

## Authors (alphabetical order)

Cristian Jara Nercasseau  
Maryla Maliszewska  
Claudio Montenegro  
Israel Osorio Rodarte  
Javiera Petersen Muga  
Raimundo Smith Mayer  
Huanjun Zhang

22<sup>nd</sup> Annual Conference on Global Economic Analysis

# Gender Disaggregated Labor Database



**WORLD BANK GROUP**  
Macroeconomics, Trade & Investment

University of Warsaw | Warsaw, Poland  
June 19, 2019

## FAQs about the impact of trade policy



- What are the consequences for sub-national region?
- How many job opportunities for younger workers will be created?
- What are the implications for women joining the labor force?
- What are the the type of skills that will be demanded?

## FAQs about the impact of trade policy

- The **distributional** aspects of trade are shaping today's policy discussions in the context of ongoing revisions to the international trade agreements
- While the overall gains from international trade are undisputed, the discussion of **identifying relative winners and losers** from trade policy remains an open debate
- To assess the distributional aspects of trade, **detailed and comparable micro-based harmonized survey is needed**. This data has to be linked to trade policy, more specifically:
  - Countries have **revealed comparative advantage** in **products** or **sub-sectors** (i.e. coffee, cocoa, fruits, vegetables, flowers, metals, machinery, textiles)
  - Trade liberalization and tariff changes occur **at the disaggregated level** (i.e. HS-6 digit code)
  - And affect **workers** differently (managers, engineers, clerks, agricultural workers) and with strong **gender** implications

## OUTLINE

- 1.** Motivation for constructing a new database
- 2.** Methodology
- 3.** Work ahead



# 1

## Motivation



## Motivation

---

(Re)-Examination of the **distributional** consequences of international trade beyond macroeconomic indicators

---

Nevertheless, detailed comparable data for **global** policy analysis is **scarce**

---

Objective: Create a **comparable micro-based** dataset that in the form of a public good, provides **detailed** accounts on employment levels **by gender**, occupation, labor income\*, employment status at a **finer level** of disaggregation in the economic activity.

---

## Applications within the World Bank

---

a) **Processed “ready-made” statistics**, i.e. World Bank - WTO 2019 Trade and Gender Report

i.e. Women's labor intensity and gender gap by sector

b) **Statistics in flexible format** to inform other databases.

