CEPS INSTEAD

WORKING PAPERS

Measuring and accounting for the deprivation gap of Portuguese immigrants in Luxembourg

Vincent A. HILDEBRAND[1]
Maria Noel Pi ALPERIN[2]
Philippe VAN KERM[2]

York University, Canada[1]
CEPS/INSTEAD, Luxembourg[2]

Working Paper No 2012-33
August 2012

Measuring and accounting for the deprivation gap of Portuguese immigrants in Luxembourg*

Vincent A. Hildebrand†
Department of Economics, Glendon College
York University, Canada

María Noel Pi Alperin‡
CEPS/INSTEAD, Luxembourg

Philippe Van Kerm§
CEPS/INSTEAD, Luxembourg

July 2012

*This research is part of the MeDIM project (*Advances in the Measurement of Discrimination, Inequality and Mobility*) supported by the Luxembourg 'Fonds National de la Recherche' (contract FNR/06/15/08) and by core funding for CEPS/INSTEAD from the Ministry of Higher Education and Research of Luxembourg. This paper was partly done when Pi Alperin was supported by a 'Bourse Formation-Recherche' (BFR08-019) from the Ministry of Higher Education and Research of Luxembourg. Comments from participants at the IARIW General Conference 2010, the CORE 2012 conference on Multidimensional Inequality and Poverty Measurement, the CAE Annual Meeting 2012, and at a CEPS/INSTEAD seminar are gratefully acknowledged.

†Vincent Hildebrand, Department of Economics, York Hall 363, Glendon College, York University, 2275 Bayview ave., Toronto, Ontario, M4N 3M6 Canada. E-mail: `vincent@econ.yorku.ca`.

‡María Noel Pi Alperin, CEPS/INSTEAD, Av. de la Fonte 3, L-4364 Esch/Alzette, Luxembourg. Associated researcher at LAMETA, University of Montpellier. E-mail: `marianoel.pialperin@ceps.lu`.

§Philippe Van Kerm, CEPS/INSTEAD, Av. de la Fonte 3, L-4364 Esch/Alzette, Luxembourg. E-mail: `philippe.vankerm@ceps.lu`.

This paper examines the relative well-being of Portuguese immigrants in Luxembourg by looking at non-monetary, or 'direct indicators' of material deprivation. The paper not only documents deprivation differentials between immigrants and natives, but also models the association between material deprivation indicators, income and population characteristics in order to shed light on the sources of differentials. In particular, we measure how much income differentials explain differences in material deprivation. We find that answer to this question depend a lot on what deprivation indicators are taken into consideration (and a little on how aggregate material deprivation indicators are constructed). Income differences explain material deprivation differences entirely when the latter is measured according the European Commission's headline indicator on material deprivation. Inclusion of housing condition indicators mitigates this relationship and we then find compelling evidence that material deprivation is not entirely accounted for by income differentials.

Keywords: material deprivation ; Portuguese immigrants; inverse generalized Lorenz curve ; reweighting ; multidimensional poverty ; PSELL

1. Introduction

Since seminal work of Chiswick (1978) and Borjas (1985), studies on the economic assimilation of immigrants have been abundant. Economic assimilation is typically measured by immigrants' ability to eliminate an earnings penalty relative to native-born workers encountered at the time of entry on their host country's labour market; see recent contributions by Adsera and Chiswick (2007), Clark and Lindley (2009), Izquierdoa, Lacuestaa, and Vegas (2009), among many others. The literature has also explored the nativity gap of other economic dimensions believed to shape foreign-born assimilation such as occupational attainment (Green, 1999, Chiswick, 2002, Frenette, Hildebrand, McDonald, and Worswick, 2003), homeownership (Sinning, 2009), and wealth (Cobb-Clark and Hildebrand, 2006a, Bauer, Cobb-Clark, Hildebrand, and Sinning, 2011).

Recent findings reveal that the extent to which foreign-born assimilate to the host society differs widely between countries of origin (Adsera and Chiswick, 2007), nativity and/or racial background (Chiswick and Miller, 2009, Izquierdoa, Lacuestaa, and Vegas, 2009), immigration cohorts (Borjas, 1995) and the outcome under study (Aleksynska and Algan, 2010).

Another strand of this literature emphasizes the importance of social and cultural dimensions and points out that economic and social assimilation are not systematically correlated. Again, it varies by immigrant groups and countries of origin and destination (Dustmann, 1996, Aleksynska and Algan, 2010). 'Successful integration' –narrowly defined as the acquisition of attitudes and competences necessary to participate in the host society like any native-born (see Bosswick and Heckmann (2006) on concepts of assimilation and integration)– ought to be determined by considering several dimensions and the relative weight given to a particular dimension can vary depending on the group of migrants considered.

This paper considers a multidimensional measure to examine the socio-economic performance of Portuguese immigrants in Luxembourg. Specifically, we assess what we refer to as the 'deprivation gap' of Portuguese immigrants, a gap with regard to non-monetary measures of deprivation. Non-monetary measures of deprivation –also referred to as measures of 'material deprivation' or 'lifestyle deprivation'– are derived from questionnaire data on an array of direct indicators of living conditions, such as the possession of durable goods (e.g., a TV set, a refrigerator), the ability to afford basic consumption goods (e.g., heating, food, leisure), the capacity to face regular expenses (e.g., electricity bills, rent, interest on mortgage), housing conditions, etc. All items are usually widespread and considered basic for living a decent life in the society; see, e.g., Townsend (1979) for an early discussion.

Such measures have gained popularity among statistical agencies and in policy debates. A measure of material deprivation has been one of the EU's official headline indicators on social protection and social inclusion since 2006 (Atkinson, Cantillon, Marlier, and Nolan, 2002, European Commission, 2009) and the recent EU 2020 target on poverty and social exclusion is, in part, expressed in terms of material deprivation (Nolan and Whelan, 2011).[1] This approach is meant to concentrate directly on the standard of living of individuals rather than on the resources required to achieve those conditions, namely earnings or income. It has grown in fashion in recognition of limitations of income to adequately identify people experiencing poverty and to capture the multi-dimensional nature of deprivation (see, e.g., Atkinson, Cantillon, Marlier, and Nolan, 2002, Nolan and Whelan, 2010). Of particular concern, is that income largely misses accumulated savings and (financial and non-financial) wealth, which may be strong determinants of living conditions, in particular for some population subgroups such as the elderly population. This is supported by empirical evidence suggesting that income and material deprivation indicators do not perfectly overlap, and identify potentially different populations as most deprived (Nolan and Whelan, 1996). Arguments about income failing to properly reflect how accumulated wealth impacts on living conditions are germane to assessments of immigrants socio-economic assimilation too as migrants typically have specific savings behaviour and lower accumulated wealth than natives (Carroll, Rhee, and Rhee, 1994, Bauer, Cobb-Clark, Hildebrand, and Sinning, 2011, Mathä, Porpiglia, and Sierminska, 2011). However, to the best of our knowledge, relatively little use has been made of such measures to study immigrants performance; see Haisken-DeNew and Sinning's (2010) analysis of deprivation and social exclusion of

[1] Material deprivation measurement is also gaining popularity in North America. In 2009, the Canadian Province of Ontario started to collect data needed to derive a multidimensional deprivation index in the context of Ontario's latest poverty reduction program initiated by McGuinty's government.

immigrants in Germany for a recent example.

Luxembourg is the country with the largest share of foreign-born residents in the EU (European Commission, 2011). Portuguese immigrants form the largest foreign community in the country, with 41% of foreign residents in 2007 (Berger, 2008). As many foreign populations elsewhere in industrialized countries, they have been consistently reported to record lower socio-economic achievements than natives (and most other foreign communities) as measured by earnings and employment (Langers, 2006), by income (Hartmann-Hirsch, 2007), education (Alieva, 2010) or by indicators of satisfaction with financial conditions (Van Kerm and Villeret, 2007). This is noteworthy since –unlike its neighbouring countries– Luxembourg deliberately chose to favour entry of immigrants sharing comparable cultural heritage to facilitate their assimilation and foster social cohesion. Portuguese migrants' poor economic performance has been documented too in other European countries with a large Portuguese community: in France (Brinbaum and Kieffer, 2009) and in Germany (Kalter and Granato, 2002).

The objective of this analysis is effectively three-fold. First, we further document the substantive case of Portuguese immigrants, exploiting the relatively large sample size of Portuguese immigrants available in our data. We show that the deprivation gap is large and significant, and is robust over a broad spectrum of definitions of material deprivation. Second, we question the actual usefulness of deprivation indicators over more conventional income-based assessments by examining how much of the deprivation gap can be explained by income differentials between immigrants and natives. We do so after controlling for differences in demographic structure and employment of the two populations using a fully non-parametric version of sample reweighting techniques à la DiNardo, Fortin, and Lemieux (1996) and Barsky, Bound, Charles, and Lupton (2002). Our prominent finding is that the contribution of income differences to the deprivation gap crucially depends on the items considered in our indices of material deprivation. The gap is entirely accounted for by income (and demographic characteristics) when based on the nine deprivation items selected by the European Commission in its official indicator of material deprivation. However, a significant residual, unexplained gap remains once additional items are included to reflect differences in housing conditions. This observation questions the relevance of the current definition of the EU indicator of material deprivation, in particular in rich countries of the EU.

Third, and finally, the paper contributes to the methodology on material deprivation analysis which, despite the growth of theoretical and empirical work, still faces a number of unresolved issues. In particular, our analysis provides further evidence on the impact of item selection and item aggregation. It also illustrates a graphical instrument for reporting and comparing deprivation distributions.

The paper is organized as follows. We present our methodological approach in Sections 2 and 3 and detail our sample data in Section 4. We discuss results from our application in section 5 and offer some concluding remarks and suggestions for future research in Section 6.

2. Measures of material deprivation

In the vast majority of studies on material deprivation, the degree of deprivation for an individual i is determined from a vector of data $\mathbf{d}_i \equiv (d_{i1}, d_{i2}, \ldots, d_{iK})$ where each d_{ik} is a binary indicator variable taking the value 1 if individual i is 'deprived' of a particular commodity (e.g., some basic consumption, or durable good) or experiences a given 'bad' (such as financial stress, etc.). There is no agreement as to what constitutes a relevant item in such a vector. For example, the EU indicator of material deprivation considers only nine items: whether a household (*i*) can face unexpected expenses; (*ii*) can afford one week annual holiday away from home; (*iii*) has the capacity to pay regular bills (mortgage or rent, utility bills or hire purchase installments); (*iv*) has the capacity to afford meat, chicken or fish every second day; (*v*) has the capacity to keep home adequately warm; (*vi*) could possess a washing machine (if desired); (*vii*) could possess a colour TV; (*viii*) could possess a telephone; and (*ix*) could possess a personal car (European Commission, 2009, Nolan and Whelan, 2011). Authors have however often considered broader sets of indicators by including, e.g., indicators of housing conditions or of social interactions (see e.g Whelan, Nolan, and Maître, 2008). In this context, despite the obvious limitation of considering only nine items, the official status of the EU definition makes it a fixed point of reference, which we will compare to some more encompassing measures (detailed in Section 4).

Summary indices of material deprivation are typically derived on the basis of sample data on $\{\mathbf{d}_i\}_1^N$:

$$S = \frac{1}{N}\sum_{i=1}^{N} s_i,$$

where $s_i = \sum_{k=1}^{K} w_k d_{ik}$ is a linear combination of the K deprivation items, and with w_k being a non-negative weight reflecting the relative importance of item k in contributing to individual-level deprivation. This approach to constructing synthetic indices of deprivation is the simplest and most common strategy in the literature; see, e.g., Deutsch and Silber (2005), Fusco (2007) for reviews.

Practitioners have also adopted an array of approaches to determine item weights; see Decancq and Lugo (2012) for a recent survey. We consider three weighting schemes.[2] The first and simplest scheme is 'equal weighting':

$$w_k^{\text{eq}} = \frac{1}{K}.$$

This is also known as the 'counting approach' (Atkinson, 2003, Alkire and Foster, 2011). The other two schemes are 'frequency-based' where item weights are proportional to the prevalence of the deprivation item in the population. The second scheme, proposed by Cerioli and Zani (1990), takes:

$$\omega_k^{\text{cz}} \propto \log\left(\frac{1}{\bar{d}_k}\right),$$

where $\bar{d}_k$ is the mean of item d_{ik} in our sample. This scheme assigns higher weight to relatively infrequent deprivation items: the weight is higher, the most 'unusual' is a deprivation. The third scheme, inspired from Betti and Verma (1998), adopts a more sophisticated double-weighting rule sensitive to both the relative frequency of items and the correlation among items. The correlation is taken into account so that two perfectly correlated items 'count as one' and only two uncorrelated items fully 'count as two'. To achieve this, Betti and Verma (1998) and Betti, Cheli, Lemmi, and Verma (2008) defined item weights as the product of two components:

$$\omega_k^{\text{bv}} \propto \left(\omega_k^a \times \omega_k^b\right),$$

[2] Without loss of generality, we normalize item weights to sum to unity in all cases so S can be interpreted as a form of average deprivation; see *supra*.

with ω_k^a being the coefficient of variation of d_{ik} acting similarly to the frequency-based weights described above, and:

$$\omega_k^b = \left(1 + \sum_{m=1}^{M} \rho_{km} \mathrm{I}\Big(\rho_{km} < \rho_H\Big)\right)^{-1} \left(\sum_{m=1}^{M} \rho_{km} \mathrm{I}\Big(\rho_{km} \geq \rho_H\Big)\right)^{-1},$$

where ρ_{km} is the correlation between items k and m and $\mathrm{I}(\cdot)$ is an indicator function equal to 1 if the expression in brackets is true and 0 otherwise. ρ_H is a pre-determined cut-off correlation level.[3] ω_k^b is the inverse of a measure of 'average correlation' of item k with all the other items. The larger is the average correlation with item k, the lower is the resulting weight for item k. In our application, for using a consistent set of schemes, we modify the Betti and Verma (1998) approach by adopting the Cerioli and Zani (1990) weight as the first component $\omega_k^a = \omega_k^{\mathrm{cz}}$.

