Module 4

The sectoral approach
1 Introduction

The aim of this module is to introduce the methodologies that can be used to analyse the effects of trade on gender at the sectoral level. In particular, we will look at how the existing empirical literature investigates the impact of trade shocks and trade policies on women engaged in specific sectors and industries of the economy. As mentioned in Module 1 of this volume, while there is value to qualitative studies on trade and gender (e.g. Shayo, 2012), the focus here is on a quantitative analysis.

Trade shocks and trade policies will affect different sectors in different ways. Some sectors will contract, as they will not be able to compete with imports, and others will expand as a result of trade-led specialization. This implies that resources, and in particular labour, will have to move from one sector to another. However, it is important to note that this process is not automatic, but is rather influenced by labour market frictions, which vary across countries and sectors and shape countries’ patterns of specialization. Sector- and country-specific labour market frictions also imply that unemployment rates vary by sector and country. As mentioned in Module 1 of Volume 1 of this teaching material, women also tend to be segregated into fewer economic sectors – what we have called “horizontal gender segregation”. Therefore, it is important to also evaluate the gender effects of trade at the sectoral level. With this type of study, we can estimate, for instance, export premiums or mobility costs by gender, and explore whether trade affects women more than men. We can also study indirect effects on important issues such as domestic labour-sharing among members of the household, investment in children’s education, and even gender selection at birth.

There has been a recent effort in economic literature to develop trade models that explore how the correlation between comparative advantage and labour market frictions at the sectoral level can explain the heterogeneous impact of trade on unemployment and gender inequality. Our objective here is not to review these new theoretical developments, but to look at some of the empirical studies that document the evidence these models try to explain.

Section 2 of this module is a summary of studies on trade and gender using the sectoral approach. In particular, we briefly introduce the literature on global value chains, which represents an additional method of exploring the relationship between trade and gender at the sectoral level beyond the study in the hands-on application, examined later in this module. Section 3 provides the intuition behind some of the methodologies applied in the empirical sectoral studies. For the hands-on application in Section 4, we have selected a paper that uses a quasi-experimental approach to study how labour income opportunities for women in a non-traditional exporting sector may affect the time women and men spend on housework. Some final remarks are offered in Section 5.

At the end of this module, students should be able to:

- Apply the sectoral approach to the research on the relationship between trade and gender;
- Review and summarize the literature exploiting the sectoral approach to study the relationship between trade and gender, including the literature on global value chains;
- Identify the differences between the sectoral approach and the microeconomic and macroeconomic approaches presented in the previous modules;
- Understand and describe the difference between truncated and censored data;
- Understand the econometric techniques for the treatment of truncated and censored data, such as the Tobit and Heckman sample selection models;
- Replicate, using Stata, the results of the paper by Newman (2002) titled “Gender, Time Use, and Change: The Impact of the Cut Flower Industry in Ecuador”.

2 Review of the literature

This section discusses a few studies that have used the sectoral approach to analyse the interlinkages between trade and gender inequality. In particular, the papers cited here look at the agricultural sector—where women represent the majority of casual and seasonal workers—and the garment sector—where women are mostly employed as subcontractors and home-based workers. This section is a non-exhaustive survey of the literature but it examines a few interesting papers that you may wish to add to your reading list.

We start by presenting the literature on global value chains, which includes a wide selection of detailed studies on trade and gender inequality at the sectoral level. The basic idea of this strand of literature is that developing countries should take advantage of global value chains to improve their economic performance and thus achieve better labour conditions or social improvements.
Participation in global value chains may on the one hand incentivize firms to produce higher quality/value-added goods, and thus foster the use of a more skilled and formalized labour force; on the other hand, firms may react to increasing global competitive pressures by cutting labour costs and increasing the flexibility of the work force. In the latter case, women, among other weaker groups of the population, would be the most affected because they represent a large portion of irregular and informal workers, and because gender constraints often limit their ability to access work opportunities in global production (Barrientos and Kabeer, 2004). Overall, the aim is to understand how both firms and workers can stand to benefit from increased participation of developing countries in global value chains (University of Manchester, 2010).

Barrientos (2013) offers a gender perspective to the analysis of global value chains. The author focuses on global value chains in the cocoa sector of Ghana and India and finds that the participation of female cocoa farmers and workers in global production contributes to their empowerment and to more sustainable and quality production. According to the study, women in both countries have been long relegated to household activities and forms of casual work due to gender social norms and practices. However, they play a relevant role in the production of quality cocoa because their work is mostly concentrated in activities (early plant care, fermentation, and dying) that are considered crucial to ensure good yields, thereby attracting foreign chocolate companies. Another paper on trade and gender inequality in the context of global value chains is Tallontire et al. (2005), who focus on the African horticultural sector, where women represent the bulk of insecure seasonal and casual workers. The paper explores whether ethical trading practices (codes of conduct covering employment conditions, and environmental and social standards) applied by European buyers in the horticultural value chain may reach women workers, among others, and whether those practices contribute to the improvement of their working conditions. The horticultural sector is also the focus of Maertens and Swinnen (2009), who present the various mechanisms through which women are directly affected by the emergence of modern supply chains. They also review existing empirical evidence and present new quantitative evidence for the high-value horticulture supply chains in Senegal. Their findings suggest that the growth of modern horticulture supply chains has been associated with direct beneficial effects on rural women, and that it has reduced gender inequalities in rural areas.

