

Updating GMig2 to GTAP 10

Angel Aguiar

February 2020

GMig2 is based on the GTAP Data Base and extends it to track international bilateral migration and remittances (Walmsley, Ahmed, and Parsons, 2007). Whenever a new version of the GTAP Data Base is released, an update of the GMig2 Data Base can be constructed. The purpose of this document is to outline how GMig2 was constructed for the GTAP 10 version.

Up to GTAP Data Base version 9, migration flows were calibrated based on decennial census data available from the World Bank. The latest available year for these data is 2000. We used to update this information with UN population data for the GTAP reference year, but since the emergence of the Knomad data, we used the 2013 bilateral migration data along with inward and outward remittances data posted on their website (<https://www.knomad.org>).¹

Knomad's bilateral migration is available for 214 countries plus two extra aggregate regions named 'other north' and 'other south'. It appears that for each country, these aggregates can contain different countries. Without any additional information, we could either discard this information or use total immigration and emigration to rebalance the original matrix of 214 by 214 countries. We opt for the latter since discarding information could really affect small countries with large rates of inward and outward migration flows.

Since the latest reference year for GTAP 10 is 2014, we need to update the 2013 Knomad migration data using the United Nations' (UN) population time series data. The UN provides data for 235 countries between 1950 and 2020.

The first step in constructing GMig2 data is to aggregate Knomad's migration data into GTAP 10's 141 regions. Using this migration information, we derive population by country of origin, by computing domestic population as a residual, given the total and foreign population.

We also collect labor force data from the International Labour Organization (ILO), via the World Bank's World Development Indicators. This is available for 186 countries. For each

¹ Currently you can download 2017 bilateral migration, which we will use for GTAP 11. Previously, we were able to download the 2013 data, which is what we use here because is closer to GTAP 10 latest reference year of 2014.

country, the participation rate is computed by dividing its labor force by its population. For countries without labor force data, we use the global participation rate. In order to classify labor force into skilled and unskilled, we use ILO's annual estimates of employment by occupation.²

As explained in Walmsley *et al.* (2007), to split the migration data by skill we use an updated version of the Docquier and Markouk (2005) database (Özden, Parsons, Schiff, and Walmsley, 2011). In the GMig2 database, skills are determined by education, resulting in total labor force being divided into three levels of education: tertiary, primary and secondary education. Skilled labor is composed of those with tertiary education, while unskilled labor is composed of those primary and secondary education.

Migrant labor force abroad is first estimated using the labor force participation rate in the country of origin multiplied by the number of migrants in the destination country (for motivation and details, see Walmsley *et al.*, 2007). For example, if in the US, participation rate is 45%, then 45% of total US migrants in Mexico are of working age. Then this number is calibrated so that the total labor force in the destination country matches official statistics.

Wage data is derived by labor type within each host region using GTAP's value of labor at the different valuations (for details, please refer to Walmsley *et al.*, 2007). The wage rate of workers with skill I from region r in region c is assumed to be equal to the wage in region r plus a proportion (β) of the difference between the wages (host and home). We consider two β s in this equation, β is 0.75 when host wage is greater than home wage and 0.2 otherwise.

The final piece of information is remittances data. As previously mentioned we collect this information from the Knomad website. We then follow Walmsley *et al.* (2007) in deriving a bilateral remittance matrix by assuming a constant share of remittances to income given the wages and number of migrants previously constructed. Finally, savings information from the GTAP Data Base is updated to account for remittances.

² ILO provides 8 categories which are aggregated into skilled and unskilled. GTAP also uses ILO data but distinguishes 5 labor categories, which we in turn aggregate into two skill types. The reason for doing this is that for migration we only have two skill categories. From ILO, we collect 2014 data.

References

- Docquier, F., and A. Markouk, (2005): "International migration by educational attainment (1990-2000) - Release 1.1, F.", In C. Özden and M. Schiff (eds). *International Migration, Remittances and Development* Palgrave Macmillan: New York, chapter 5, 151-199.
- Özden, C., C. Parsons, M. Schiff, and T. Walmsley (2011), "Where on Earth is Everybody? The Evolution of Global Bilateral Migration 1960-2000", *The World Bank Economic Review*, 2011.
- Walmsley, T., Ahmed, S., & Parsons, C. (2007). *A Global Bilateral Migration Data Base: Skilled Labor, Wages and Remittances* (GTAP Research Memorandum No. 06). Purdue University, West Lafayette, IN: Global Trade Analysis Project (GTAP). Retrieved from https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=1880