In all three cases, normalization to unity is achieved by setting:

$$w_k = \frac{\omega_k}{\sum_{k=1}^{K} \omega_k}.$$

Note that we evaluate the deprivation of all individuals (natives and immigrants) on the basis of a common set of item weights. We discard any potential interpersonal or inter-group (cultural) differences in social perceptions about the importance of different items. We take it as desirable in our context as we do not want the immigrants/natives comparison to be mitigated (or exacerbated) by adopting different sets of item weights (but see Haisken-DeNew and Sinning (2010) for a different point of view).[4]

In sum, the aggregate index S is an average of individual-level deprivation contributions s_i, which itself is a linear combination of specific deprivation items d_{ik}. We provide a complementary, graphical illustration of the distribution of the individual-level s_i, as it may be argued that the details of the distribution of individual s_i are lost in the averaging over individuals. Behind a given value of S, it may happen that a small number of individuals are deprived on many items, or that a larger population is deprived on just a few items. Whether the former is preferable to the latter is related

[3] ρ_H separates high and low correlations. Betti and Verma (1998) suggest setting ρ_H as to divide the ordered set of correlations at the point of the largest gap.

[4] See Dickes, Fusco, and Marlier (2010) on variations in social perceptions of deprivation items in the European Union, and Bellani and D'Ambrosio (2011) for a methodological discussion.

to discussion about the 'union' vs. 'inter-section' approaches to aggregating deprivation indicators (see e.g Duclos, Sahn, and Younger, 2006).

As a graphical instrument for depicting the distribution of s_i, we opt for the inverse generalized Lorenz (IGL) curve introduced for comparing distributions of poverty gaps in Spencer and Fisher (1992) and Jenkins and Lambert (1997). The IGL curve plots the cumulative share of the sample $p = \frac{i}{N}$ against the cumulative deprivation accumulated by the fraction p of the most deprived individuals:

$$\mathrm{IGL}(p) = \frac{1}{N} \sum_{i=1}^{pN} s_{(i)},$$

where the sample values s_i are ordered such as $s_{(1)} \geq s_{(2)} \geq \cdots \geq s_{(N)}$. As discussed in Jenkins and Lambert (1997), the IGL curve provides compact description of the incidence, intensity and inequality dimensions of the distribution of individual deprivation.[5] The value of the IGL curve at $p = 1$ gives index S, namely the average deprivation score. The point S_0 on the horizontal axis at which the IGL curve becomes flat gives the fraction of individuals deprived on at least one item ($S_0 = \inf\{p \text{ s.t. } \mathrm{IGL}(p) = \mathrm{IGL}(1)\}$). Also, the greater the curvature of the IGL curve the more is deprivation concentrated on few individuals with high degree of deprivation –in other words, the more is deprivation unequally distributed in the population; see Aaberge and Peluso (2011) for a discussion and dominance criteria.

3. Accounting for the 'deprivation gap'

We refer to the difference in our aggregate measures of deprivation between Portuguese immigrants and Luxembourg natives as the 'deprivation gap' of Portuguese immigrants. As we show *supra* the deprivation gap is substantial. But the populations compared are also different, with Portuguese immigrants being in general much younger and with lower levels of educational achievements, for example. Levels of income are also different. Assessing how much of the deprivation gap is just a reflection of an 'income gap' is central to judging the usefulness of deprivation measures for policy as a complement to income-based indicators.

[5] Spencer and Fisher (1992) and Jenkins and Lambert (1997) in fact consider distributions of (income-based) poverty gaps. The analogy to distributions of deprivation score is direct.

To capture the extent to which the gap can be accounted for by differences in population characteristics, we implement a reweighting-based standardization approach similar to DiNardo, Fortin, and Lemieux (1996), Barsky, Bound, Charles, and Lupton (2002) or Cobb-Clark and Hildebrand (2006b). In a nutshell, the procedure involves determining adjustment weights for each observation in our sample of natives where weights are calibrated so that the weighted sample of natives has the same characteristics as the sample of immigrants on a set of key control covariates. Computing the deprivation gap with the adjusted sample identifies a gap netted out of the effect of population characteristics. We apply this procedure to control for population differences with respect to the age of household head, household composition, the education of household head, the labour market attachment of the household, and finally to differences in household income. The adjustments are sequential and cumulative. We first control for age differences, then for household composition conditional on age adjustments, and so on, until we finally control for income differences conditionally on all previous characteristics. We therefore treat the dependence between covariates using an explicit chain of conditional distributions and consider the marginal impact of netting out each of the covariates in turn to identify their impact on the deprivation gap.

Consider the vector $\mathbf{d} \equiv (d_{i1}, d_{i2}, \ldots, d_{iK})$ of binary values on K deprivation items for an individual i.[6] We view $\mathbf{d}$ to be a realization from a K-variate multinomial random variable Θ with associated probability distribution $p_g(\cdot)$:

$$p_g(\mathbf{d}) \equiv \Pr[\Theta = \mathbf{d} \mid I = g],$$

where I is a group indicator denoting the population to which observation i belongs to – i.e., $I = 0$ for natives and $I = 1$ for Portuguese immigrants. Covariates are introduced by expressing $p_g(\mathbf{d})$ as a function of *conditional* deprivation probabilities, averaged over covariate distributions:

$$p_g(\mathbf{d}) = \sum_{z \in \Omega_Z} \Pr[\Theta = \mathbf{d} \mid Z = z,\ I = g] \times \Pr[Z = z \mid I = g],$$

where Ω_Z is the 5-dimensional domain of definition for the combination of the five

[6] In what follows, the subscript i is omitted from $\mathbf{d}$ for clarity.

covariates considered and $\Pr[Z = z \mid I = g]$ is the probability of observing a particular configuration of covariates z in population g.[7]

To identify the distinct impact of netting out differences in each of our five covariates in turn, we further factorize $\Pr[Z = z \mid I = g]$ into a sequence of five conditional probabilities:

$$
\begin{aligned}
p_g(\mathbf{d}) = \sum_{s\in\Omega_A}\sum_{t\in\Omega_H}\sum_{u\in\Omega_E}\sum_{v\in\Omega_L}\sum_{w\in\Omega_Y} & \Pr[\Theta = \mathbf{d} \mid A = s, H = t, E = u; L = v, Y = w, \;\; I = g] \\
& \times \Pr[A = s \mid \;\; I = g] \\
& \times \Pr[H = t \mid A = s, \;\; I = g] \\
& \times \Pr[E = u \mid A = s, H = t, \;\; I = g] \\
& \times \Pr[L = v \mid A = s, H = t, E = u, \;\; I = g] \\
& \times \Pr[Y = w \mid A = s, H = t, E = u, L = v, \;\; I = g], \quad (1)
\end{aligned}
$$

where Ω_A, Ω_H, Ω_E,Ω_L and Ω_Y are the (unidimensional) domains of definition of each of the five covariates. This factorization of the probability distribution $p_g(\cdot)$ provides a direct way to construct counterfactual deprivation distributions for natives 'as if' they had the covariate distributions of immigrants. We define the generic counterfactual distribution:

$$
\begin{aligned}
p_g^{ahely}(\mathbf{d}) = \sum_{s\in\Omega_A}\sum_{t\in\Omega_H}\sum_{u\in\Omega_E}\sum_{v\in\Omega_L}\sum_{w\in\Omega_Y} & \Pr[\Theta = \mathbf{d} \mid A = s, H = t, E = u; L = v, Y = w, \;\; I = j] \\
& \times \Pr[A = s \mid \;\; I = a] \\
& \times \Pr[H = t \mid A = s, \;\; I = h] \\
& \times \Pr[E = u \mid A = s, H = t, \;\; I = e] \\
& \times \Pr[L = v \mid A = s, H = t, E = u, \;\; I = l] \\
& \times \Pr[Y = w \mid A = s, H = t, E = u, L = v, \;\; I = y],
\end{aligned}
$$

where each of the superscripts a, h, e, l, y can take on values of 0 (for using the

[7] All covariates considered in the paper are ordinal or nominal by construction; see Section 4.

conditional probabilities for natives) or 1 (for using the conditional probabilities for Portuguese immigrants). The probability distribution for natives is therefore $p_0(\mathbf{d}) = p_0^{00000}(\mathbf{d})$, and that for Portuguese immigrants is $p_1(\mathbf{d}) = p_1^{11111}(\mathbf{d})$. Mixed patterns on the '*ahely*' parameters correspond to counterfactual distributions, e.g., $p_0^{10000}(\mathbf{d})$ is the deprivation distribution that would be observed among natives if they had the same age distribution as immigrants but all else remained unchanged; $p_0^{11110}(\mathbf{d})$ is the counterfactual deprivation distribution 'as if' natives had the same age, household structure, education and labour market attachment as immigrants but retained their income level and deprivation conditionally on the other characteristics; $p_0^{11111}(\mathbf{d})$ is the counterfactual deprivation distribution 'as if' natives had the same characteristics as immigrants, yet retained their deprivation probabilities given covariates.

Noting that all our measures of interest –aggregate deprivation indices or coordinates of IGL curves– are functionals of the probability distributions of deprivation vectors, we can use the counterfactual distributions constructed to generate counterfactual indices or IGL curves. We denote generally any such measure as $\theta(p_g^{ahely}) \equiv \theta_g^{ahely}$.

The impact of, say, factor a (the difference in age composition) on the aggregate measure is then estimated by $\theta_g^{1hely} - \theta_g^{0hely}$ (and similarly for all other factors, e.g., $\theta_g^{ahe1y} - \theta_g^{ahe0y}$ for factor l (differences in labour market attachment given age and household structure)), that is by the marginal change of the index when a particular element of the covariate distribution is changed from that of natives to that of immigrants.

Note that there is a standard index problem in this approach: we have not specified at what baseline values for *hely* and g the marginal effect of a is evaluated (or at what values of *ahey* and g for the effect of l, etc.). In principle, any of the $2^5 = 32$ possible combinations of reference covariates is valid. In similar situations, some authors have suggested to estimate the marginal effect on all possible combinations (or some relevant subset thereof) and average over resulting estimates to bypass the problem of selecting relevant baseline combinations (see Cobb-Clark and Hildebrand, 2006b, Devicienti, 2010). In our application, we argue that a sequential approach is most attractive because of the large differences in population composition, and the restricted range of covariate values observed for the immigrant group. As illustrated below, Portuguese immigrants are severely under-represented among the elderly, the highly educated and high income deciles, and they are absent from numerous 'cells' when characteristics are combined. We therefore proceed by taking the natives sample and sequentially reweighting its covariate distribution to that of immigrants, factor after factor. In effect this also implies gradually discarding observations from the natives sample that have no 'twins' in the immigrants sample. Alternative evaluations requiring 'inflating' the immigrants sample to the natives sample distribution would be infeasible in our case as many covariate configurations are absent from the immigrants sample (in particular for high income or old people) and it is therefore impossible to replicate the natives sample characteristics from the immigrants sample characteristics by reweighting.[8]

In sum, we first adjust the natives distribution for age, then consider the finer partition by age and household structure, then the finer partition by age, household structure and education, and so forth. We estimate the effect of age differences on the deprivation gap as $C_a^\theta = (\theta_0^{10000} - \theta_0^{00000})$, household structure as $C_h^\theta = (\theta_0^{11000} - \theta_0^{10000})$, education as $C_e^\theta = (\theta_0^{11100} - \theta_0^{11000})$, labour market attachment as $C_l^\theta = (\theta_0^{11110} - \theta_0^{11100})$, and income as $C_y^\theta = (\theta_0^{11111} - \theta_0^{11110})$. Finally, we define the residual, unexplained difference as $C_r^\theta = (\theta_1^{11111} - \theta_0^{11111})$ which reflects the deprivation gap that would be observed if natives had the same characteristics as immigrants (including the same income). Combining these components leads to a simple, 'natural' additive decomposition of the deprivation gap:

$$\Delta \equiv \theta_1 - \theta_0 = (C_a^\theta + C_h^\theta + C_e^\theta + C_l^\theta + C_y^\theta) + C_r^\theta, \tag{2}$$

where the key components of interest to us are C_y^θ (the impact of income differences on the deprivation gap) and C_r^θ the unexplained gap.