i.e. the **Global Trade Analysis Project Database**

Correct homogenous skill-intensity within manufacturing sector

c) **Micro-data** for World Bank Simulation Models

**For example:** Impact of trade policy on WB client countries

AfCFTA, CP-TPP, Brexit, NAFTA2.0

---

# 2

## Methodology



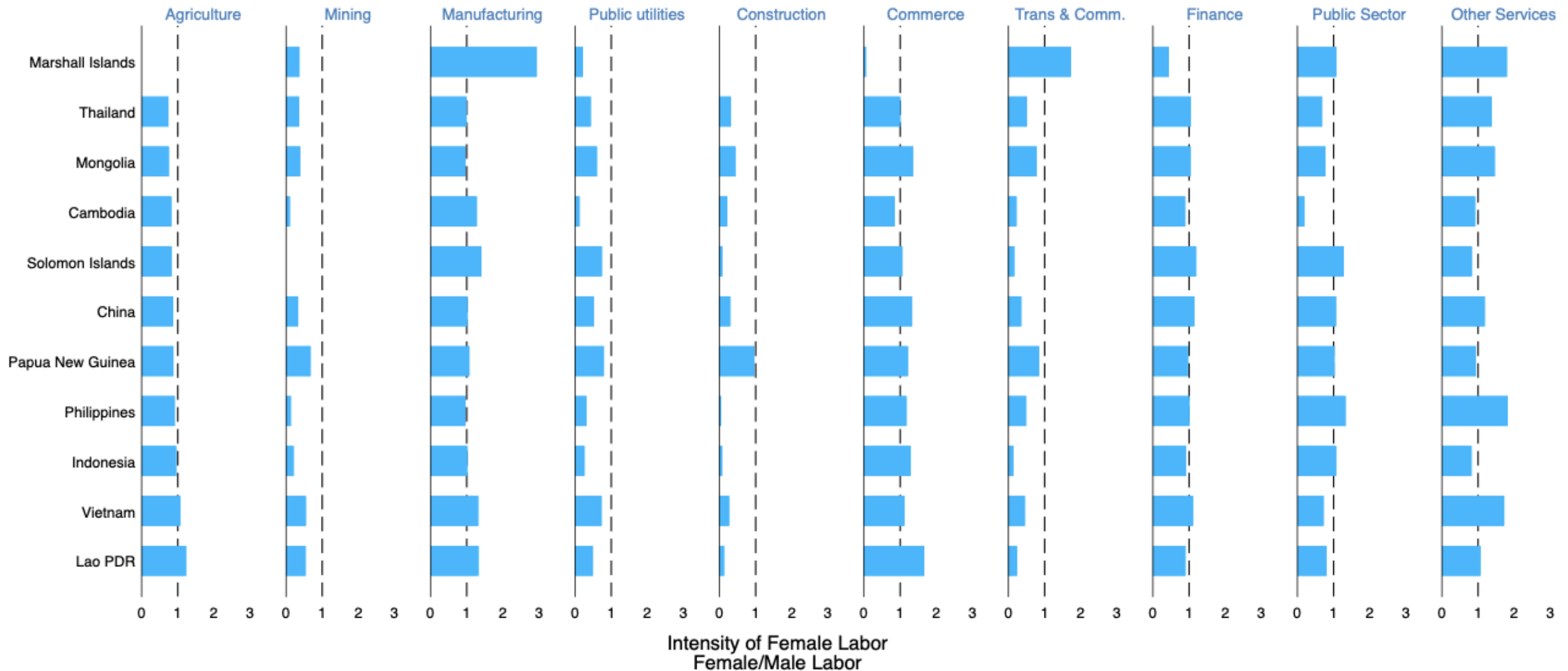


## Starting point

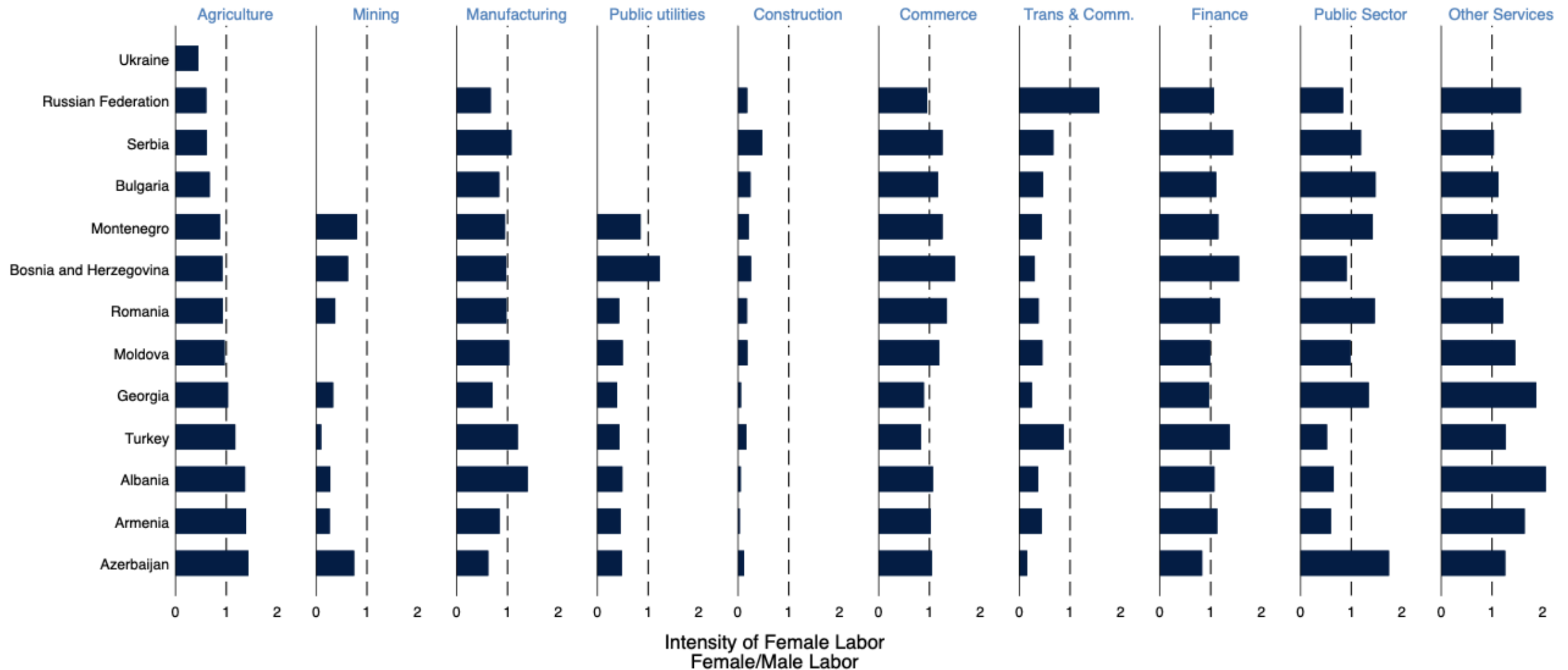
- The World Bank has harmonized data for 138 countries with **broad** industry and occupation variables that account for 90% of population and GDP
- By linking WB international trade models (CGE) with existing harmonized household survey data, for instance, the effect on the WB **twin-goals** can be assessed
- Beyond the twin-goals, some information can be used to build labor volumes and relative wages, by sector and gender