With regard to the textiles and garment sector, the integration of export-oriented firms in global value chains has been found to worsen the working conditions of women. Dedeoglu (2010) studies the case of Turkey, where garment exporters have increasingly relied on informal labour through subcontracted and home-based female workers as a strategy to reduce production costs. On a similar note, Rossi (2001) finds that in the garment sector of Morocco, women, who generally have less bargaining power, are mostly engaged to perform unskilled activities, including packing and loading. In conclusion, firms in the low-value segments of global value chains, which are mostly concentrated in developing countries, make use of informal and low-paid jobs, for which female workers are usually preferred, to compete in foreign markets.

We have already mentioned a couple of other types of sectoral studies in Section 2.3 of Module 1 of this volume. Here we include a few more, including Porto et al. (2011), who look at how the internal structure of agricultural export markets and the level of competition affect poverty and welfare in rural areas of Africa. They conduct twelve case studies and find that in nine of the twelve simulations the increased competition has a larger positive income effect on male-headed households than on female ones. This is not surprising, given that women normally face entry barriers to participation in cash crop production (Vargas-Hill and Vigneri, 2011). Ackah and Aryee et al. (2012) instead focus on the cocoa sector. They assess whether cocoa production, and therefore the income generated by it, is controlled by males and whether this in turn causes gender inequalities to be reinforced by the promotion of cash cropping in rural areas. They find that this is not the case, at least not for cocoa in Ghana.

Of the three methodological approaches described in this volume, the sectoral approach is perhaps the most heterogeneous from a technical point of view. We have seen that sectoral studies on trade and gender can be carried out with a gender analysis of global value chains through simulations or with more traditional econometric techniques. The paper reviewed in Section 4 employs yet another technique based on quasi-experimental data. If your research question on trade and gender requires the adoption of a sectoral approach, you should also think about the most appropriate technique to use.
3 Methodological approach

While the microeconomic approach explained in Module 2 in this volume was taken from the trade and poverty literature, and the macroeconomic approach described in Module 3 was mainly based on the growth, trade openness, and women’s empowerment literature, empirical analyses on the relationship between trade and gender at the sectoral level can be carried out in many different ways.

Therefore, our focus in this section will be on providing the intuition of a series of econometric methods that are often used in case studies, but without exhausting the toolkit available to the researcher. These methods relate to the statistical treatment of truncated and censored data as described below.

3.1 Truncation and censoring

We define a variable as truncated or censored when we cannot observe all the possible values that this variable takes. Specifically, a variable is censored when we only know the true value of $Y$ for a restricted range of observations. Values of $Y$ are in a certain range and reported as a single value or there is significant clustering around a particular value, for instance zero. An example of a censored variable is when we consider data on consumers and prices paid for a good: if a consumer’s willingness to pay for a particular good is negative, we will have observations with consumers’ information but not the “real” price for that good, as price observations are censored at zero.

On the other hand, a variable is truncated when we only observe values of $X$ in case $Y$ is not censored. In this case, we do not have a full sample for $(Y, X)$ as we exclude observations based on the characteristics of $Y$. The truncation is a result of sampling only part of the distribution of the outcome variable. A variable can be truncated because of the survey design (you only have a sample of women who work) or because of a particular incidence (if you are studying the wage offered to married women, you only have wage information for those women who actually are in a job). Truncated data differ from censored data in that for the latter we still observe values for $X$ when $Y$ is censored.

3.2 Tobit model

The Tobit model is a censoring model applied to a linear model with normal residuals. The textbook example of a Tobit model is that of the labour supply of married women. In fact, labour supply is a two-stage decision process. In the first stage, a woman has to decide whether or not to work (is the salary higher than her reservation wage?). This is a probit model because her participation decision is a binary outcome variable ($o = \text{no work}, 1 = \text{work}$). If she decides to work, the second decision is about the number of hours she is prepared to work. This can be considered as a linear regression model where several factors can explain how many hours (a continuous variable) she decides to work. Both decisions are related: factors that make a married woman more likely to participate in the labour market tend to make her work more hours. Typically, as regressors we would include variables such as education level, non-labour income, spouse’s income, number of children, general economic conditions. However, to estimate the labour supply, we require information on wage offers and we do not observe this information for those women who are not working. Moreover, a wage offer is also likely to be related to unobservable characteristics that also affect the decision to work. If we use only the observed wages and labour supply decisions to estimate the regression, the estimated coefficient will be inconsistent under ordinary least squares because of the selection bias. The Tobit model proposes an alternative to this, as it is a combination of a linear regression model for the estimation of the variables that influence the number of hours supplied and a probit model for the estimation of the likelihood for an individual to participate in the labour market. This model, which uses a full sample (i.e. both women who work and those who do not), can be estimated using maximum likelihood. Under normality, it provides consistent estimators.