Practically, we follow DiNardo, Fortin, and Lemieux (1996), Barsky, Bound, Charles, and Lupton (2002), Cobb-Clark and Hildebrand (2006b) to construct sets of reweighting factors which, when applied to the natives sample, results in a (weighted) sample having the same covariate distribution as the immigrants sample. The procedure (which does not require estimation the $p_g^{ahely}(\cdot)$ counterfactual distributions directly) is detailed in Appendix A. Counterfactual values for aggregate deprivation measures and IGL curves

[8] In addition, in cases where some 'twins' are observed, but are rare in the natives sample, they would be assigned very large reweighting factors, therefore leading to large sampling variability. See Barsky, Bound, Charles, and Lupton (2002) for a similar discussion. This problem is also closely related to the common support restrictions in matching procedures (Lechner, 2008).

($\hat{\theta}_0^{ahely}$) are estimated by incorporating these reweighting factors into equation (1) and (2) in a straightforward manner (re-weighting factors are treated just like sampling weights).

The sampling variability of all estimates presented in the paper is estimated using a non-parametric block bootstrap resampling procedure detailed in Appendix B. Resampling methods make it possible to assess the sampling variability of the relatively complex statistics considered here and to take into account the relatively complex dependence of the sample data (induced by the stratified PSELL survey design and our pooling of multiple waves of data described shortly).

4. Data and descriptive statistics

4.1. Sample

We use data from the Panel Socio-Economique *Liewen zu Lëtzebuerg* (PSELL-3), a general purpose panel survey carried out annually since 2003 with an initial sample of over 3,500 households representative of the population living in private dwellings in Luxembourg. Analysis is conducted on pooled data from waves 3, 4 and 5 (covering the 2005–2007 period) which contain comparable deprivation indicators related to the enforced lack of a combination of items depicting material living conditions.[9] 'Enforced' is understood as lacking possession due to insufficient financial resources, not by choice (see Nolan and Whelan, 1996, McKay, 2004, for further discussion).

We restrict our sample to all native and Portuguese households whose reference person is 16 years old or more. We eliminate all 'mixed households' so that both partners of couple-headed households are Luxembourg citizens born in Luxembourg while both partners in couple-headed Portuguese households are Portuguese citizens. After excluding all observations with missing values on any of the variables used in our empirical analysis, our estimation sample includes 5,020 native and 1,321 Portuguese household-year observations.

[9] Earlier waves of data either did not contain all deprivation items considered or used a different wording of questions which lead to slight inconsistencies.

4.2. Deprivation indicators

We exploit a total of seventeen deprivation indicators of three broad categories: economic strain, non-possession of common durable goods, and housing conditions. Economic strain is related to the inability to afford most basic needs including the capacity to face unexpected expenses; to eat meat or fish every second day (if the households wanted to); to pay for a week of annual holiday away from home; to keep home (household's principal residence) adequately warm; and, the inability to meet scheduled payment such as mortgage payments, accommodation or hire purchase installments. Non-possession of durables is related to the lack of widely desired durable goods: a TV set, a phone, a computer, a dishwasher, a car or van for private use. Housing conditions capture both the absence of basic housing amenities and the existence of serious problems associated with the family home including having a leaky roof, having damp walls, windows or grounds, having rot in walls, windows or grounds, having non-hermetic windows and doors, not having double glazing windows and not having an outdoor space.

All indicators are binary with a value of 1 indicative of deprivation and 0 otherwise. Sample means of all indicators are reported, separately for native and Portuguese households, in Table 1. Portuguese households are more frequently deprived than native households for each item. The difference in the capacity to face unexpected expenses is particularly striking: 45% of all Portuguese households report difficulties to face unexpected expenses compared to only 12% amongst natives. Percentage point differentials in the possession of durables is smaller than for economic strain indicators for all items but the possession of a computer. But the number of households lacking common durables is generally trivial (except for the possession of a computer). By contrast, we observe large differences between the two groups for all items within the 'housing conditions' dimension with the lack of possession of an outdoor space being the most striking one – about 30% among Portuguese households versus only 5% among natives from Luxembourg.

Decision over the subset of items to consider for computing the aggregate index S is a non trivial and largely unsettled issue. In the absence of consensus or compelling arguments about particular choices (see, e.g., Klasen, 2000, Guio, Fusco, and Marlier, 2009, Nolan and Whelan, 2010) our application reports estimates based on four different

Table 1: Deprivation item means by nationality

	Luxembourgish	Portuguese	Diff.
1. Capacity to face unexpected expenses	0.120	0.453	0.333*
2. Capacity to eat meat/fish	0.010	0.026	0.016*
3. Capacity paying a week holiday	0.064	0.208	0.144*
4. Capacity to keep house warm	0.002	0.016	0.014*
5. Difficulty paying bills	0.015	0.044	0.029*
6. Inability to pay rent/mortgage	0.008	0.036	0.029*
5.+6. Inability pay rent/mortgage or bills (combined)	0.017	0.056	0.039*
7. Possession of TV set	0.000	0.001	0.001*
8. Possession of phone	0.000	0.001	0.001
9. Possession of dishwasher	0.001	0.003	0.002
10. Possession of computer	0.009	0.093	0.085*
11. Possession of car	0.007	0.023	0.016*
12. Leaky roof	0.040	0.054	0.014
13. Rot in house	0.058	0.128	0.070*
14. Damp in house	0.094	0.135	0.040*
12.+13.+14. Leaky roof/Damp/rot (combined)	0.123	0.184	0.062*
15. Double glazing	0.132	0.219	0.087*
16. Hermetic doors/windows	0.102	0.140	0.039*
17. Outdoor space	0.051	0.305	0.254*

Note: Stars indicate that differences are statistically significant at 90 per cent confidence levels. Combined items are equal to 1 if any of the deprivation items combined is equal to 1.

subsets defined as follows (the exact list of items used in each of the four sets is in Appendix Table A1).

The first set –which we refer to as 'EU set'– includes the nine items that have been selected to construct the official EU indicator of material deprivation and now headline indicator in the EU Horizon 2020 strategy. This is a small set of basic deprivation items. It does not take any of our housing conditions indicators into account. The second set –which we refer to as our 'minimal' set– also relies on just a few items which we consider important and exhibit significant differences between natives and Portuguese immigrants. In contrast to the EU set, it includes one item on housing condition but excludes items related to the possession of a TV set and a phone. The latter two indicators are discarded because these deprivations are almost absent from our samples (e.g., only 0.2% of Portuguese households are deprived of a TV set). The third set –which we will refer to as our 'maximal' set– includes all seventeen available indicators. Fourth and finally, we chose an 'intermediate' set as a middle range between the 'minimal' and 'maximal' sets. The 'intermediate' set balances the number of items from the three broad dimensions, and only includes items which are frequently considered relevant in similar studies (see for example, Layte, Whelan, Maître, and Nolan, 2001, Guio, Fusco, and Marlier, 2009, Pi Alperin, 2010).

For the 'EU set', we will only consider the equal weighting case to keep in line with the EU indicator. For 'minimal', 'maximal' and 'intermediate' sets we experiment with the three weighting schemes presented in Section 2. This results in ten alternative versions of the aggregate summary index S (shown in Section 5), spanning a range of positions about how to select and weight deprivation items in the construction of a synthetic indicator. Respective item weights w_k in these ten schemes are reported in Appendix Table A1.

Because Portuguese immigrants fare worse than Luxembourg natives in *all* items taken separately (Table 1), the synthetic indicator S will obviously be higher for the latter than for the former, irrespectively of the aggregation function used. However, (i) the magnitude of the difference will vary with the weighting scheme, and (ii) the degree to which differences are explained by socio-economic characteristics and income will also differ significantly with the weighting scheme as shown shortly.

4.3. Demographic characteristics

Differences in deprivation level between Luxembourgish and Portuguese households are marked, but our sample composition with respect to basic socio-economic household characteristics also differ widely. Household characteristics in our sample are summarized in Table 2. Portuguese households are younger than natives. Pensioner households are significantly more prevalent among natives (18% versus 2%) while Portuguese are more likely to live in families with children (73% versus 48%).

It is therefore natural to start our decomposition by accounting for differences in households' age (a) and family types (h).[10] In addition, as Portuguese are younger than natives and wages are expected to be an increasing function of labour market experience, identifying the independent effect of age on the deprivation gap is interesting in itself.

Portuguese exhibit significantly lower educational attainment than the natives with only 19% of Portuguese households having at least completed secondary school compared to 68% of natives. It is now well established that poor educational attainment has long term negative consequences on individual well-being (Blanchflower and Oswald, 2004). Hence, we consider the role of human capital (e) as a third factor explaining the deprivation gap. We create three indicator variables capturing the highest level of education completed by the head of household including primary education (or without any formal education), high-school education, and post-secondary education.

Respondents from Portuguese households also report higher labour market participation than their native counterparts with about 60% of the latter having more than one active member participating in the labour market compared to less than 40% amongst natives. Greater labour force participation of Portuguese households could reflect age differences and/or a behavioural response of immigrants to compensate for the significantly lower compensation level of their active members. As a result, we consider the role of labour market participation (l) as another potentially relevant explanatory factor. To this end, we create three indicator variables (i) households without any active respondents, (ii) households with one active respondent and (iii) households with two or more active respondents.

[10] We consider four different household types: (i) couples without any children and single persons, (ii) single parents, (iii) couples with two or fewer children and, (iv) couples with three or more children.

Finally, the degree of income polarization by nativity group is stunning. Over 89% of Portuguese households are found in the bottom four deciles of the income distribution while about 60% of natives are in the top four. Given the contention –widely conveyed in the European Union by the objectives set by the 'Lisbon Treaty' to promote social inclusion (Cantillon, Verschueren, and Ploscar, 2012)– that income based measures are unlikely to fully capture all facets of poverty and deprivation, we consider the impact of income (y) on the nativity gap.[11] We purposely consider income as our last factor in the sequence used to implement our decomposition exercise as it allows us to measure the impact of income purged of basic demographic differences in the populations.

Table 2: Sample composition by nationality group (in percentage)

	Luxembourgish	Portuguese
Age of head of household (a)	*100*	*100*
Below 25	1	2
25–49	48	78
50–64	29	17
65+	22	3
Household structure (h)	*100*	*100*
Pensioners households (1 or 2 person aged 65+, no kids)	18	2
Single adults (with or without kids)	12	6
Couples (no kids)	22	19
Families (couples with kids)	48	73
Education level of head of household (e)	*100*	*100*
Lower secondary or below	32	81
Upper secondary	44	18
Higher education	24	1
Employment intensity in household (l)	*100*	*100*
No active adult	26	5
One active adult	36	31
Two or more active adults	38	63
Income decile group (y)	*100*	*100*
1st decile group	6	27
2nd decile group	6	25
3rd decile group	8	16
4th or 5th decile group	21	20
6th or 7th decile group	23	8
8th, 9th or 10th decile group	36	4
Sample size (# household-years)	5020	1321

[11] We use a measure of annual equivalent household income derived by applying the conventional modified-OECD equivalence scale to make household income levels comparable across household types.

5. The deprivation gap of Portuguese immigrants in Luxembourg

We now bring empirical evidence on our three main questions of interest. We document the size and structure of the deprivation of Portuguese immigrants, assess how much this can be accounted for by differences in population characteristics, employment and income, and finally discuss how much results are affected by item selection and choices on item weights – with particular scrutiny on the EU official indicator of material deprivation. We first briefly report on aggregate indicators and then provide a more detailed discussion on the basis of inverse generalized Lorenz curves.

As discussed *infra*, our analysis considers a total of ten different indices with alternative composition and weighting schemes. Index values are reported in Table 3. The first line shows the index value computed on the raw data, the second line gives the deprivation gap (the difference between estimates for Portuguese and for (reweighted) Luxembourg natives), the third line gives the marginal impact of introduction of each factor in turn and the last line gives variability bands thereof. Stars indicate that bootstrap variability bands for the deprivation gap or marginal impact do not include zero.

We find a large positive and significant raw gap for all of the ten indices (compare the first and last columns or the 2nd line of column 1 from Table 3). There are however variations in the levels of the aggregate deprivation indicators and in the magnitude of the differentials according to the index composition and weighting scheme. Estimates are not substantially affected by either using the 'Cerioli and Zani' (C–Z) weighting scheme or the 'modified Betti-Verma' (B–V) scheme for the three item sets for which they are considered. This is in contrast to equal weighting indicators which, by construction of the weights, records higher levels of deprivation.