## Starting Point: East Asia & The Pacific

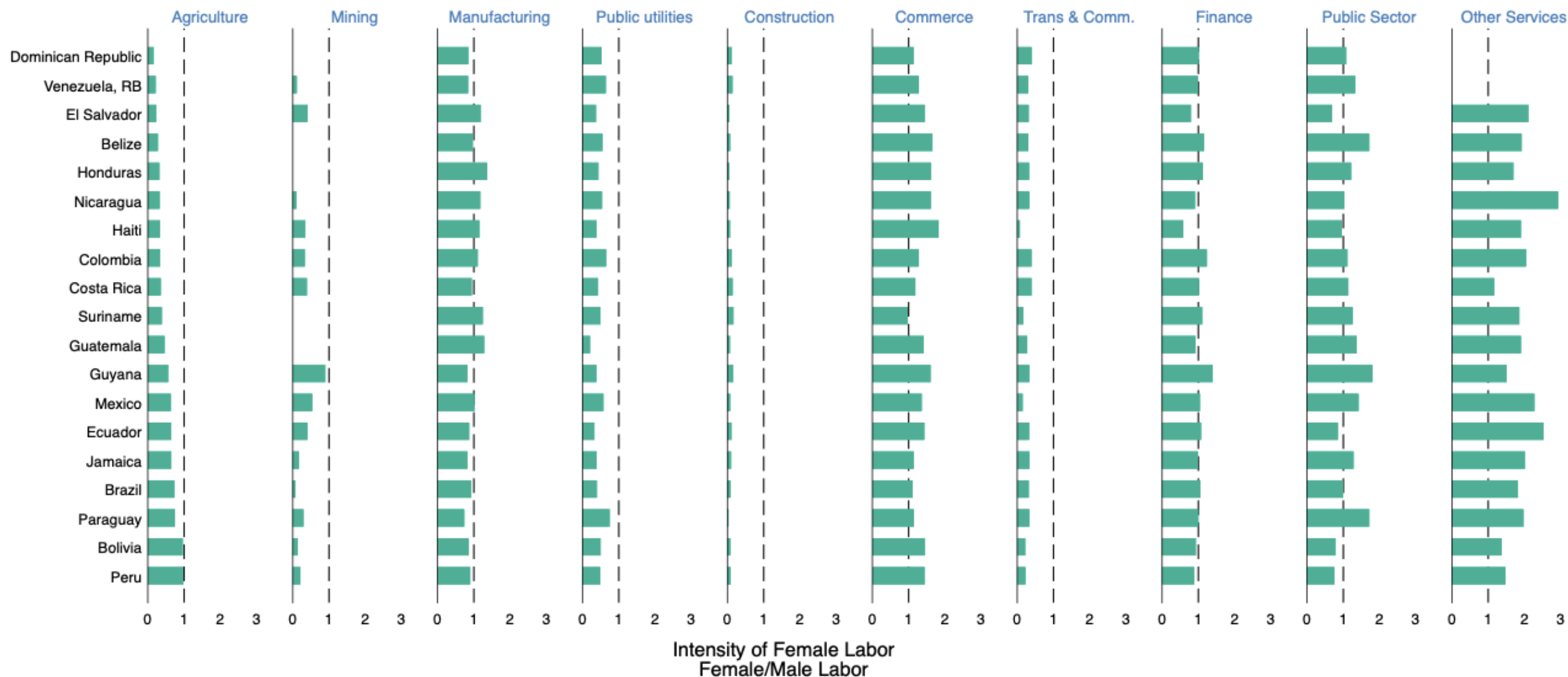


## Starting Point: Eastern Europe and Central Asia



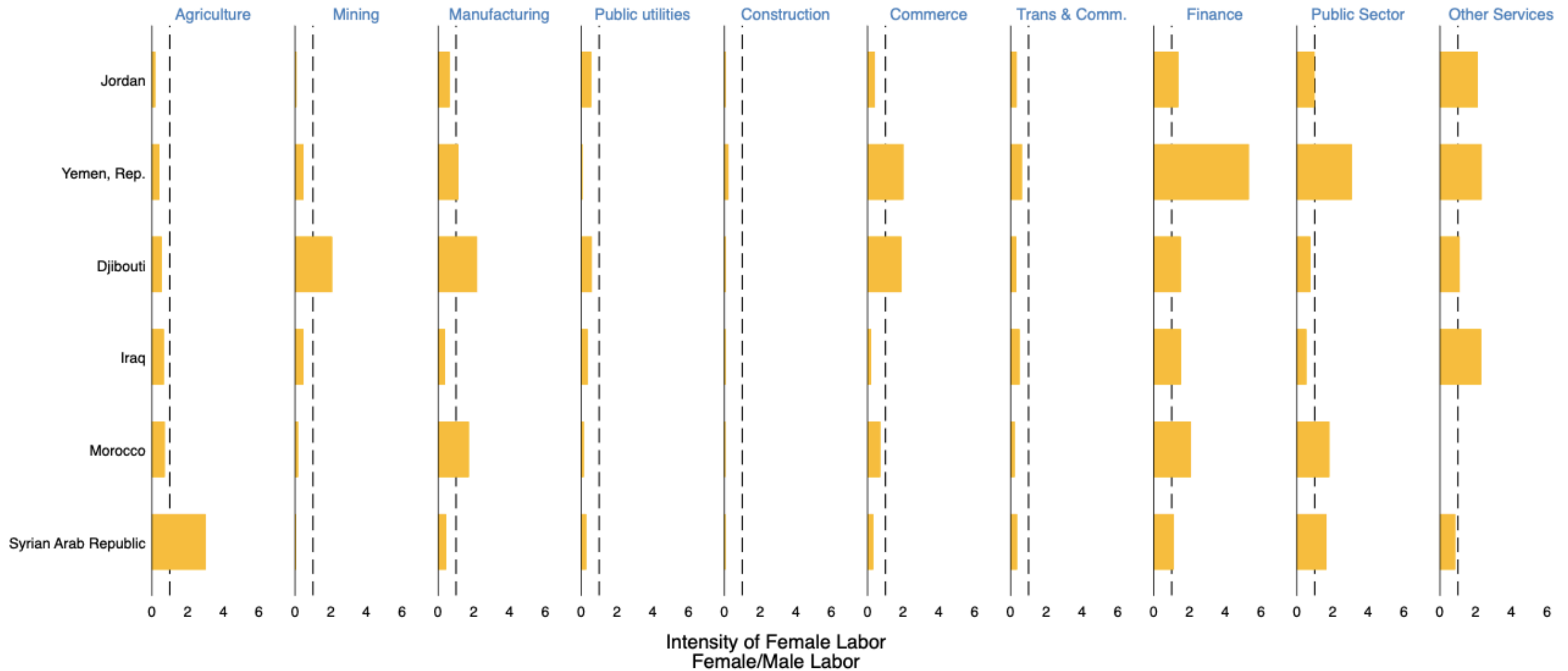


## Starting Point: Latin America & The Caribbean



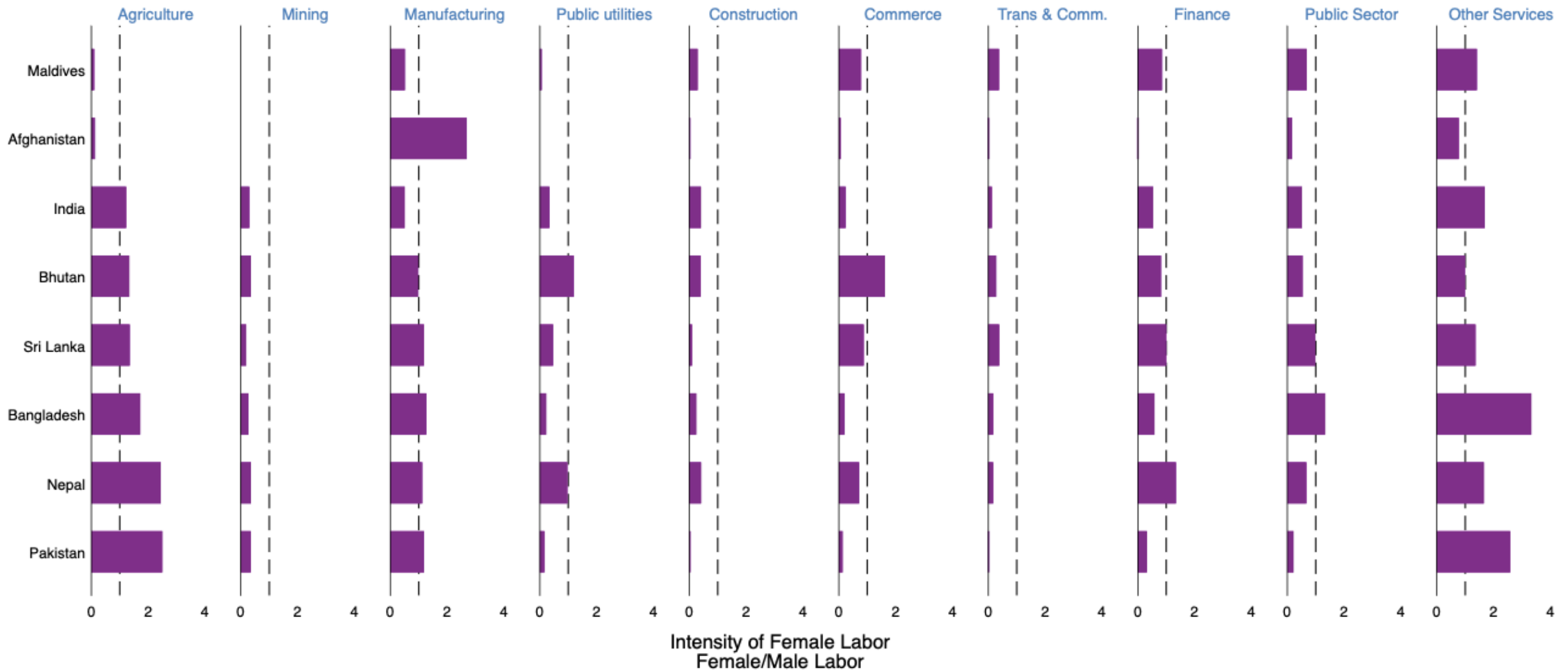


## Starting Point: Middle East & North Africa



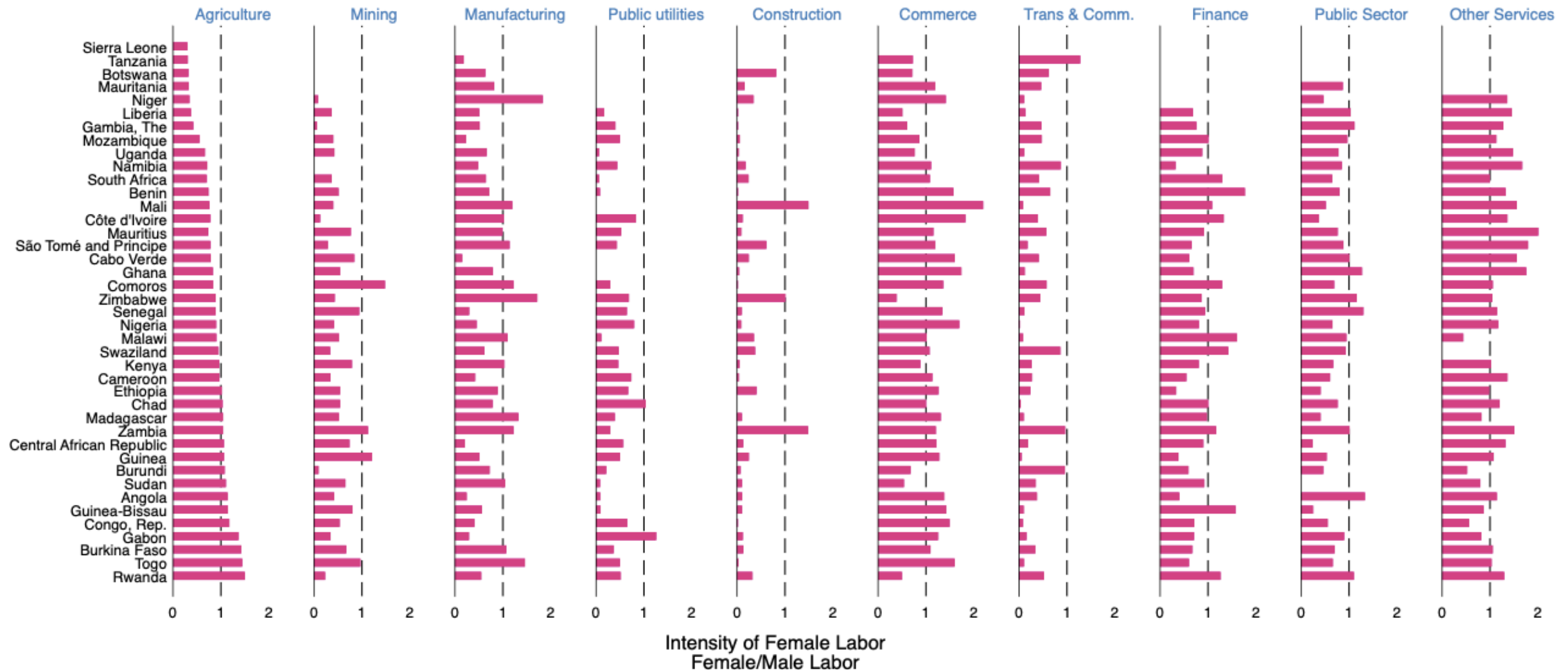


## Starting Point: South Asia



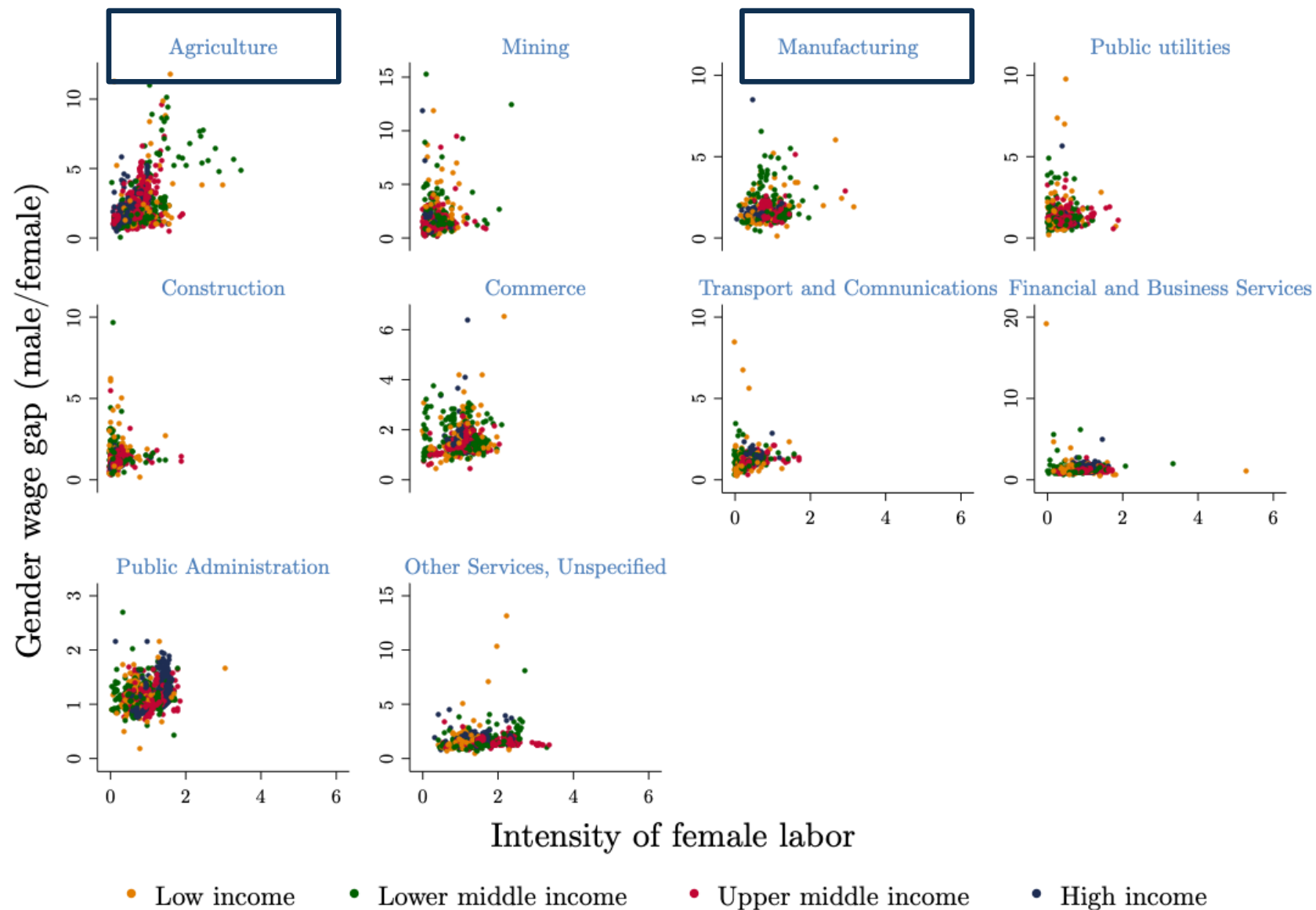


## Starting Point: Sub-Saharan Africa





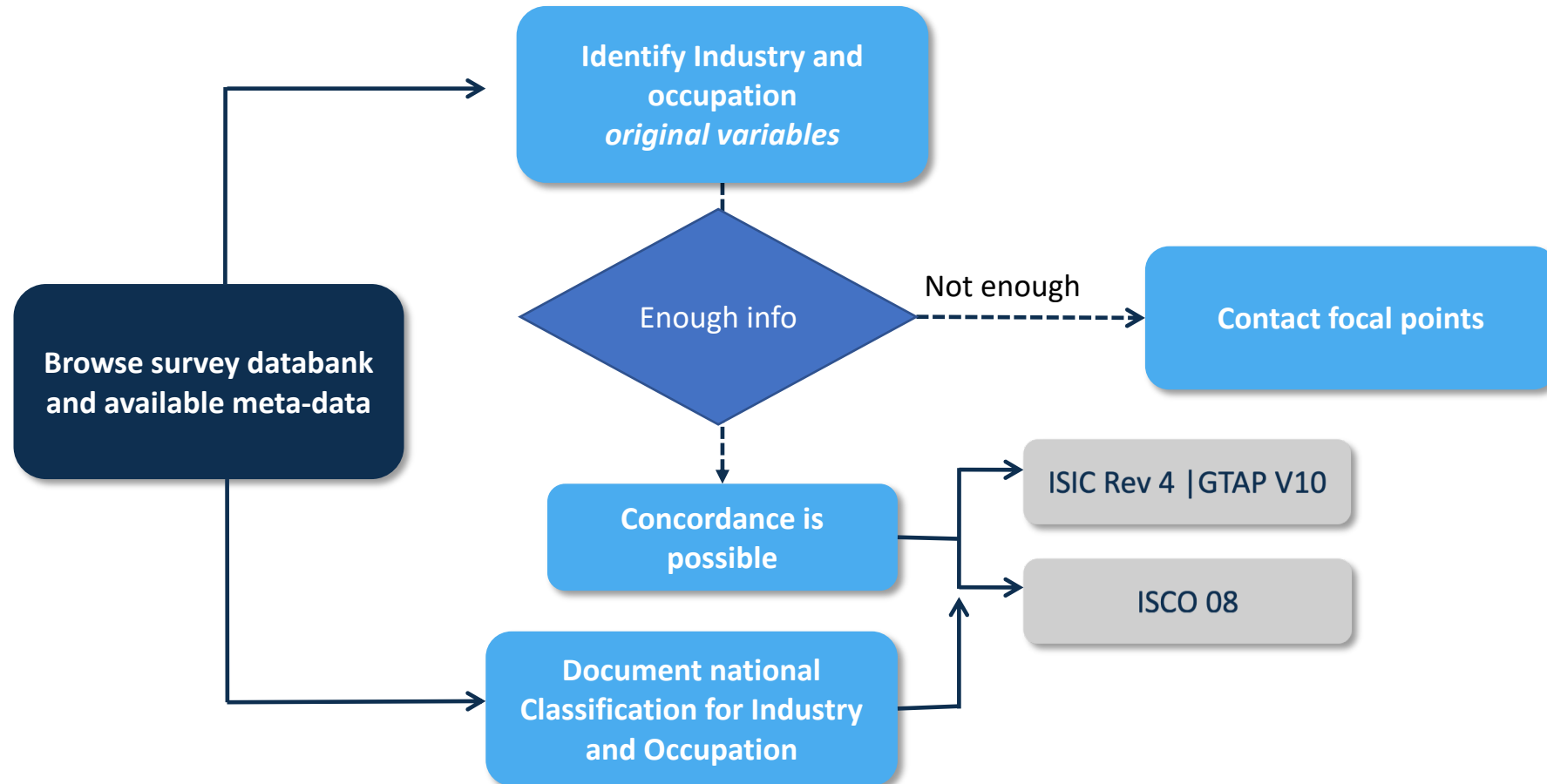
