Report on “Essays on Social Welfare Systems, Education and Agglomeration across the EU,” CERGE-EI PhD dissertation by Lubomira Anastassova-Chirmiciu

This dissertation comprises two essays on the labor-market outcomes of immigrants and one on the relationship between agglomeration and labor productivity. The three chapters are well written and competently researched. Overall, I believe this piece of work will merit the award of a PhD but changes should be made in line with my comments below.

Comments on Chapter 1:

Chapter 1 analyzes the factors that determine the extent to which individuals receive welfare payments, focusing in particular on the differences between natives on the one hand, and EU and non-EU immigrants on the other. The findings indicate that non-EU immigrants display a much higher dependency on welfare payments than either natives or EU immigrants.

The concept of social income as it is used in this chapter is somewhat confusing. It is defined as any benefits excluding pensions and labor income -- but then why not call it simply ‘benefits’ or ‘welfare dependence‘?

Related to the above point, the author claims that focusing on ‘social income’ is the main novelty of the paper. Yet, there are a number of papers that focus on immigrants’ dependence on welfare benefits; some of those papers are cited in the literature review. The unique and novel contribution of this chapter should be therefore clarified. Is the concept of ‘social income’, as it is applied here, broader than the concept of ‘welfare benefits’ used in the other papers?

The author argues that another original feature of the analysis is that it separates between EU and non-EU immigrants. Indeed, it is reasonable to expect that immigrants from developed countries are fundamentally different from those in less developed countries. But in that case, differentiating between OECD and non-OECD immigrants would make better sense -- and this would also make it possible to carry out a similar analysis for the US where the EU vs non-EU dichotomy is irrelevant.

It is common for this kind of analysis to be carried out separately for males and females, which is not the case here. Households headed by a female are likely to be significantly different from those headed by a male. This should be remedied, or if such separate analysis is not possible (for instance, due to low number of observations to allow splitting the sample), then at least the reasons for not separating the data by gender should be discussed. If the number of female-headed households it too low, then at the very least a separate analysis should be presented for male household heads.

The analysis only accounts for education by means of a dummy for university degree. Surely, adding dummies for the other categories (primary and secondary, at least) must be better.

The analysis fails to address a related and relevant question: are immigrants attracted by the generosity of the welfare state? If so, then the coefficient estimated for

immigrant status is biased. However, addressing this would be difficult if not impossible in the current analytical framework as all immigrants within a given country are subject to the same welfare state rules. To address this issue, the author could merge the data for the various countries -- to utilize the variation across countries in the generosity of the welfare state -- and then instrument immigrant status. Doing so would be clearly difficult because the analysis only includes five countries and because finding suitable instruments may be quite a challenge. But the results would e rewarding.

While I like distinguishing between EU and non-EU immigrants (or, similarly, between OECD and non-OECD), I do not understand why separate regressions are run for the two categories. It would appear much more natural to me to run a single regression while accounting for the two different categories of immigrants.

The analysis could go further in explaining the difference between EU and non-EU (or OECD and non-OECD)immigrants. For example, the immigrant-status dummies could be replaced by dummies for individual countries of origins or groups of countries. This would reveal whether, for example, immigrants from Southern or Eastern European countries are more similar to those from Northern European countries or to those from non-EU countries.

It is not clear from the discussion whether the analysis includes only households that receive welfare benefits (i.e. have positive social income) or all households. If it is the former, then I would like to know why the households not receiving benefits were omitted. If it is the latter, then it would be interesting to know what share of native and immigrant households receive any benefits (and perhaps a separate analysis of the probability of receiving benefits should be carried out).

Comments on Chapter 2:

This chapter analyzes the differences in the return to education between natives and immigrants in three European countries. In doing so, the analysis first looks at the overall return to education and then distinguishes between returns to ‘typical’ education and over/under education. Immigrants are found to have lower overall return to education. When distinguishing between typical education and over/under education, this gap in return to education is found to be largely due to lower the return to over-education.

The discussion of regression results frequently makes comparisons of the results obtained for France, Germany and Austria with those for the US and UK. Yet, the last two countries are not included in this analysis and instead the comparison is with results published elsewhere. This is confusing as this distinction is not always made very clear. Moreover, the results for the US and UK are obtained with different datasets and different methodologies and therefore are really not comparable.

As I commented on Chapter 1, it is conventional to carry out this kind of analysis separately for males and females. That is not the case here. Why not? Women tend to earn less than men and therefore separating the genders would appear rather important here.

There is much literature arguing that when estimating returns to education, the analysis must account for the possible endogeneity of education. Why is this not done here?

It would be interesting to see whether there is a difference in returns to education between EU and non-EU (or OECD vs. non-OECD) immigrants. And this would also draw a nice parallel between the methodological approaches in Chapters 1 and 2.

Regression coefficients for some control variables are not reported. They should be.

The way 'typical' years of education have been determined appears dubious, as large fractions of workers seem to be either overeducated or undereducated. This is most notable in France where more than half of immigrants are undereducated and even slightly more natives are undereducated than correctly educated. Indeed, one might argue that being overeducated is typical for France. Determining the 'typical' education by means of bands of a few years around the average education for each occupation is less likely to lead to this kind of strange outcomes.

Comments on Chapter 3:

The final chapter considers the relationship between agglomeration and productivity to explain the large differences in productivity across regions. It uses data for NUTS3 and NUTS4 regions in the UK. To account for the possible endogeneity of productivity with respect to agglomeration, the relationship is instrumented using regional population in 1801 and land area. The results suggest that there is a significant and positive relationship between agglomeration and productivity and that the nature of this relationship is more or less the same for both kinds of regions considered.

The novel contributions of this chapter include comparing results obtained with two different levels of regional aggregation, using 1801 population as an instrument (previous literature used population from a later period) and estimating separate regression for urban and non-urban regions. However, in the first two of these cases, the methodological innovation does not seem to make much difference. This is disappointing.

The level of regional aggregation, for instance, should make a difference; otherwise the analysis would be powerless in terms of explaining regional differences in productivity. It is disappointing that the paper does not purse this question further: what is the level of regional aggregation that matters for measuring the relationship between aggregation and productivity? And, related to that, at which level of regions does aggregation affect productivity the most?

Estimating separate regression for urban and non-urban regions yields dramatically different estimates. This should be pursued further. For example, is this because the relationship is in fact non-linear? Is so, then the estimation should involve the quadratic polynomial of aggregation rather than stipulate a linear relationship.

Occasionally, the chapter is rather unclear about what exactly is going on. On p. 80,

for example, it is stated that specification controls for age and gender -- yet given that the analysis is carried out at the regional level, it is not clear what exactly this means.

General Comments:

The dissertation overall is written well and demonstrates a high degree of intuitive understanding and technical competence. However, the analysis often comes across as being rather mechanical, as if the research contained therein was carried out mainly for the sake of doing research (and, in turn, for the sake of earning an advanced research degree). What I miss is a perception that the author herself is excited about her results and that she believes that her work yields some important new findings that move the frontier of our knowledge further out. Furthermore, the analysis also sometimes appears superficial: in Chapters 1 and 2, the analysis is carried out for both genders together but the next logical steps, splitting the analysis for males and females, is not taken.