Preliminary Draft: Do Not Cite

Gendered Employment Data for Global CGE Modeling

Betina Dimaranan, Kathryn Pace, and Alison Weingarden

Abstract

The gender-differentiated impacts of trade reforms and other economic shocks have been assessed in several studies with the use of single-country CGE models. These studies have found that men and women may be affected by trade liberalization differently as the factors of production are reallocated among sectors that employ men and women in different intensities. However, a global approach which could allow for the comparison of gendered economic impacts across countries is hampered by the lack of a gendered employment dataset for global modeling.

The paper builds on work done by Weingarden and Tsigas (2010) to generate employment shares across five occupational levels for 48 countries for use in updating the skilled and unskilled labor splits in the GTAP database. This study uses employment and wages data, differentiated by gender, from the International Labor Organization (ILO). Gendered employment data and industry average wages, across industry and occupations, for each country, is obtained from the Yearbook of Labor Statistics. Wages by job, by gender, are obtained from the ILO October Inquiry data. Following Weingarden and Tsigas (2010), for each country, missing data for wages by occupation and industry are imputed using a minimization of square errors approach.

From the employment (quantity) and wages (prices) data, employment value shares for men and women for five occupational levels and 10-15 industries for several countries are used to econometrically estimate data across all sectors and countries in the GTAP database. Construction of the genderedd database is part of a larger effort to analyze the gendered impacts of trade reforms and other economic shocks across countries with the use of a global CGE database and model.

DRAFT | Do Not Cite

Introduction

The gender-differentiated impacts of trade reforms and other economic shocks have been assessed in several studies with the use of single-country CGE models. These studies have found that men and women may be affected by trade liberalization differently as the factors of production are reallocated among sectors that employ men and women in different intensities. However, a global approach which could allow for the comparison of gendered economic impacts across countries is hampered by the lack of a gendered employment dataset for global modeling.

The paper builds on work done by Weingarden and Tsigas (2010) to generate employment shares across five occupational levels for 48 countries for use in updating the skilled and unskilled labor splits in the GTAP database. This study uses employment and wages data, differentiated by gender, from the International Labor Organization (ILO). Gendered employment data and industry average wages, across industry and occupations, for each country, is obtained from the Yearbook of Labor Statistics. Wages by job, by gender, are obtained from the ILO October Inquiry data. Following Weingarden and Tsigas (WT) (2010), for each country, missing data for wages by occupation and industry are imputed using a minimization of square errors approach.

The GTAP database currently includes a labor component with splits into two categories: skilled and unskilled. The current labor data also is also for both genders, without any distinction. We extend these splits into 5 different skill levels, implementing the method of WT, while also including gender-specific labor data.

Source Data and Issues

Gendered wage data by job is sourced from Harsch and Kleinert (2011). We are using their cleaned data set which is originally based from the ILO October Inquiry data set. They cleaned the data to remove any inconsistencies in wages over time for each country. Harsch and Kleinert do not correct for time differences in reporting (average wages per hour versus maximum wages per month) so we use country level correction factors from Oostendorp (2005), Appendix A, to correct wages to average monthly wage as necessary. We also have correction factors based on gender, so countries that have male gender data only or insufficient female information can be corrected to create a female wage rate. After accounting for available correction factors, we counted the number of occupations with wages available for each country in each year for each gender to determine the “best” years. We determined best year first by most recent year and then by the year with the most information available.

Gendered wage data by industry (ISIC) is sourced from the ILO Yearbook Table 5A. We corrected industry codes to the ISIC Rev 3. codes, cleaned the data, and again counted the number of industries per country per year for each gender to determine best years first in terms of most recent data and then by amount of information. Best years were only kept for countries with sufficient data for both males and females.

Gendered employment data by industry (ISIC) and occupation (ISCO) is sourced from the ILO Yearbook Table 1E. ISIC codes for this set were also corrected to the ISIC Rev. 3 codes and the data was cleaned and counted by industry as previously mentioned.

These three data sets have been merged, and the resulting best years for wage and employment data are recorded in Tables 1 (male) and 2 (female). All of the following 21 countries have sufficient data (at least 10 jobs, 3 occupations, and 3 industries for both genders) from all three sources. Only best years for occupations and employment data from the ILO Yearbook are included in the tables. We attempted

DRAFT | Do Not Cite

to keep the years from all three data sets within a reasonable range of each other, but in certain cases more recent data was not available.