## Starting point: Global





## Process for Constructing the Gender Disaggregated Labor Database

Gender Disaggregated Labor Database creates a finer level of disaggregation for “industry” and “occupation”.



## Building a Disaggregated Labor Database, limitations

- On the microeconomic side, it requires a re-harmonization of industry and occupation data relying on a sub-sample of the World Bank collection of household surveys
- The strategy has been to find initial documentation for a cross-section of countries
- Finding disaggregated data at industry level in household surveys is not easy, because of:
  - **Focus.** The World Bank efforts to harmonize household survey data are concentrated on collecting data for monitoring “global poverty”. This has created a gap in survey coverage for **a.** data on high-income countries, and **b.** an emphasis in poverty, rather than in the production of labor statistics.
  - **Survey design.** Either when the classification system present in the household survey cannot be further disaggregated. Due to sampling, most surveys are not representative at the highest disaggregation level
  - **Data quality:** Even if data can be disaggregated, there are some internal inconsistencies, inherent to data collection in difficult environments and with teams with low-statistical capacity
  - **Inconsistent meta-data:** Comparable international information about national industry and occupation classifications doesn't exist and country information is not always easily available

## Building a Disaggregated Labor Database, progress on meta-data collection

- Meta-data about national industry and occupation classification has been collected for 96 countries/surveys, including number of digits in the survey
- Missing meta-data for 38 countries/surveys
- Missing meta-data for 29 countries/surveys in both industry & occupation:
  - 12 in SSA (AGO, BEN, CMR, COM, CIV, LSO, MDG, MWI, NMB, TCD, SEN, TGO)
- Missing meta-data for 5 countries/surveys for occupation:
  - 2 in SSA (NER, MLI)
- Missing meta-data for 4 countries/surveys for industry:
  - 3 in SSA (ETH\*, BFA, LBR)