The deprivation index of Portuguese immigrants is about 150 per cent larger than that of natives for the 'maximal set' (0.111 vs. 0.042), up to approximately 250 per cent larger for the 'EU set' (at 0.087 vs. 0.025). Accounting for differences in household's age, household type, educational attainment and labour market participation (AHEL) explains only between 10 and 20 per cent of this deprivation gap. This reduction is almost entirely explained by differences in educational attainment. Income differences (after controlling for the previous four factors) explain a much larger portion of the

Table 3: Aggregate deprivation indices on raw and reweighted samples

	LU						PT
	Raw	A	AH	AHE	AHEL	AHELY	Raw
EU material deprivation set							
Equal weights	0.025	0.030	0.026	0.048	0.046	0.080	0.087
	0.063*	0.057*	0.061*	0.040*	0.042*	0.008	
		0.005*	−0.004*	0.021*	−0.002	0.034*	
		[0.003;0.008]	[−0.006;−0.002]	[0.014;0.029]	[−0.006;0.002]	[0.017;0.046]	
Minimal item set							
Equal weights	0.056	0.065	0.058	0.090	0.086	0.139	0.189
	0.133*	0.124*	0.131*	0.099*	0.103*	0.050*	
		0.010*	−0.007*	0.032*	−0.005	0.053*	
		[0.006;0.014]	[−0.011;−0.004]	[0.019;0.045]	[−0.011;0.003]	[0.030;0.073]	
C–Z weights	0.041	0.048	0.042	0.066	0.062	0.105	0.145
	0.104*	0.098*	0.104*	0.080*	0.083*	0.040*	
		0.007*	−0.006*	0.024*	−0.004	0.043*	
		[0.004;0.010]	[−0.009;−0.003]	[0.015;0.033]	[−0.010;0.002]	[0.025;0.056]	
B–V weights	0.039	0.045	0.040	0.060	0.056	0.096	0.133
	0.094*	0.088*	0.093*	0.074*	0.077*	0.037*	
		0.006*	−0.005*	0.020*	−0.003	0.040*	
		[0.003;0.009]	[−0.008;−0.003]	[0.011;0.029]	[−0.009;0.002]	[0.023;0.053]	
Maximal item set							
Equal weights	0.042	0.044	0.041	0.057	0.054	0.090	0.111
	0.069*	0.067*	0.070*	0.054*	0.057*	0.021*	
		0.002	−0.003*	0.016*	−0.003	0.036*	
		[−0.000;0.004]	[−0.005;−0.002]	[0.009;0.023]	[−0.007;0.001]	[0.019;0.049]	
C–Z weights	0.024	0.026	0.024	0.033	0.031	0.055	0.066
	0.042*	0.040*	0.042*	0.033*	0.035*	0.011	
		0.001*	−0.002*	0.009*	−0.002	0.024*	
		[0.000;0.003]	[−0.003;−0.001]	[0.005;0.014]	[−0.004;0.001]	[0.012;0.031]	
B–V weights	0.021	0.021	0.020	0.027	0.026	0.045	0.056
	0.035*	0.035*	0.037*	0.029*	0.030*	0.011*	
		0.001	−0.002*	0.008*	−0.002	0.020*	
		[−0.001;0.002]	[−0.003;−0.001]	[0.004;0.011]	[−0.004;0.001]	[0.009;0.026]	
Intermediate item set							
Equal weights	0.059	0.061	0.056	0.080	0.075	0.124	0.174
	0.115*	0.112*	0.118*	0.094*	0.099*	0.050*	
		0.003	−0.006*	0.025*	−0.005	0.049*	
		[−0.001;0.007]	[−0.009;−0.003]	[0.014;0.036]	[−0.011;0.001]	[0.027;0.067]	
C–Z weights	0.041	0.044	0.039	0.058	0.054	0.095	0.131
	0.089*	0.087*	0.091*	0.073*	0.077*	0.035*	
		0.002*	−0.005*	0.019*	−0.004	0.041*	
		[0.000;0.005]	[−0.007;−0.003]	[0.011;0.027]	[−0.009;0.001]	[0.021;0.054]	
B–V weights	0.044	0.045	0.041	0.056	0.052	0.090	0.130
	0.085*	0.085*	0.089*	0.074*	0.078*	0.040*	
		0.001	−0.004*	0.015*	−0.004	0.038*	
		[−0.002;0.004]	[−0.007;−0.003]	[0.008;0.024]	[−0.009;0.001]	[0.020;0.052]	

Notes: For each index, the first row gives the index value, the second row gives the difference between the index for Portuguese and the (reweighted) natives, the third row shows the marginal reduction of the gap with the column component added and the fourth row in brackets gives 90% bootstrap variability bands for the latter. Stars indicate that bootstrap variability bands for the deprivation gap or marginal effects do not include zero.

deprivation gap, with impact ranging broadly between 40 and 60 percent of the raw gap. Portuguese immigrants are much more concentrated towards the lowest deciles of the household income distribution than natives (even after controlling for labour market participation and education) and this translates directly on the deprivation gap. While this result is not extremely surprising, the large magnitude of its effect is worth noticing. The deprivation gap is more than halved when the native population is fully re-weighted to the characteristics of Portuguese immigrants. Interestingly, while a statistically significant portion remains unexplained for eight of the ten indices, the gap fully disappears for the official 'EU set'.

In what follows, we further examine graphically the extent to which our five factors account for the raw gap. As discussed earlier, the latter allows us to provide a more comprehensive representation of the deprivation gap over the entire distribution. For the sake of brevity, we restrict our focus to two item sets: the 'EU set' and the 'intermediate' set with Betti–Verma weights.[12] We first consider the 'EU set' with equal weighting of items. This indicator is the official EU indicator of material deprivation. It does not include any of the housing quality indicators and focuses on items reflecting the strongest degree of deprivation. Accordingly, in a country like Luxembourg, the level of deprivation implied by this indicator is comparatively low, especially among natives. The lion's share of the contribution to this aggregate indicator is due to the 'ability to face unexpected expenses' and to a lesser extend the 'capacity of paying bills' (see Table 1).

Figure 1 displays (i) the inverse generalized Lorenz (IGL) curves from the raw samples (top left) and the difference between the two curves with bootstrap variability bands (bottom left) and (ii) counterfactual IGL curves from the reweighted native sample at the AHE (the conditional distribution of education), the AHEL (the conditional distribution of employment) and the AHELY (the conditional distribution of income) steps (top right) with the bottom right panel showing the remaining, 'unexplained' difference between the Portuguese IGL curve and the AHELY curve (also with bootstrap variability bands). The vertical bars indicate the proportion of the populations with non-zero deprivation (that is, deprived on at least one item) while the end-value of the

[12] Figures for the 'minimal' and 'maximal' sets are provided in Appendix E.

curves give mean deprivation levels.

Figure 2 shows the marginal reduction of the difference between IGL curves at each of three steps of the sequence of introduction of our control factors. The top left quadrant gives the reduction in the gap after controlling for age, household structure and education ($C_e^\theta = (\theta_0^{11100} - \theta_0^{11000})$); the top right quadrant gives the additional reduction of the gap observed after controlling for (conditional) employment differences ($C_l^\theta = (\theta_0^{11110} - \theta_0^{11100})$); the bottom left quadrant gives the partial effect of income differences ($C_y^\theta = (\theta_0^{11111} - \theta_0^{11110})$) netted out of differences in conditional education and employment; the bottom right quadrant gives the remaining, unexplained deprivation gap ($C_r^\theta(\theta_1^{11111} - \theta_0^{11111})$).

The raw deprivation gap is large and significant. More than 50 percent of Portuguese immigrants experience deprivation in (at least) one of the nine items compared to less than 20 percent of natives. The expected deprivation score is above 0.08 for Portuguese against just above 0.02 for natives. The most striking result, however, is that this gap is fully accounted for by the explanatory factors. The unexplained portion of the deprivation gap is never significantly different from zero. Income accounts for the largest part of the reduction in the deprivation gap. This should come as no surprise given the definition of items included in the EU indicator; most items refer to elements that "money can buy". This finding questions the value-added over common 'income-based' indicators of the official indicator of material deprivation in general and in a country like Luxembourg in particular.

Figures 3 and 4 show results for the aggregation based on the 'intermediate' set of items and the 'modified Betti-Verma' weighting scheme.[13]

The 'intermediate' set of items reflects our preferred choice of items in which we grouped some items and excluded items that most households possess making the information of enforced lack irrelevant (such as the possession of a TV set or phone or the capacity to eat meat/fish).

In the raw samples, the IGL curve for Portuguese lies everywhere above the curve for Luxembourgish. The deprivation gap is again unambiguous and significant. More than 70 per cent of Portuguese immigrants experience at least one of the deprivations

[13]See Table A1 for the composition of the index and item weights.

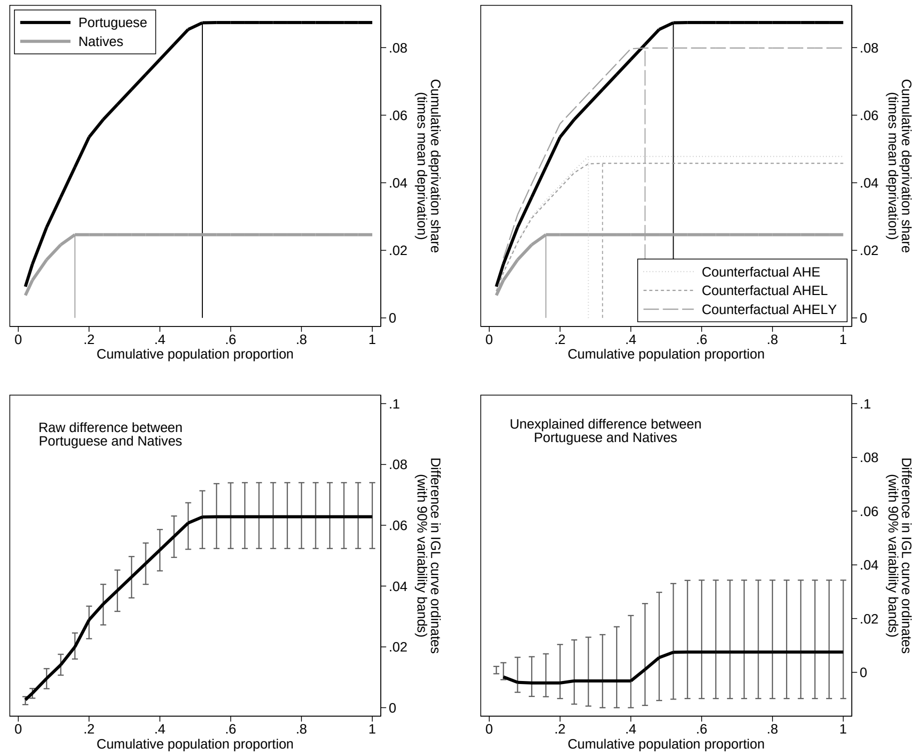

Figure 1. IGL curves (top) and curve differences (bottom) for the EU index of material deprivation

