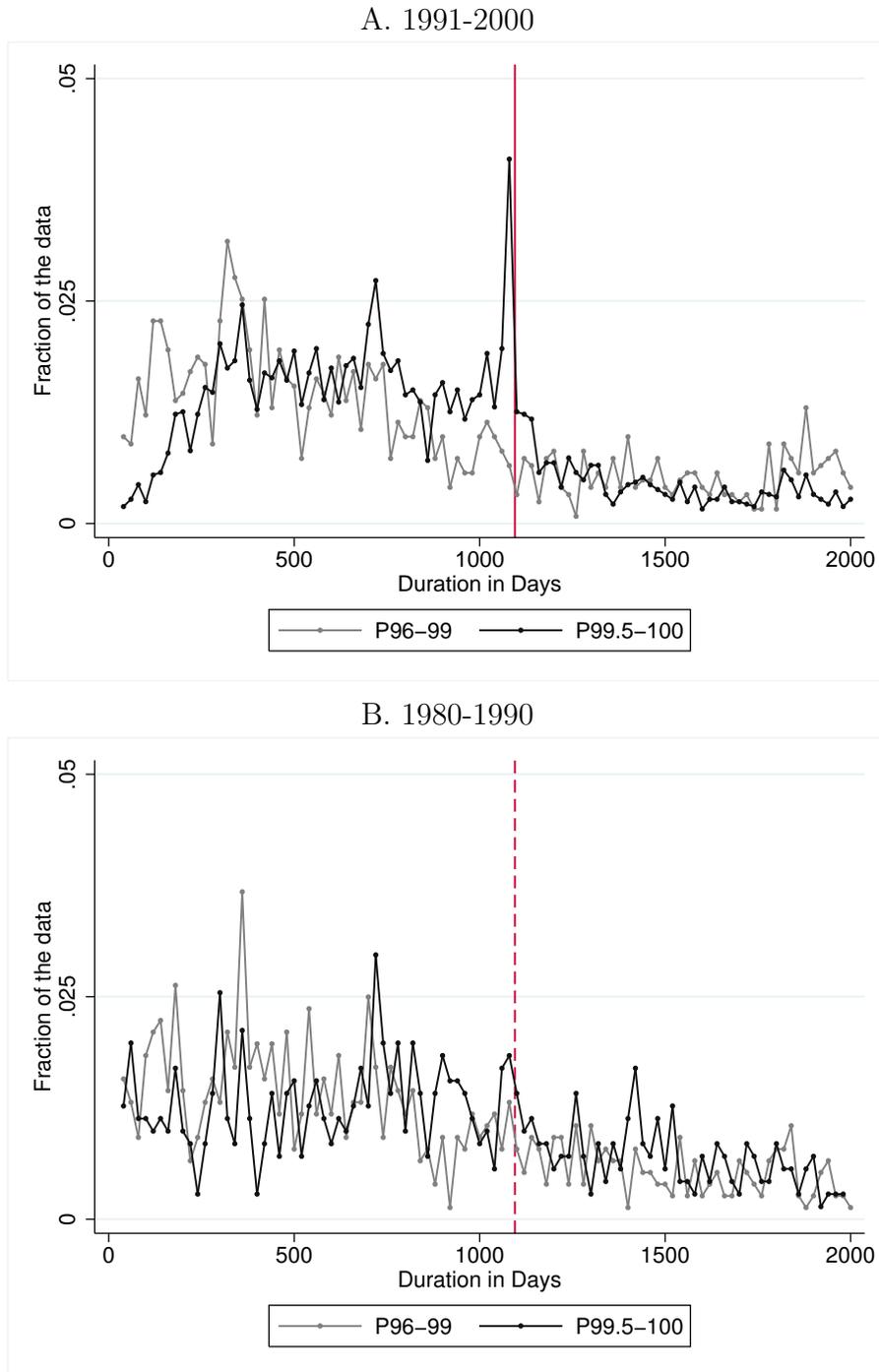


Figure 11: Duration of Stay Density for Foreigners

The Figure reports the density of duration of foreigners with (annualized) earnings above percentile 99.5th (P99.5-100) and hence above the eligibility threshold for the scheme and earnings between percentile 96 and percentile 99 (P96-99) and hence below the eligibility threshold for the scheme. Panel A is for years 1991-2000 (after the scheme was implemented) while Panel B is a control graph for years 1980-1990 (before the scheme was implemented). The vertical line denotes the 3-year threshold (maximum duration of the scheme). The spike at 3 year in Panel A (but not in Panel B) shows that the scheme has an impact on duration of stay along the intensive duration margin.