## Building a Disaggregated Labor Database, progress on performing conversion

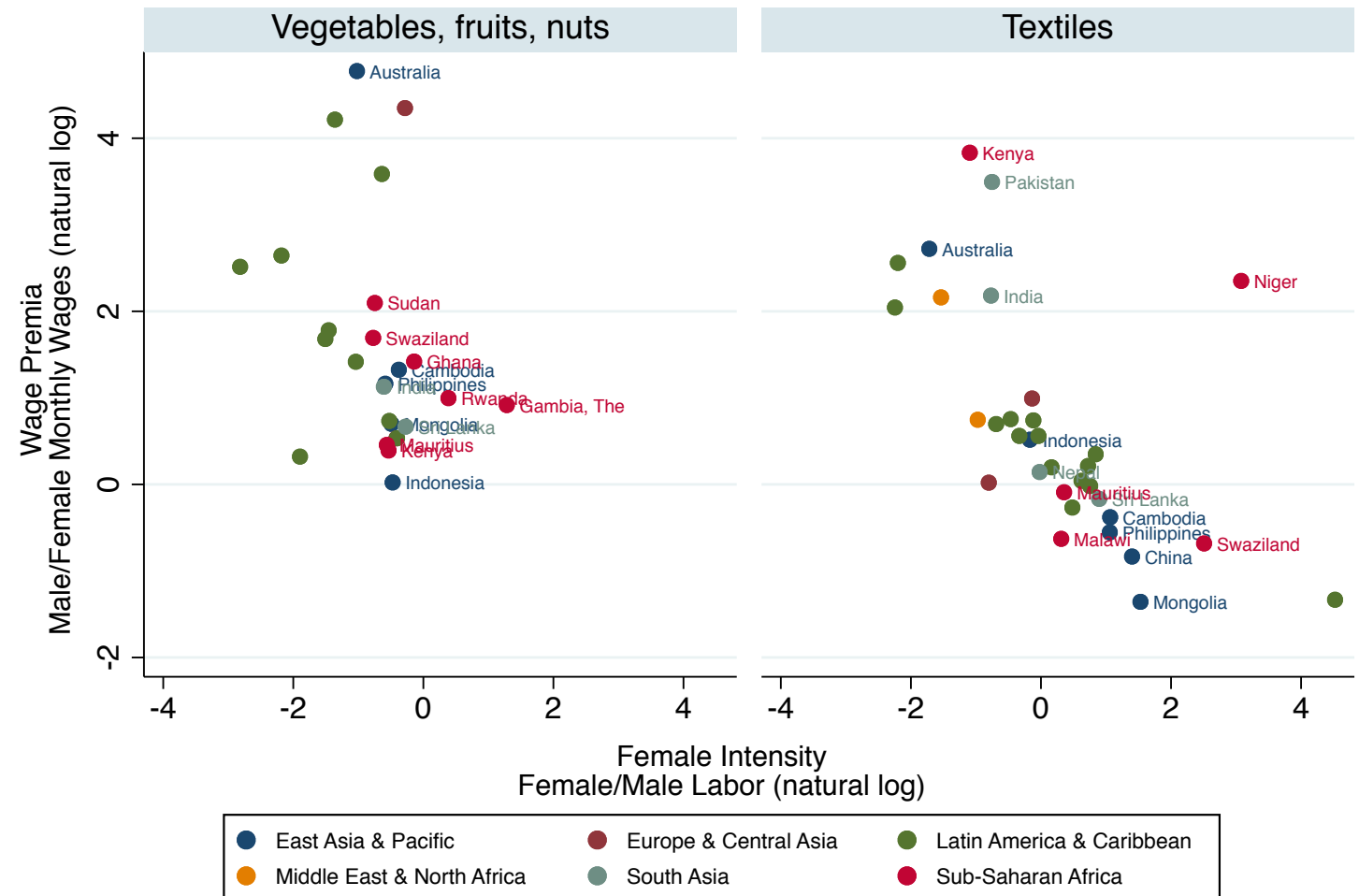
- **a.** ISIC Rev 4 and GTAP 10 codes and/or **b.** ISCO 08 codes have been processed for:
  - 67 countries/surveys with “some” access to microdata
  - Statistics for 27 European Union countries/surveys (based on NACE 2.0 and ISCO 08) were obtained from the Luxembourg Income Study
  - 10 countries/surveys are currently in the pipeline due to need to build specific (non-ISIC/non-ISCO) classification systems

## Summary of Data Processing

TYPE	SUB TYPE	EAP	ECA	LAC	MNA	NAM	SAS	SSA	Total
Processed	Only partial statistics								
	Both industry and occupation	9	4	12	4	1	6	12	48
	only industry	1	1	1				1	4
	only occupation			2				4	6
	<b>TOTAL</b>	<b>10</b>	<b>5</b>	<b>15</b>	<b>4</b>	<b>1</b>	<b>6</b>	<b>17</b>	<b>58</b>
Incomplete Metadata	No industry / No Occupation	2	8	5	2			12	29
	Only industry		3					2	5
	Only occupation	1						3	4
	<b>TOTAL</b>	<b>3</b>	<b>11</b>	<b>5</b>	<b>2</b>			<b>17</b>	<b>38</b>
Currently Processing		1	27+2	2	2	1	2	2	39

## Building a Disaggregated Labor Database, what can be done

- Two prominent options for diversification in Africa are a. Vegetables and Fruits and b. Textiles
- Vegetables and fruits have higher value added than traditional agricultural products such as crops, coffee, cocoa
- Textiles, on the other hand, have been the base of manufacturing expansion in East Asia
- Mainly due to the intensity of female labor and gender wage premia, both sectors can have different implications in job creation, by gender



Graphs by GTAP(65) industry classification

**3**

**Work ahead**



## Work ahead

Finalize the processing of surveys (early July 2019)

Establish methods to fill-in the gaps for missing information (August, 2019)

Make clean data and documentation available to the public (December, 2019)

Collaboration with GTAP community to link with GTAP Database (-> June 2020)

Generate a set of policy recommendations and work with development partners to inform statistical offices improve data collection and meta-data documentation process



## Better quality data is translated into more effective policy-making

- The Open Data Initiative established institutional rules to transform World Bank data into a global public good
  - Microeconomic household survey data is a sub-set of the large collection of data within the World Bank Group
- The World Bank harmonizes cross-country microeconomic survey data to:
  - Monitor twin-goals of reducing poverty and sharing prosperity
  - Support World Bank operations
  - Perform economic modeling, research and inform flagships
- The harmonization of household survey data is a labor-intensive continuous process
- In part thanks to World Bank engagement with statistical offices, the quality and frequency of household survey data (especially in SSA is improving), but still there are important gaps in terms of comparability, documentation and quality controls with respect to international standards.



**WORLD BANK GROUP**  
Macroeconomics, Trade & Investment

[www.worldbank.org](http://www.worldbank.org)