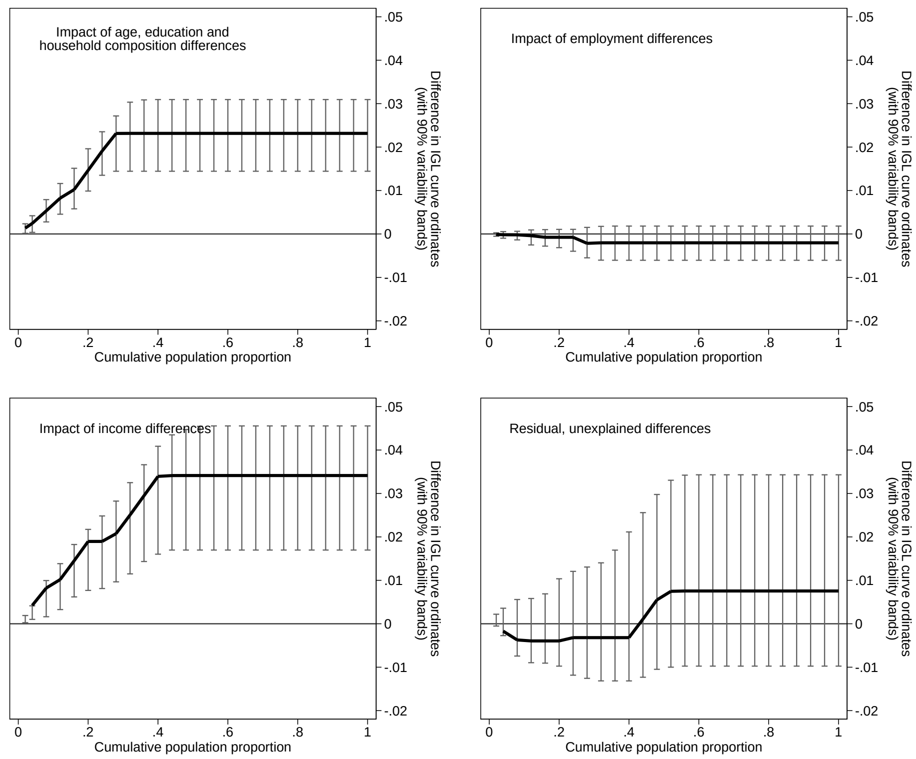

Figure 2. Marginal effects of components

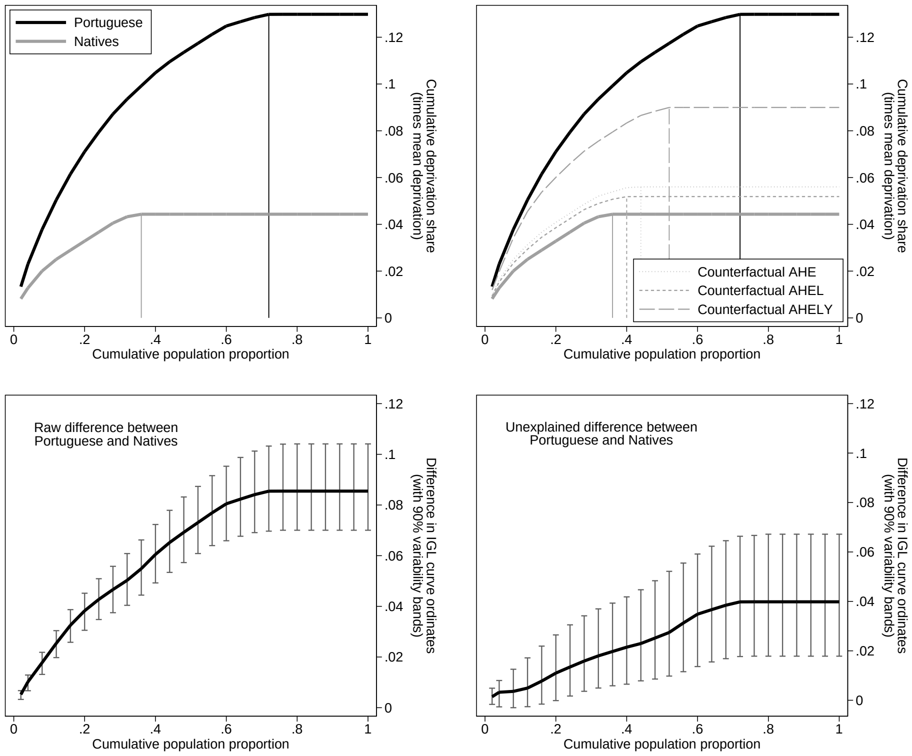

Figure 3. IGL curves (top) and curve differences (bottom) for the intermediate item set and Betti-Verma weighting scheme

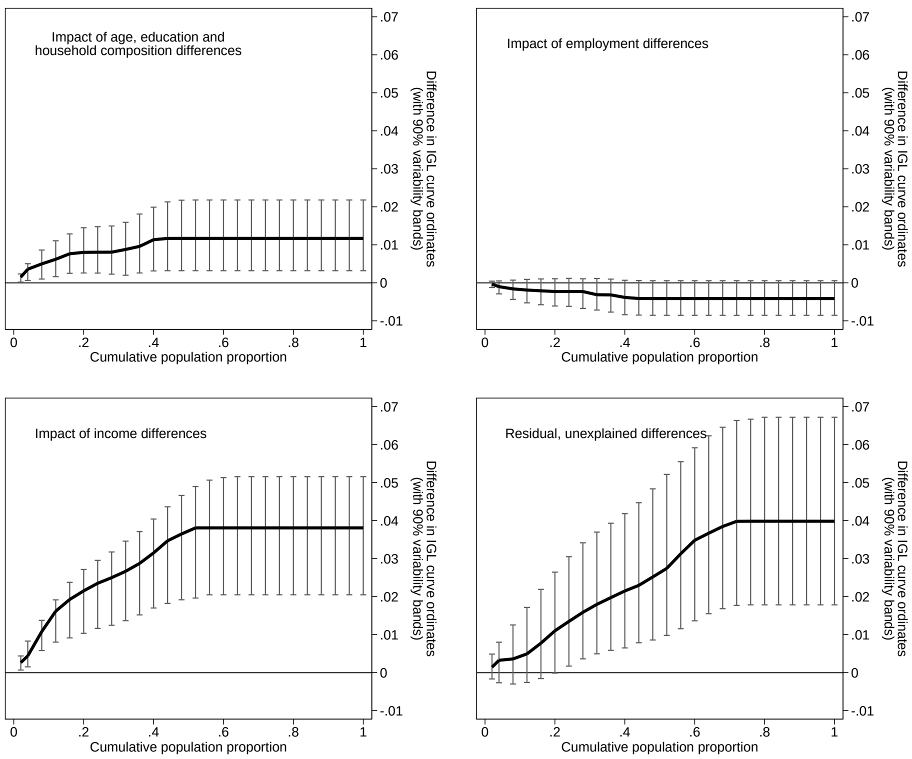

Figure 4. Marginal effects of components

considered, against less than 40 per cent of natives. The mean deprivation score is 0.13 for immigrants against 0.04 for natives (see Table 3).[14]

Just over half of the deprivation gap is accounted for by the five control factors (55%). Age, household structure, education and employment differences only account for just under 10% of the gap. Controlling for differences in income also leads to a marked reduction of the 'unexplained' part. Nevertheless, even after controlling for differences in income the deprivation gaps now remains positive and significant leaving a significant part unexplained, unlike what is observed with the 'EU set'. A similar pattern is found with the alternative 'minimal' and 'maximal' item sets (see Appendix E).

Before concluding, we look back at single items. The latter provide additional clues on the significance and relative importance of single items in our aggregated results. Table 4 shows the proportion of the population experiencing specific deprivations in the native and the Portuguese samples, as well as in the reweighted native sample. Numbers in brackets show the item-level deprivation gaps of Portuguese at each stage of the reweighting sequence. Stars indicate that the deprivation gap is statistically significant at 90 per cent confidence level. For instance, the first row repeats the results of Table 1 that 12% of Luxembourg natives and 45% of Portuguese lack the "*capacity to face unexpected expenses*" resulting in a large and statistically significant deprivation gap (at the 10% level of significance) of 33 percentage points. Columns in between these results show the expected deprivation at each stage of the five reweighting stages.

In line with our aggregated results, the size of the gap for each item is reduced when population characteristics of Luxembourg natives are substituted for those of Portuguese i.e. when the Luxembourg sample is reweighted to reflect Portuguese characteristics.

For all items within the economic strain dimension, before conditioning on income, differences in education explain most of the reduction in item-level gap. Once income differences are accounted for (AHELY), we no longer find sizable systematic differences between natives and Portuguese on these items. Differences in item means become small (except perhaps for the 'capacity to face unexpected expenses') and none remain

[14]Since the curves do not cross, the ranking would remain unchanged for generalized means of any positive power (Jenkins and Lambert, 1997, Aaberge and Peluso, 2011).

Table 4: Deprivation item means in raw and reweighted samples and item-level deprivation gaps of Portuguese immigrants

	LU						PT
	Raw	A	AH	AHE	AHEL	AHELY	Raw
1. Capacity to face unexpected expenses	0.12	0.15	0.13	0.25	0.24	0.36	0.45
	[0.33]∗	[0.30]∗	[0.32]∗	[0.20]∗	[0.21]∗	[0.09]	
2. Capacity to eat meat/fish	0.01	0.01	0.01	0.02	0.02	0.05	0.03
	[0.02]∗	[0.01]∗	[0.01]∗	[0.00]	[0.00]	[−0.02]	
3. Capacity paying a week holiday	0.06	0.08	0.07	0.11	0.11	0.19	0.21
	[0.14]∗	[0.13]∗	[0.14]∗	[0.09]∗	[0.10]∗	[0.02]	
4. Capacity to keep house warm	0.00	0.00	0.00	0.00	0.00	0.01	0.02
	[0.01]∗	[0.01]∗	[0.01]∗	[0.01]∗	[0.01]∗	[0.01]	
5. Difficulty paying bills	0.01	0.02	0.02	0.03	0.02	0.07	0.04
	[0.03]∗	[0.03]∗	[0.03]∗	[0.02]	[0.02]	[−0.03]	
6. Inability to pay rent/mortgage	0.01	0.01	0.01	0.02	0.02	0.05	0.04
	[0.03]∗	[0.03]∗	[0.03]∗	[0.02]∗	[0.02]∗	[−0.02]	
5.+6. Inability pay rent/mortgage or bills (combined)	0.02	0.02	0.02	0.03	0.03	0.08	0.06
	[0.04]∗	[0.04]∗	[0.04]∗	[0.03]∗	[0.03]∗	[−0.02]	
7. Possession of TV set	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	[0.00]∗	[0.00]∗	[0.00]∗	[0.00]∗	[0.00]∗	[0.00]	
8. Possession of phone	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	[0.00]	[0.00]∗	[0.00]∗	[0.00]∗	[0.00]∗	[0.00]	
9. Possession of dishwasher	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	[0.00]	[0.00]∗	[0.00]∗	[0.00]∗	[0.00]∗	[0.00]∗	
10. Possession of computer	0.01	0.01	0.01	0.01	0.01	0.02	0.09
	[0.08]∗	[0.08]∗	[0.09]∗	[0.08]∗	[0.08]∗	[0.07]∗	
11. Possession of car	0.01	0.01	0.00	0.01	0.01	0.02	0.02
	[0.02]∗	[0.01]∗	[0.02]∗	[0.01]∗	[0.01]∗	[−0.00]	
12. Leaky roof	0.04	0.04	0.05	0.04	0.04	0.06	0.05
	[0.01]	[0.01]	[0.01]	[0.01]	[0.01]	[−0.01]	
13. Rot in house	0.06	0.06	0.07	0.07	0.07	0.12	0.13
	[0.07]∗	[0.06]∗	[0.06]∗	[0.06]∗	[0.06]∗	[0.01]	
14. Damp in house	0.09	0.11	0.11	0.12	0.12	0.17	0.13
	[0.04]∗	[0.03]	[0.02]	[0.02]	[0.02]	[−0.03]	
12.+13.+14. Leaky roof/Damp/rot (combined)	0.12	0.14	0.14	0.14	0.14	0.20	0.18
	[0.06]∗	[0.05]∗	[0.05]∗	[0.05]	[0.05]	[−0.01]	
15. Double glazing	0.13	0.09	0.09	0.08	0.07	0.10	0.22
	[0.09]∗	[0.13]∗	[0.13]∗	[0.14]∗	[0.15]∗	[0.12]∗	
16. Hermetic doors/windows	0.10	0.10	0.10	0.13	0.13	0.20	0.14
	[0.04]∗	[0.04]∗	[0.04]∗	[0.01]	[0.01]	[−0.06]	
17. Outdoor space	0.05	0.06	0.04	0.07	0.07	0.09	0.30
	[0.25]∗	[0.25]∗	[0.27]∗	[0.23]∗	[0.24]∗	[0.21]∗	

Notes: Figures in square brackets are differences between raw item means for target sample (Portuguese) and item means for (reweighted) reference sample (Luxembourgish). Stars indicate that these differences are statistically significant at 90 per cent confidence levels. Combined items are equal to 1 if any of the deprivation items combined is equal to 1.

statistically significant. This finding is rather intuitive validating the critical importance of the association between income and financial distress.

We observe comparable results on items reflecting possession of common durable goods. Observed differences between the two groups disappear once income differences are controlled for, with the exception of the possession of a dishwasher and, most notably, of a computer. Interestingly, the gap of the latter is largely unexplained – our factors only explain one of the eight percentage point differences between the two groups. Other factors, perhaps cultural, appear to be at play on this particular item. Note that the gap in all items included in the official 'EU set' fully disappears once differences in income are accounted for.

Results for items reflecting housing conditions are more contrasted. While, raw differences on all items are to the disadvantage of Portuguese, the gap turns to their advantage on four of the seven items including having a leaky roof, damp and rot in the house or lacking hermetic doors and windows. This finding suggests that the Portuguese appear to have better housing conditions than natives with similar characteristics. This could be potentially explained by the large contingent of Portuguese workers in the construction sector. Differences are not significantly different from zero however.

One notable exception is the 'availability of outdoor space' which remains significantly to the disadvantage of Portuguese. This can allegedly be related to the time-invariant character of the presence of outdoor space in an accommodation, as this cannot be 'fixed' by repair or transformation work but is tied to the initial investment. Despite its high housing costs, Luxembourg is considered a successful model of residential integration as foreign-born are not more likely to reside in subsidized public housing than natives (Fetzer, 2010). However, while Portuguese are much less likely to own their principal residence than natives (Berger, 2004), this housing gap is reduced by half once including their housing owned overseas[15] – about 20% of Portuguese households who do not own their residence in Luxembourg own one abroad, most likely in their country of origin. Given the high private ownership of land in Luxembourg (Cahill and McMahon, 2010), it is likely that in-vivo transfers and inheritance ease Luxembourg natives' access to

[15] See Berger (2004) for home ownership statistics in Luxembourg.

housing including properties with an outdoor space (which is likely to proxy some of the aforementioned factors). Overall, these considerations are likely to loosen the link to education and income. Material deprivation indicators that include housing indicators are therefore less closely determined by income.