Table 1: Male Most Recent then Complete Data from ILO Yearbook Tables 1E and 5A

Country	Best Years		# Industries
	w (ind)	N (occ, ind)	
Austria	2003	2008	15
Bahrain	2008	1991	17
Bolivia	2008	2007	10
Brazil	2002	2007	17
Colombia	2007	1994	10
Costa Rica	2008	2008	17
Cyprus	2006	2008	15
Denmark	2007	2002	12
Finland	2008	2008	14
Hong Kong	2008	2008	14
Japan	2008	2008	11
Korea (South)	2007	2007	12
Lithuania	2008	2008	15
Luxembourg	2008	1991	12
Peru	1995	2008	13
Portugal	2008	2008	12
Romania	2007	2008	15
Singapore	2008	2008	11
Thailand	2008	1998	14
Turkey	1997	2008	16
United Kingdom	2008	2006	17

DRAFT | Do Not Cite

Table 2: Female Most Recent then Complete Data from ILO Yearbook Tables 1E and 5A

Country	Best Years		# Industries
	w (ind)	N (ind, occ)	
Austria	2003	2008	15
Bahrain	2008	1991	17
Bolivia	2008	2007	10
Brazil	2002	2007	17
Colombia	2007	1994	10
Costa Rica	2008	2008	17
Cyprus	2006	2008	15
Denmark	2007	2002	12
Finland	2008	2008	14
Hong Kong	2008	2008	14
Japan	2008	2008	11
Korea (South)	2007	2007	12
Lithuania	2008	2008	15
Luxembourg	2008	1991	11
Peru	1995	2008	11
Portugal	2008	2008	12
Romania	2007	2008	15
Singapore	2008	2008	11
Thailand	2008	1998	14
Turkey	1997	2008	16
United Kingdom	2008	2006	17

We have 17 countries with wages by job and employment data, but insufficient occupational wage data. These countries can be found by gender in Tables 3 and 4. We will attempt to match this data with reasonable estimates of wages to complete the imputation.

DRAFT | Do Not Cite

Table 3: Male Most Recent then Complete Employment Data from ILO Yearbook Table 1E

Country	Best Years N (ind, occ)	# Industries
Bangladesh	1991	17
Belize	2005	17
Canada	2008	16
Chile	2007	16
Czech Republic	2008	15
Estonia	2008	11
Iceland	2000	16
Kuwait	1995	18
Malawi	1998	16
Malaysia	2008	17
Nicaragua	2006	17
Poland	2008	15
Russia	2007	16
Slovakia	2008	16
Trinidad and Tobago	1999	17
United States	2008	12
Venezuela	1997	17

Table 4 Male Most Recent then Complete Employment Data from ILO Yearbook Table 1E

Country	Best Years N (ind, occ)	# Industries
Bangladesh	1991	16
Belize	2005	17
Canada	2008	16
Chile	2007	16
Czech Republic	2008	16
Estonia	2008	12
Iceland	2000	15
Kuwait	1995	18
Malawi	1998	16
Malaysia	2008	17
Nicaragua	2006	17
Poland	2008	16
Russia	2007	16
Slovakia	2008	16
Trinidad and Tobago	1999	17
United States	2008	12
Venezuela	1997	17

DRAFT | Do Not Cite

In addition to the countries for which data by gender is listed above, we will need to look for further outside sources, particularly for larger countries such as China, India, and the United States.

Data Methodology

We picked one country with all three portions of data available, Brazil, to test the imputation method of WT, which creates labor splits and occupational shares for use in GTAP. We did, however, modify the degree of belief parameter from WT due to the use of monthly, rather than hourly wages which caused the WT degree of belief calculation on the gendered data to be extremely small. We chose to calculate a coefficient of variation using the standard deviation of job level wage data within each industry and occupation divided by the mean of that same wage data. We found that approximately 75% of the data had a coefficient of variation above .1, so any observation without a coefficient of variation was assigned a value of .1.

We use the methods from WT to generate matrices for wages and employment with dimensions occupation by industry. First, a many-to-one concordance is used to match the October Inquiry wage data to ISCO occupation and ISIC industry classifications. Then, the October Inquiry wage data is then averaged within each occupation j, industry i combination to estimate a wage by occupation and industry w_{ij}. Each initial wage w_{ij} is weighted according to our degree of belief parameter $[b_{ij}]$, with more weight given to more accurate observations. For those w_{ij} that do not have underlying wage observations, we assign an initial value equal to the observed median wage by occupation, which is assigned a low degree of belief (0.1) and then adjusted by the minimization of squared errors procedure.