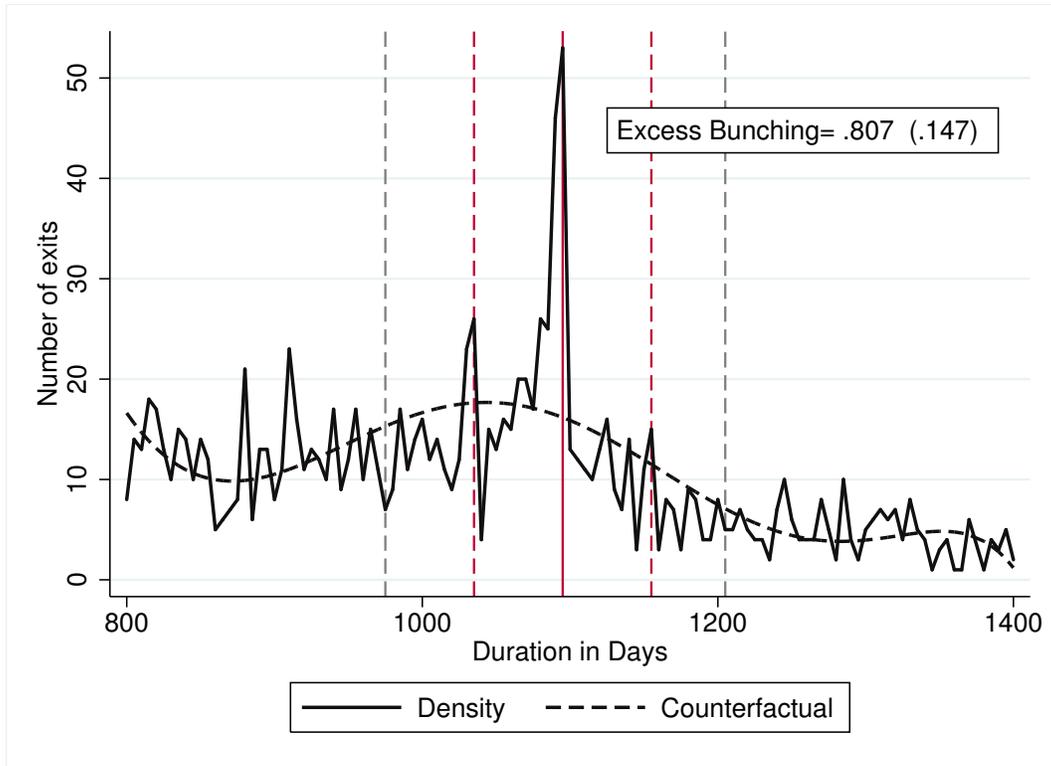


Figure 12: Excess Bunching at the 3-year Cut-off Duration

The Figure reports the density of duration of foreigners with (annualized) earnings above percentile 99.5th for years 1991-2005, zooming around the 3-year threshold (maximum duration of the scheme denoted in a vertical line). The spike at 3 year shows that the scheme creates excess bunching and hence that the scheme has an impact on duration of stay along the intensive duration margin.

Table 1: Difference-In-Difference Estimates of the Effect of the Tax Scheme on the Total Number of Foreigners

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All Foreigners				Nordic Countries	Continen- tal Europe	English Speaking Countries	Other
α_{tr*rf}	748.7*** (77.02)	421.8*** (53.44)	-30.7 (28.4)	706.5*** (48.6)	253.3*** (29.64)	202.1*** (20.87)	253.0*** (25.50)	40.30*** (10.13)
α_{tr}	-100.5* (43.05)	-100.5** (27.60)	-191.1*** (14.7)	177.44*** (27.19)	-17.19 (16.57)	-39.36** (11.67)	-32.31* (14.26)	-11.64* (5.662)
α_{rf}	194.8** (54.46)	24.27 (37.79)	17.77 (20.13)	102.94** (34.4)	34.03 (20.96)	49.04** (14.76)	82.66*** (18.03)	29.03*** (7.162)
α_0	734.1*** (30.44)	734.1*** (19.51)	462.62*** (10.39)	353.8*** (19.23)	290.3*** (11.72)	231.6*** (8.249)	174.9*** (10.08)	37.32*** (4.004)
$\frac{\alpha_{tr*rf}}{\alpha_{tr} + \alpha_{rf} + \alpha_0}$.90 (.16)	.65 (.11)	-.104 (.097)	1.11 (.16)	.825 (.16)	.84 (.15)	1.12 (.24)	.74 (.29)
Treatment	I>T		.9*T< I <T	I >P99.5	I>T			
Control	.8*T< I<T		.8*T< I <.9*T	P96< I <P99	.8*T< I <T			
Pre-reform	1980-1990	1980-1990	1980-1990	1980-1990	1980-1990	1980-1990	1980-1990	1980-1990
Post-reform	1995-2000	1991-1995	1991-1995	1991-1995	1995-2000	1995-2000	1995-2000	1995-2000
N	32	30	30	30	32	32	32	32

This table presents difference-in-difference estimates of the number of foreigners above the eligibility threshold (treatment) relative to foreigners between 80 and 100% of the threshold (control) before the scheme and after the scheme. The regression specification is $Y_{it} = \alpha_0 + \alpha_{tr} \cdot \mathbb{1}[i = 1] + \alpha_{rf} \cdot \mathbb{1}[t > 1991] + \alpha_{tr*rf} \cdot \mathbb{1}[i = 1] \cdot \mathbb{1}[t > 1991] + \nu_{it}$. The first row in the top panel presents the DD estimate α_{tr*rf} . The second panel presents the percent increase of foreigners due to the scheme, $\frac{\alpha_{tr*rf}}{\alpha_0 + \alpha_{rf} + \alpha_{tr}}$, where $\alpha_0 + \alpha_{rf} + \alpha_{tr}$ is the counterfactual number of foreigners that would have prevailed in the treatment group after 1991 in the absence of a reform. The bottom panel specifies the DD groups. Col. (3) is a placebo comparing two control groups.