6. Concluding remarks

The assimilation of immigrants to their host destination has been the object of numerous studies on income, earnings or employment differentials. Consistent with the growing recognition of the multidimensionality of well-being, recent studies have also started to treat the socio-economic assimilation of immigrants as an inherently multi-dimensional process (Haisken-DeNew and Sinning, 2010, Aleksynska and Algan, 2010).

This study presented a methodological framework which incorporates these recent developments to examine the degree of material deprivation of Portuguese migrants in several non-monetary dimensions. In this context, the degree of assimilation is measured by the distance between native- and foreign-born in terms of material deprivation. We find that material deprivation among Portuguese is non-negligible and the nativity gap is large. This finding corroborates Haisken-DeNew and Sinning (2010) who show that immigrants in Germany are more severely deprived than natives. It is also consistent with the few studies reporting that Portuguese in Luxembourg lag behind Luxembourg natives in income, employment and educational attainments (see, Langers, 2006, Hartmann-Hirsch, 2007, Alieva, 2010, Van Kerm and Villeret, 2007). To the extent that the poor performance of Portuguese immigrants does not appear to be limited to Luxembourg, as suggested by Kalter and Granato (2002) or Brinbaum and Kieffer (2009), our application provides an assessment which could potentially benefit policy makers in larger, neighbouring countries which also share an important population of Portuguese residents.

We have assessed the robustness of our findings to the choice of weighting schemes and dimensions used in defining material well-being. With the exception of our so-called 'EU set' which exclusively focuses on the most basic deprivation items following the practice of the European Commission, we find compelling evidence that material deprivation is not entirely accounted for by income differentials (conditional on population composition), in particular when housing conditions are taken into account. This observation gives further support to the use of a multidimensional approach encompassing income to examine well-being. At the same time, it cast serious doubt on the value added by the official EU indicator –in its current format– over the simple use of an income based indicator in a country like Luxembourg.

Findings from our application suggest that the extent to which income differences account for the deprivation gap is more sensitive to the choice of individual items used to define deprivation than to the weighting scheme considered. This may not be fully surprising as we imposed a common weighting of items for both natives and Portuguese immigrants and implicitly assumed common 'preferences' over deprivation items in the two groups. However, our results could be affected if native- and foreign-born households assess their well-being differently, with respect, e.g., to idiosyncratic reference groups (Haisken-DeNew and Sinning, 2010, Dickes, Fusco, and Marlier, 2010, Bellani and D'Ambrosio, 2011). Adjusting the structure of the deprivation index to subgroup deprivation levels in subgroup comparisons within a country is however normatively debatable.

We note finally that the methods developed in the paper have potential application more generally. There is a large consensus supporting the view that deprivation (or poverty) is an inherently multidimensional concept which encompasses a broad spectrum of dimensions, implying that the sole use of income-based indicators would likely fail to capture important elements of human well-being; see, e.g., Sen (1992) or Bourguignon and Chakravarty (2003). As a result, multidimensional and non-monetary deprivation indicators are gaining popularity in Europe and elsewhere to better understand and identify poverty, the feeling of poverty and social exclusion (see OECD (2011) or Nolan and Whelan (2010) for a recent discussion). Our methodological framework could be adapted to compare such multidimensional outcomes across population groups –substituting the material deprivation indicators we focus on by broader, multidimensional socio-economic outcome indicators. In particular, given evidence of variations in the degree of assimilation according to alternative dimensions (Aleksynska and Algan, 2010), integrated, multidimensional analysis of immigrants assimilation along these lines is an avenue for future research.

References

Aaberge, R., and E. Peluso (2011): "A Counting Approach for Measuring Multidimensional Deprivation," Working Papers 07/2011, Universita di Verona, Dipartimento di Scienze economiche.

Adsera, A., and B. Chiswick (2007): "Are there gender and country of origin differences in immigrant labor market outcomes across European destinations?," *Journal of Population Economics*, 20(3), 495–526.

Aleksynska, M., and Y. Algan (2010): "Assimilation and Integration of Immigrants in Europe," IZA Discussion Paper 5185, Institute for the Study of Labor, Bonn.

Alieva, A. (2010): *Educational Inequalities in Europe: Performance of Students with Migratory Background in Luxembourg and Switzerland*. International Academic Publishers.

Alkire, S., and J. Foster (2011): "Counting and multidimensional poverty measurement," *Journal of Public Economics*, 95(7-8), 476–487.

Atkinson, A. B. (2003): "Multidimensional Deprivation: Contrasting Social Welfare and Counting Approaches," *Journal of Economic Inequality*, 1, 51–65.

Atkinson, A. B., B. Cantillon, E. Marlier, and B. Nolan (2002): *Social Indicators: The EU and Social Inclusion*. Oxford University Press.

Barsky, R., J. Bound, K. K. Charles, and J. P. Lupton (2002): "Accounting for the Black-White Wealth Gap: A Nonparametric Approach," *Journal of the American Statistical Association*, 97, 663–673.

Bauer, T. K., D. A. Cobb-Clark, V. A. Hildebrand, and M. G. Sinning (2011): "A Comparative Analysis of the Nativity Wealth Gap," *Economic Inquiry*, 49(4), 989–1007.

Bellani, L., and C. D'Ambrosio (2011): "Deprivation, Social Exclusion and Subjective Well-Being," *Social Indicators Research*, 104(1), 67–86.

Berger, F. (2004): "Le Patrimoine Immobilier des Ménages: Qui Possède Quoi ?," Population & Emploi 4, CEPS/INSTEAD, Differdange, Luxembourg.

Berger, F. (2008): "Zoom sur les Primo-Arrivants Portuguais et leurs Descendants," Vivre au Luxembourg 49, CEPS/INSTEAD, Differdange, Luxembourg.

Betti, G., B. Cheli, A. Lemmi, and V. K. Verma (2008): "The Fuzzy Set Approach to the Multidimensional Poverty: The Case of Italy in the 1990s," in *Quantitative Approaches to Multidimensional Poverty Measurement*, ed. by N. Kakwani, and J. Silber, pp. 30–48. Palgrave Mac Millan, New York.

Betti, G., and V. K. Verma (1998): "Measuring the degree of poverty in a dynamic and comparative context: A multi-dimensional approach using fuzzy set theory," Working Paper 22, Dipartimento di Metodi Quantitativi, Universita di Siena.

Blanchflower, D. G., and A. Oswald (2004): "Well-being over time in Britain and the USA," *Journal of Public Economics*, 88, 1359–1386.

Borjas, G. J. (1985): "Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants," *Journal of Labor Economics*, 3(4), 463–89.

Borjas, G. J. (1995): "Assimilation and Changes in Cohort Quality Revisited: What Happened to Immigration Earnings in the 1980s?," *Journal of Labor Economics*, 13(2), 201–245.

Bosswick, W., and F. Heckmann (2006): "Integration of Migrants: Contribution of Local and Regional Authorities," European Foundation for the Improvement of Living and Working Conditions, Dublin.

Bourguignon, F., and S. R. Chakravarty (2003): "The Measurement of Multidimensional Poverty," *Journal of Economic Inequality*, 1, 25–49.

Brinbaum, Y., and A. Kieffer (2009): "Trajectories of Immigrants' Children in Secondary Education in France: Differentiation and Polarization," *Population-E*, 64, 507–554.

Cahill, K., and R. McMahon (2010): *Who Owns the World: The Surprising Truth About Every Piece of Land on the Planet*. Grand Central Publishing.

Cameron, A. C., and P. K. Trivedi (2005): *Microeconometrics: Methods and Applications*. Cambridge University Press, Cambridge.

Cantillon, B., H. Verschueren, and P. Ploscar (2012): *Social Inclusion and Social Protection in the EU: Interactions between Law and Policy*. Intersentia.

Carroll, C. D., B.-K. Rhee, and C. Rhee (1994): "Are There Cultural Effects on Saving? Some Cross-Sectional Evidence," *Quarterly Journal of Economics*, 109(3), 685–99.

Cerioli, A., and S. Zani (1990): "A fuzzy approach to the measurement of poverty," in *Income and wealth distribution, inequality and poverty*, ed. by C. Dagum, and M. Zenga, pp. 272–84. Springer Verlag, Berlin.

Chiswick, B. (2002): "Longitudinal Analysis of Immigrant Occupational Mobility: A Test of the Immigrant Assimilation Hypothesis," IZA Discussion Paper 452, Institute for the Study of Labor, Bonn.

Chiswick, B., and P. W. Miller (2009): "Citizenship in the United States: The Roles of Immigrant Characteristics and the Country of Origin," *Research in Labor Economics*, 29, 91–130.

Chiswick, B. R. (1978): "The Effect of Americanization on the Earnings of Foreign-born Men," *Journal of Political Economy*, 86(5), 897–921.

Clark, K., and J. Lindley (2009): "Immigrant Assimilation Pre and Post Labour Market Entry: Evidence from the UK Labour Force Survey," *Journal of Population Economics*, 22, 175–198.

Cobb-Clark, D., and V. Hildebrand (2006a): "The Wealth and Asset Holdings of U.S.-Born and Foreign Born Households: Evidence from SIPP Data," *Review of Income and Wealth*, 52(1), 17–42.

——— (2006b): "The Wealth of Mexican Americans," *Journal of Human Resources*, 41(4), 841–868.

Decancq, K., and M. Lugo (2012): "Weights in Multidimensional Indices of Well-Being: An Overview," Forthcoming in *Econometrics Reviews*.

Deutsch, J., and J. Silber (2005): "Measuring Multidimensional Poverty: An Empirical Comparison of Various Approaches," *Review of Income and Wealth*, 51(1), 145–174.

Devicienti, F. (2010): "Shapley-Value Decompositions of Changes in Wage Distributions: A Note," *Journal of Economic Inequality*, 8, 35–45.

Dickes, P., A. Fusco, and E. Marlier (2010): "Structure of national perceptions of social needs across EU countries," *Social Indicators Research*, 95, 143–67.

DiNardo, J., N. M. Fortin, and T. Lemieux (1996): "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, 64(5), 1001–1044.

Duclos, J.-Y., D. E. Sahn, and S. D. Younger (2006): "Robust Multidimensional Poverty Comparisons," *Economic Journal*, 116, 943–968.

Dustmann, C. (1996): "The social assimilation of immigrants," *Journnal of Population Economics*, 54, 9–37.

European Commission (2009): "Portfolio of indicators for the monitoring of the European strategy for social protection and social inclusion – 2009 update," Directorate for Employment, Social Affairs and Equal Opportunities, European Commission, Brussels.

——— (2011): "Migrants in Europe – A statistical portrait of the first and second generation," Statistical Book, Eurostat, The Statistical Office of the European Union, Luxembourg.

Fetzer, J. S. (2010): "Housing and the Residential Integration of Immigrants in Luxembourg and the EU-15," Working Paper 547, Luxembourg Income Study.

Frenette, M., V. Hildebrand, J. McDonald, and C. Worswick (2003): "Job and Occupational Mobility of Immigrants in Australia, Canada and the United States," in *Canadian Immigration Policy for the 21st Century*. John Deutsch Institute for the Study of Economic Policy.

Fusco, A. (2007): *La Pauvreté - Un Concept Multidimensionnel*. L'Harmattan, Paris.

Green, D. (1999): "Immigrant Occupational Attaintment: Assimilation and Mobility over Time," *Journal of Labor Economics*, 17(1), 49–79.

Guio, A.-C., A. Fusco, and E. Marlier (2009): "A European Union Approach to Material Deprivation using EU-SILC and Eurobarometer data," IRISS Working Paper 2009-19, CEPS/INSTEAD, Differdange, Luxembourg.

Haisken-DeNew, J. P., and M. Sinning (2010): "Social Deprivation and Exclusion of Immigrants in Germany," *Review of Income and Wealth*, 56(4), 715–733.

Hartmann-Hirsch, C. (2007): "Les etrangers et le marche de l'emploi: politiques migratoires et immigrations," Cahier PSELL 157, CEPS/INSTEAD, Differdange, Luxembourg.

Izquierdoa, M., A. Lacuestaa, and R. Vegas (2009): "Assimilation of immigrants in Spain: A Longitudinal Analysis," *Labour Economics*, 16, 669–678.

Jenkins, S. P., and P. J. Lambert (1997): "Three I's of Poverty Curves, With An Analysis of UK Poverty Trends," *Oxford Economic Papers*, 49(?), 317–327.

Kalter, F., and N. Granato (2002): "Demographic Change, Educational Expansion, and Structural Assimilation of Immigrants: The Case of Germany," *European Sociological Review*, 18(2), 199–216.

Klasen, S. (2000): "Measuring Poverty and Deprivation in South Africa," *Review of Income and Wealth*, 46(1), 33–58.

Langers, J. (2006): "Nationaux et étrangers résident au Luxembourg: Analyse statistique de quelques différences socio-démographiques," Economie et Statistique 2, Statec, Luxembourg.

Layte, R., C. T. Whelan, B. Maître, and B. Nolan (2001): "Explaining Levels of Deprivation in the European Union," *Acta Sociologica*, 44(2), 105–121.