The following equations allow us to minimize the belief-weighted distance between the initial wage matrix $[w_{ij}]$ and the final wage matrix $[W_{ij}]$.[1] The final employment-weighted average industry wage is equal to the industry average wage as given in the ILO data $[w_{.j}]$:

$$\min \sum_{i \in Occ} \sum_{j \in Ind} b_{ij}[w_{ij} - W_{ij}]^2 + \sum_{i \in Occ} \sum_{j \in Ind} [w_{.i} - W_{ij}]^2$$

$$s.t. \sum_{i \in Occ} W_{ij} \cdot N_{ij} = \sum_{i \in Occ} w_{.j} \cdot N_{ij}, i \in Occ, j \in Ind$$

The adjusted wage input matrices (based on average manufacturing wage) for Brazil can be found in Table 5 and output for Brazil can be found in Table 6.

Table 5 Brazil Input Tables

Initial matrix of wages															
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
onetwo	2.93	2.93	5.59	0.53	2.93	2.93	2.93	2.93	2.93	4.41	2.93	2.93	1.29	2.82	2.93
three	1.74	1.74	1.74	1.74	1.74	1.74	1.74	1.74	1.74	1.74	1.74	1.74	1.74	1.74	1.74
four	1.65	1.65	1.65	1.65	1.65	1.65	1.65	1.65	1.12	2.39	1.65	1.44	1.65	1.65	1.65
five	1.33	1.33	1.33	1.33	1.33	1.33	1.33	1.33	1.33	1.33	1.33	1.33	1.33	1.33	1.33
sixtonine	1.02	0.62	2.54	0.54	1.02	1.02	0.75	1.02	0.63	1.02	1.02	1.02	1.02	1.02	1.02
avg by industry	0.41	0.42	1.51	1.00	2.31	0.62	0.60	0.45	0.92	2.50	0.85	1.38	1.41	1.03	0.80

[1] An additional term in the objective function minimizes the distance between each wage observation and its average wage $[w_{i.}]$ by occupation.

DRAFT | Do Not Cite

Degree of belief in wage estimates															
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
onetwo	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.6	0.1	0.1
three	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
four	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.4	0.1	0.2	0.1	0.1	0.1
five	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
sixtonine	0.1	0.1	1.0	0.6	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Number of workers															
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
onetwo	92	1	30	877	38	179	1,111	191	195	198	815	402	542	320	376
three	17	-	29	658	64	117	761	11	140	128	617	425	276	165	307
four	12	-	12	587	34	58	643	92	316	209	399	381	101	107	134
five	5	-	2	223	3	5	2,965	1,055	248	22	583	291	19	46	278
sixtonine	5,200	312	99	1,649	65	1,888	987	16	53	3	65	56	7	11	16

Table 6 Brazil Output Wages

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
onetwo	2.917617	2.928287	2.882414	0.627055	3.355885	2.862417	2.057742	2.363374	2.312136	3.713898	1.424248	2.323933	1.259956	1.101547	1.393675
three	1.742335	1.744471		1.814404	2.462445	1.700659	1.147549	1.711329	1.301706	1.322243	0.606138	1.104483	1.641255	0.859676	0.490264
four	1.645706	1.647242	0.567004	1.709569	2.032861	1.625673	1.142945	1.374665	0.256984	2.235636	0.9107	1.126654	1.609397	1.071716	1.099627
five	1.331289	1.331935	1.146306	1.355599	1.364039	1.329976			0.548175	1.258979	0.255797	0.894694	1.324963	1.087424	0.19665
sixtonine	0.359404	0.408049	1.669354	0.572837	1.740871	0.31205		0.969426	0.457739	1.006404	0.896398	0.932388	1.014151	0.959531	0.949603

We have compared these results to the results of WT, and we see that the changes are in line with what was found in that project. In general, wages are slightly higher in this example compared to the total male and female wages used in WT, but they follow the overall pattern of both the input and output in the WT project. We intend to continue using this method for the remaining countries with sufficient data available and a similar manner to WT.

Once this is complete we will calculate shares and actually implement the idea of WT by adding this gendered data to GTAP. We will expand the current labor splits to 5 more specific skill-related splits and add in the gender component that GTAP is currently lacking.

DRAFT | Do Not Cite

References

Harsch, D. and J. Kleinert, 2011. An Almost Ideal Wage Database Harmonizing the ILO October Inquiry. IAW Discussion Paper No. 71.

Oostendorp, R., 2005. The Standardized ILO October Inquiry 1983-2003. Mimeo. Free University Amsterdam.

Weingarden, A. and M. Tsigas, 2010. Labor Statistics for the GTAP Database. Paper presented at the 13th Annual Conference on Global Economic Analysis, Penang, Malaysia.

DRAFT | Do Not Cite