Welfare, Labor Supply and Heterogeneous Preferences: Evidence for Europe and the US

Dirk Neumann
CGS, University of Cologne

Olivier Bargain

UC Dublin
Andre Decoster
KU Leuven

Mathias Dolls
CGS, University of Cologne

Andreas Peichl
IZA, Bonn

Sebastian Siegloch
IZA, Bonn

In course of the Stiglitz report (Stiglitz et al., 2009), the question of how to measure and compare well-being has gained increasing interest again. The main idea is to move 'beyond GDP' and to recognize the multidimensional character of welfare. Recent studies evaluating well-being across different countries are mostly based on aggregate willingness-to-pay (WTP) measures assuming homogeneous preferences across individuals. This can give a rough guess of preference differences across populations for various non-income dimensions (see Fleurbaey and Gaulier, 2009). The use "of individual data, however seems necessary in order to obtain a reasonable description of the distribution of equivalent incomes" (Fleurbaey, 2008b, p. 21). In (theoretical) labor and public economics, individual preference heterogeneity is a standard assumption. However, the concept is hard to be retained in (empirical) welfare analysis as this requires interpersonal comparisons of utilities what is well-known to be difficult when individual's indifference curves cross. This paper suggests a first step to overcome this problem. We compare countries by means of orderings using individual welfare measures that fully retain preference heterogeneity. Our analysis suggests that respecting differences in tastes may substantially influence how we evaluate individual well-being across countries.

In order to move 'beyond GDP', the present paper extends standard country comparisons to a multidimensional context while we exemplary focus on the leisure-consumption domain. So far, when assessing welfare in a work-leisure space, the problem of interpersonal comparability has usually been addressed by using fixed reference preferences and prices (e.g. Aaberge et al., 2004). This obviously amounts to removing preference heterogeneity from the normative part of the analysis and implies a certain ethical choice, namely that of the reference agent. We make this choice explicit by contrasting the standard approach to multidimensional welfare measures that account for different individual labor-leisure preferences (Fleurbaey, 2006, 2008a). More precisely, they can be interpreted as holding individuals to a varying degree responsible for their distaste for work (taste for leisure). Decoster and Haan (2010) show how to empirically implement these measures relying on German data. We extend their analysis, compute equivalent incomes for individuals across certain countries and rank them globally.

Unfortunately, no tax reform - which is similar for all countries - can be found that would allow identifying labor supply parameters using a natural experiment approach. In this situation, the only way to consistently compare countries is to rely on a common structural model retrieving individual preferences over the consumption-leisure domain. We opt for a flexible discrete choice model and a harmonized approach for 11 European countries and the US. This distinguishes our study from well-being assessments based on indices as GDP in at least four key features. First, using highly disaggregated household data, we are able to focus on income on the micro level rather than using aggregated indices. Second, we broaden our well-being measure to non-market dimensions and activities. Both aspects together allow us to compute equivalent income on an individual (rather than a population) level. Third, we use only ordinal preference information and impose no assumption on the aggregation of individual measures. Fourth, by applying measures based on different rationales about how to treat heterogeneity in tastes, we highlight the importance of the normative choice of the metric in order to rank individual situations.

The main results of our analysis go as follows. For the subsample of married women, we first show that taking preference heterogeneity into account clearly leads to different conclusions with regard to well-being compared to the use of pure income or classic money metric utility measures based on homogeneous preferences. Second, once heterogeneity in tastes is accounted for, our findings suggest that the resulting ranking of individuals across countries remarkably depends on the normative choice related to the metric at use. For example, when looking at income only, married women from the US are best off on average due to high average working hours and wages. The same holds true (in terms of individual well-being) for the metric which assigns “minimal” responsibility for effort distaste. But when moving to measures, that increasingly attach higher responsibility for work aversion, the ranking of individuals across countries changes. For the metric which implies “maximal” responsibility for leisure preferences, women from countries where average working hours are relatively low (Ireland and some Continental European countries) are now better off on average than women from the US (i.e. are considered less needy). We find that these differences in orderings are indeed due to heterogeneous work preferences (or non-captured institutions) to a substantial part. Even though we do not aim at providing complete measures of well-being in order to rank countries, the analysis clearly shows that respecting individual and country specific preferences may have substantial influences when assessing multidimensional well-being in a cross-country perspective.