Lechner, M. (2008): "A Note on the Common Support Problem in Applied Evaluation Studies," *Annales d'Economie et de Statistique*, (91-92), 217–235.

Mathä, T. Y., A. Porpiglia, and E. M. Sierminska (2011): "The Immigrant/Native Wealth Gap in Germany, Italy and Luxembourg," BCL working paper 57, Central Bank of Luxembourg, Luxembourg.

McKay, S. (2004): "Poverty of Preference: What do Consensual Deprivation Indicators really Measure," *Fiscal Studies*, 25, 201–224.

Nolan, B., and C. T. Whelan (1996): *Resources, Deprivation and Poverty.* Clarendon Press.

——— (2010): "Using Non-Monetary Deprivation Indicators to Analyze Poverty," *Journal of Policy Analysis and Management*, 29, 305–325.

——— (2011): "The EU 2020 Poverty Target," GINI Discussion Paper 19, AIAS, Amsterdam.

OECD (2011): *How's Life? Measuring Well-Being.* OECD Publishing.

Pi Alperin, M. N. (2010): "Comparing non-monetary deprivation and inequality levels in the EU countries," CEPS/INSTEAD Working Paper 2010-27, CEPS/INSTEAD, Differdange, Luxembourg.

Saigo, H., J. Shao, and R. R. Sitter (2001): "A Repeated Half-Sample Bootstrap and Balanced Repeated Replications for Randomly Imputed Data," *Survey Methodology*, 27(2), 189–196.

Sen, A. K. (1992): *Inequality re-examined. Cambridge, MA: Harvard University Press.* Cambridge, MA: Harvard University Press.

Sinning, M. (2009): "Home Ownership and the Economic Performance of Immigrants in Germany," *Urban Studies*, 5, 1–23.

Spencer, B. D., and S. Fisher (1992): "On comparing distributions of poverty gaps," *Sankhya: The Indian Journal of Statistics, Series B*, 54, 114–126.

Townsend, P. (1979): *Poverty in the United Kingdom.* Penguin, Harmondsworth.

Van Kerm, P., and A. Villeret (2007): "Difficile de joindre les deux bouts? La satisfaction des ménages luxembourgeois quant à leur situation financière," Vivre au Luxembourg 36, CEPS/INSTEAD, Differdange, Luxembourg.

Whelan, C. T., B. Nolan, and B. Maître (2008): "Measuring Material Deprivation in the Enlarged EU," ESRI Working Paper 249, Economic and Social Research Institute, Dublin, Ireland.

Appendix A Construction of reweighting factors

Reweighting factors are constructed as follows. Consider the counterfactual distribution:

$$p_0^{11111}(\mathbf{d}) = \sum_{s\in\Omega_Z} \Pr[\Theta = \mathbf{d} \mid Z = s,\ I = 0] \times \Pr[Z = s \mid I = 1],$$

where Z is the full set of all five covariates for notational clarity. This can be written as a weighted version of the observed p_0:

$$p_0^{11111}(\mathbf{d}) = \sum_{s\in\Omega_Z} \Pr[\Theta = \mathbf{d} \mid Z = s,\ I = 0] \times \psi^{11111}(s) \times \Pr[Z = s \mid I = 0],$$

where

$$\psi^{11111}(s) = \frac{\Pr[Z = s \mid I = 1]}{\Pr[Z = s \mid I = 0]},$$

is a reweighting factor.

As observed by DiNardo, Fortin, and Lemieux (1996), invoking Bayes' rule, $\psi^{11111}(s)$ can be equivalently expressed as:

$$\psi^{11111}(s) = \frac{\Pr[I = 1 \mid Z = s]}{\Pr[I = 0 \mid Z = s]}\, \frac{\Pr[I = 0]}{\Pr[I = 1]}.$$

Both formulations are equivalent, but the latter involves only univariate probabilities which can be easily modeled using standard binary choice models. This is particularly useful when the dimension of Z is large. In our application however, given the categorical nature of our covariates, we estimate the reweighting factor using estimates of probabilities directly by 'cell means', e.g.,

$$\widehat{\Pr[I = j \mid Z = s]} = \frac{1}{\sum_{i=1}^{N} \mathbf{1}\{z_i = s\}} \sum_{i=1}^{N} \mathbf{1}\{I_i = j\}\, \mathbf{1}\{z_i = s\},$$

where I_i is an immigrant indicator and z_i is a vector of covariates for individual i ($\mathbf{1}\{\text{cond}\}$ evaluates to 1 if cond is true and 0 otherwise). This avoids any parametric assumption about the distribution of covariates in each of the two groups, at the cost of inflated sampling variability.[16]

Other generic counterfactual distributions adjustments for only a subset of covariates

[16]Experiments with a naive probit model (with no interactions between covariates) to predict probabilities yield similar estimates.

are constructed using similar reweighting procedures. The counterfactual distribution corresponding to assigning a subset of covariates from our comparison group (Portuguese migrants) to our reference group (natives from Luxembourg) can be expressed by reweighting the probability distribution for natives as follows:

$$\begin{aligned}
p_0^{ahely}(\mathbf{d}) = \sum_{s\in\Omega_A}\sum_{t\in\Omega_H}\sum_{u\in\Omega_E}\sum_{v\in\Omega_L}\sum_{w\in\Omega_Y} \psi_0^{ahely}(s,t,u,v,w) \\
\Pr[\Theta = \mathbf{d} \mid A = s, H = t, E = u; L = v, Y = w, \quad I = 0] \\
\Pr[A = s \mid \quad I = 0] \\
\Pr[H = t \mid A = s, \quad I = 0] \\
\Pr[E = u \mid A = s, H = t, \quad I = 0] \\
\Pr[L = v \mid A = s, H = t, E = u, \quad I = 0] \\
\Pr[Y = w \mid A = s, H = t, E = u, L = v, \quad I = 0],
\end{aligned}$$

where the reweighting function is now a product of ratios of univariate probabilities which can also be estimated by cell-means computations or from a parametric model:

$$\begin{aligned}
\psi_0^{ahely}(s,t,u,v,w) = \left(\frac{\Pr[I=1 \mid A=s]}{\Pr[I=0 \mid A=s]}\ \frac{\Pr[I=0]}{\Pr[I=1]}\right)^{\mathbf{1}\{a=1\}} \\
\left(\frac{\Pr[I=1 \mid A=s, H=t]}{\Pr[I=0 \mid A=s, H=t]}\ \frac{\Pr[I=0 \mid A=s]}{\Pr[I=1 \mid A=s]}\right)^{\mathbf{1}\{h=1\}} \\
\left(\frac{\Pr[I=1 \mid A=s, H=t, E=u]}{\Pr[I=0 \mid A=s, H=t, E=u]}\ \frac{\Pr[I=0 \mid A=s, H=t]}{\Pr[I=1 \mid A=s, H=t]}\right)^{\mathbf{1}\{e=1\}} \\
\left(\frac{\Pr[I=1 \mid A=s, H=t, E=u, L=v]}{\Pr[I=0 \mid A=s, H=t, E=u, L=v]}\ \frac{\Pr[I=0 \mid A=s, H=t, E=u]}{\Pr[I=1 \mid A=s, H=t, E=u]}\right)^{\mathbf{1}\{l=1\}} \\
\left(\frac{\Pr[I=1 \mid A=s, H=t, E=u, L=v, Y=w]}{\Pr[I=0 \mid A=s, H=t, E=u, L=v, Y=w]}\ \frac{\Pr[I=0 \mid A=s, H=t, E=u, L=v]}{\Pr[I=1 \mid A=s, H=t, E=u, L=v]}\right)^{\mathbf{1}\{y=1\}} .
\end{aligned}$$

Appendix B Bootstrap inference

Sampling variability of all estimates is estimated from bootstrap resampling. To deal with the dependence of sample observations, we implement a non-parametric block bootstrap procedure, as described by Cameron and Trivedi (e.g., 2005, Chapter 11). Resampling is done independently from households interviewed at wave 1 of the survey within each of the sampling strata. We also resample households selected into the survey at subsequent waves from the 'new immigrants' samples added at each wave. Sampling 'at wave 1' for these households corresponds to sampling at the wave of their selection into the 'immigrants' sample. Each 'new immigrants' sample has its own set of strata from which resampling is done.

To maintain dependence of observations belonging to a common household and dependence over time in the pooled sample, the full response history for waves 3–5 of all members of the selected households (plus members of associated split-off households and all respondents that later joined these households) is then selected to generate a bootstrap replicate of our original, working sample. To take into account complex survey design features (potentially small stratum sizes specifically), resampling is based on the repeated half-sample bootstrap algorithm of Saigo, Shao, and Sitter (2001).

Denote by $\mathbf{X}^0$ our $(n^0 \times q)$ working data matrix. n^0 is the total number of observations on q variables –which include a nationality indicator, individual and household characteristics along with a set of K deprivation indicators (and an individual sample weight)– in the pooled 2004–2006 sample. Denote by $\mathbf{X}^b$ one bootstrap replication b of $\mathbf{X}^0$ constructed as described above. Define also $\mathrm{PT}(\mathbf{X}^b)$ and $\mathrm{LU}(\mathbf{X}^b)$ (for $b \in \{0, 1, \ldots, B\}$) as the subsets of $\mathbf{X}^b$ containing only observations with, respectively, Portuguese and Luxembourgish nationality.

All statistics of interest in this paper (including coordinates of the IGL curves at any p) are estimated on $\mathbf{X}^0$ and on each of 500 replicate subsamples $\mathbf{X}^b$. Denote generally any such estimate as $\hat{\theta}^b_{\mathrm{PT}} \equiv \theta(\mathrm{PT}(\mathbf{X}^b), \mathbf{X}^b)$ (if the statistic is estimated for Portuguese immigrants) and as $\hat{\theta}^b_{\mathrm{LU}} \equiv \theta(\mathrm{LU}(\mathbf{X}^b), \mathbf{X}^b)$ otherwise. The two arguments of θ in this generic notation are meant to reflect that some statistics of interest are functions of data from both the subsample of interest and the full sample; e.g., for computing the Cerioli-Zani or Betti-Verma weights, or for calculating reweighting factors as described in Section 3. Note that both arguments use data from the bootstrap replication b. This emphasizes the point that we fully incorporate sampling variability associated with the estimation of item weights or of reweighting factors. All comparisons of Portuguese immigrants and natives are calculated as: $\hat{\Delta}^b \equiv \theta(\mathrm{PT}(\mathbf{X}^b), \mathbf{X}^b) - \theta(\mathrm{LU}(\mathbf{X}^b), \mathbf{X}^b)$.

The B replicates of $\hat{\theta}_j^b$ (for $j \in \{\mathrm{PT}, \mathrm{LU}\}$) are used to estimate the standard error of our point estimate $\hat{\theta}_j^0$ as:

$$\hat{s}(\hat{\theta}_j^0) = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}(\hat{\theta}_j^b - \bar{\theta}_j)^2}, \tag{3}$$

where $\bar{\theta}_j$ is the average of $\hat{\theta}_j^b$ over the B replications. When we report sampling variability bands, these are derived from the largest and smallest vintiles of the B replications for $\hat{\theta}_j^b$. Standard errors and variability bands for $\hat{\Delta}^0$ are derived analogously from the B replicates of $\hat{\Delta}^b$. Variability bands for $\hat{\Delta}^0$ that do not include zero are interpreted as evidence that the observed difference between natives and Portuguese immigrants is significantly different from zero (at a 90 per cent confidence level).

Appendix C Item sets and weighting schemes

Table A1 reports individual item weights for each of the four item sets and three weighting schemes described in the main text.

Table A1: Item weights for alternative sets and weighting schemes

Item	EU set	Minimal set			Maximal set			Intermediate set		
	Eq.	Eq.	C–Z	B–V	Eq.	C–Z	B–V	Eq.	C–Z	B–V
1. Capacity to face unexpected expenses	0.111	0.143	0.082 (0.002)	0.056 (0.005)	0.059	0.025 (0.001)	0.018 (0.003)	0.111	0.063 (0.002)	0.045 (0.005)
2. Capacity to eat meat/fish	0.111	-	-	-	0.059	0.063 (0.002)	0.056 (0.009)	-	-	-
3. Capacity paying a week holiday	0.111	0.143	0.115 (0.003)	0.079 (0.007)	0.059	0.035 (0.001)	0.024 (0.004)	0.111	0.087 (0.002)	0.062 (0.008)
4. Capacity to keep house warm	0.111	-	-	-	0.059	0.078 (0.003)	0.088 (0.014)	-	-	-
5. Difficulty paying bills	-	-	-	-	0.059	0.058 (0.002)	0.053 (0.009)	0.111	0.146 (0.004)	0.130 (0.011)
6. Inability to pay rent/mortgage	-	-	-	-	0.059	0.064 (0.002)	0.056 (0.010)	0.111	0.162 (0.004)	0.134 (0.011)
5.+6. Inability pay rent/mortgage or bills (combined)	0.111	0.143	0.183 (0.005)	0.203 (0.011)	-	-	-	-	-	-
7. Possession of TV set	0.111	-	-	-	0.059	0.129 (0.009)	0.163 (0.041)	-	-	-
8. Possession of phone	0.111	-	-	-	0.059	0.116 (0.009)	0.156 (0.036)	-	-	-
9. Possession of dishwasher	0.111	-	-	-	0.059	0.098 (0.006)	0.093 (0.014)	-	-	-
10. Possession of computer	-	0.143	0.179 (0.006)	0.179 (0.010)	0.059	0.054 (0.002)	0.048 (0.006)	0.111	0.137 (0.004)	0.144 (0.009)
11. Possession of car	0.111	0.143	0.232 (0.007)	0.246 (0.017)	0.059	0.070 (0.003)	0.061 (0.009)	0.111	0.176 (0.006)	0.197 (0.017)
12. Leaky roof	-	-	-	-	0.059	0.046 (0.002)	0.046 (0.006)	-	-	-
13. Rot in house	-	-	-	-	0.059	0.039 (0.001)	0.024 (0.005)	-	-	-
14. Damp in house	-	-	-	-	0.059	0.033 (0.001)	0.021 (0.004)	-	-	-
12.+13.+14. Leaky roof/Damp/rot (combined)	-	0.143	0.097 (0.003)	0.121 (0.009)	-	-	-	0.111	0.074 (0.002)	0.097 (0.008)
15. Double glazing	-	-	-	-	0.059	0.028 (0.001)	0.031 (0.005)	0.111	0.070 (0.002)	0.096 (0.008)
16. Hermetic doors/windows	-	-	-	-	0.059	0.032 (0.001)	0.031 (0.004)	-	-	-
17. Outdoor space	-	0.143	0.112 (0.003)	0.115 (0.008)	0.059	0.034 (0.001)	0.033 (0.005)	0.111	0.085 (0.003)	0.096 (0.008)

Note: Eq. refers to the equal weight scheme. C–Z refers to the Cerioli-Zani approach and B–V refers to the Betti-Verma. Figures in brackets are bootstrapped standard errors of item weights (the equal weight scheme has no sampling variability).

Appendix D Reweighted sample comparisons

The performance of the reweighting procedure in balancing the samples can be assessed from Table A2. The table shows the distribution of covariates in the natives and Portuguese samples and in the reweighted natives sample at each stage of the reweighting sequence. The reweighted natives sample closely resembles the target Portuguese sample on all five covariates. Although this can not be read from the table, it is worth mentioning that the samples are also close in terms of the joint distributions of covariates too.

Sample size statistics reported for the reweighted samples exclude observations that are reweighted down to zero. Such observations correspond to native households that have characteristics which are not observed in the Portuguese sample, that is observations that are not on a common support with the Portuguese sample. So the effective sample size of Luxembourg natives is reduced from 5,020 households to 2,473 households with characteristics on all five factors that are observed in the Portuguese sample.

Table A2: Characteristics of reweighted samples

	LU						PT
	Raw	A	AH	AHE	AHEL	AHELY	Raw
Age of head of household (a)	*100*	*100*	*100*	*100*	*100*	*100*	*100*
Below 25	0.6	2.3	2.1	1.8	1.7	1.1	2.3
25–49	48.1	78.0	78.1	78.4	78.4	81.2	78.0
50–64	29.2	16.9	16.9	16.9	17.0	15.2	16.9
65+	22.2	2.9	2.9	2.9	2.9	2.6	2.9
Aggregate diff. to target sample	0.477	0.000	0.003	0.008	0.009	0.048	0.000
Household structure (h)	*100*	*100*	*100*	*100*	*100*	*100*	*100*
Pensioners households (1 or 2 person aged 65+, no kids)	18.4	2.6	1.8	1.8	1.8	2.1	1.8
Single adults (with or without kids)	11.8	15.8	5.9	5.9	5.9	6.7	5.9
Couples (no kids)	22.2	18.1	18.9	19.0	19.0	16.4	18.9
Families (couples with kids)	47.6	63.4	73.3	73.2	73.3	74.9	73.4
Aggregate diff. to target sample	0.395	0.166	0.001	0.002	0.001	0.040	0.000
Education level of head of household (e)	*100*	*100*	*100*	*100*	*100*	*100*	*100*
Lower secondary or below	32.1	24.7	24.9	80.8	80.8	78.7	80.6
Upper secondary	43.7	47.6	47.4	17.7	17.7	19.7	17.9
Higher education	24.2	27.7	27.6	1.5	1.5	1.6	1.5
Aggregate diff. to target sample	0.711	0.819	0.816	0.002	0.002	0.027	0.000
Employment intensity in household (l)	*100*	*100*	*100*	*100*	*100*	*100*	*100*
No active adult	26.4	8.2	6.5	8.7	5.2	5.7	5.2
One active adult	35.6	43.9	40.6	36.7	31.2	33.3	31.4
Two or more active adults	38.0	47.8	52.9	54.6	63.5	61.1	63.4
Aggregate diff. to target sample	0.389	0.220	0.147	0.128	0.002	0.033	0.000
Income decile group (y)	*100*	*100*	*100*	*100*	*100*	*100*	*100*
1st decile group	5.9	6.4	5.3	8.3	7.3	28.0	27.4
2nd decile group	6.4	6.7	6.5	8.6	7.8	21.7	24.8
3rd decile group	8.2	8.1	8.4	10.2	9.2	16.5	15.9
4th or 5th decile group	20.6	20.3	20.9	25.9	25.0	20.2	20.2
6th or 7th decile group	23.1	21.8	22.3	20.9	21.9	8.9	7.7
8th, 9th or 10th decile group	35.8	36.6	36.6	26.2	28.7	4.7	4.0
Aggregate diff. to target sample	0.836	0.823	0.840	0.712	0.762	0.053	0.000
Aggregate diff. to target sample	2.809	2.028	1.807	0.853	0.776	0.201	0.000
Sample size (# household-years)	5020	5020	4998	4113	3594	2473	1321

Notes: Stars indicate that differences between the (reweighted) natives sample and the Portuguese sample are statistically significant at 90 per cent confidence level based on bootstrap variability bands (see text).

Appendix E Additional figures

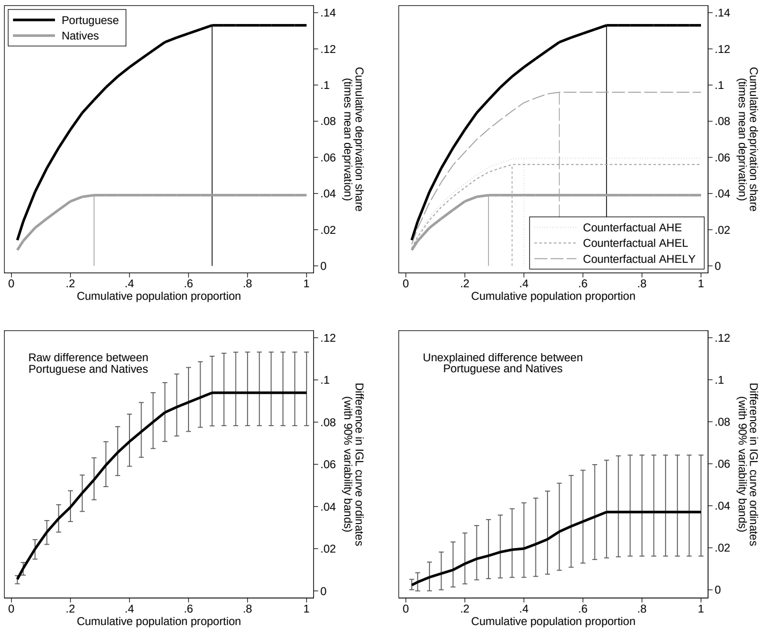

Figure A1. IGL curves (top) and curve differences (bottom) for the 'minimal' item set and Betti-Verma weighting scheme

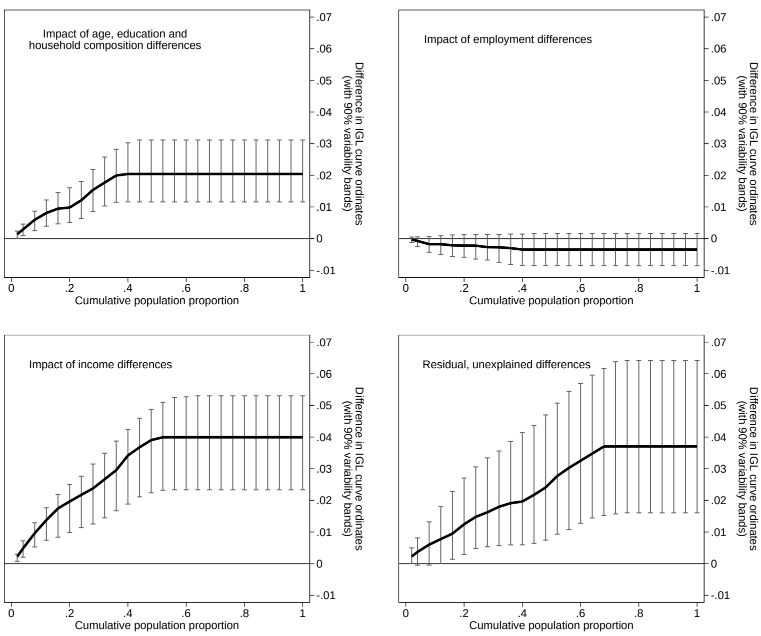

Figure A2. Marginal effects of components

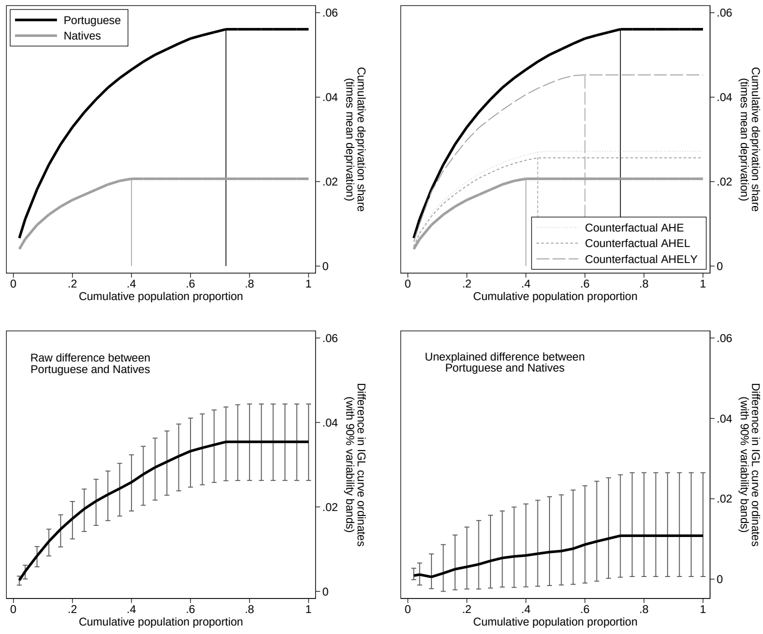

Figure A3. IGL curves (top) and curve differences (bottom) for the 'maximal' item set and Betti-Verma weighting scheme

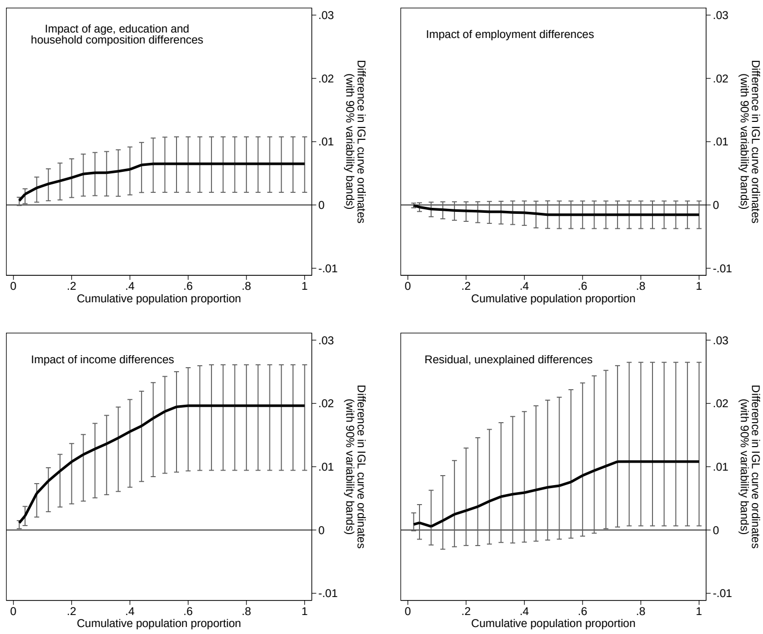

Figure A4. Marginal effects of components

CEPS
INSTEAD

3, avenue de la Fonte
L-4364 Esch-sur-Alzette
Tél.: +352 58.58.55-801
www.ceps.lu