3.3 Heckman sample selection model

The Heckman (1979) sample selection model is a type of Tobit model in the context of our example, the equation that estimates the labour supply of women (or the wage paid to those who work) is:

$$w_i = \beta X_i + u_i \quad (1)$$

where $w_i$ represents women’s (hourly) wages, $X_i$ is a set of individual characteristics, and $u_i$ is the error term. However, there is a second equation that determines if a person is willing to work or not. This second equation is called the sample selection equation and it takes the following form:

$$h_i \equiv \gamma_i (w_i - w^*_i) = \pi Z_i + \epsilon_i \quad (2)$$

Here, the reservation wage $w^*_i$ is the minimum wage at which individual $i$ is willing to work and
The sectoral approach

module

variable in equation (1) above, working – i.e. we are only observing the wages as a whole but only that of women who are Therefore, we are not observing the population progressing wages on characteristics (equation (1))

Heckman sample selection model is that by re-

The key problem that we try to address with the

\[ \text{The Heckman model assumes:} \]

\[ \begin{align*}
\{e, u\} & \sim N(0, 0, \sigma^2_\varepsilon, \sigma^2_u, \rho_{\varepsilon u}) \quad \text{both error terms} \{e\}
\text{and } u \text{ are normally distributed, with mean equal to } 0 \text{ and variances } \sigma^2_\varepsilon \text{ and } \sigma^2_u, \text{ respectively. The error terms are also correlated and } \\
\rho_{\varepsilon u} & \text{indicates the correlation coefficient;}
\{e, u\} & \text{is independent of } X \text{ and } Z. \text{The error terms are independent of both sets of explanatory variables.}
\end{align*} \]

The Heckman sample selection model is that by re-

\[ \begin{align*}
\text{The CLAD estimation method was proposed by Powell (1984). In its linear model version, the method of least absolute deviations produces regression coefficient estimates by minimizing the sum of absolute residuals:}
\sum^n_{i=1} |y_i - \max(0, x_i' \beta)|
\end{align*} \]

Chay and Powell (2000) explain the intuition of the methodology. CLAD is similar to a typical regression model but in the context of censored data. When you run a standard linear regression model, you are looking at sample mean relationships between variables \( y \) and \( x \). In CLAD, you are looking at the sample median (instead of the mean) relationship between the variables in a context where the dependent variable is censored.

Let us consider a latent variable model where the variable \( y \) is only observed when it is positive.

\[ y^* = x \beta + \varepsilon \quad (4) \]

\[ y = \max(o, y^*) \quad (5) \]

Assume that conditional on \( x \) the error term \( \varepsilon \) has a median of zero: \( \text{med}(\varepsilon | x) = 0 \). This assumption implies that the median of \( y \) conditional on \( x \) is \( x \beta \) if \( x \beta > 0 \) and zero if \( x \beta < 0 \). Then, because not all true values of the dependent variable \( y^* \) are observed, we cannot directly use the least absolute deviation method to estimate the unknown coefficients. The CLAD estimator applies extra censoring (e.g. applies censoring from above if the outcome is already censored from below) to a least absolute deviations (median regression) estimator. Unlike least squares
regression, least absolute deviations regression does not have an analytical solving method. Therefore, an iterative approach is required. For this reason, and as we will see in the hands-on application in this module, the CLAD command needs to be run more times before arriving at the final estimation output.

One of the solving methods is the iterative linear programming algorithm proposed by Buchinsky (1991). This is the procedure used by the command `clad` in Stata. The method consists of estimating successive quantile regressions, dropping in each estimation the observations for which the predicted value of the dependent variable is less than the censoring value (in our example, zero). The procedure is stopped when no negative predicted values are obtained in two consecutive estimations.

4 Hands-on application: “Gender, time use, and change: Impacts of the cut flower industry in Ecuador” (Newman, 2002)

4.1 Context and overview

Newman (2002) studies women’s household time allocation in a context where economic reforms have led to the growth of a non-traditional agricultural export sector, specifically the Ecuadorian cut flower industry. The paper uses quasi-experimental data from Ecuador to understand the impact of an increase in women’s employment opportunities in the cut flower industry on household paid and unpaid labour allocation. For this purpose, the paper compares data from the cut flower industry area (with high demand for female labour) with data from areas that are culturally and ecologically similar to the previous one, but that have been less influenced by the boom of the non-traditional agricultural exporting sector.

The paper addresses two main questions concerning the changes in household time allocation as a result of the expansion of the cut flower industry. The first is whether women who work in the flower industry are working more hours as a result of combining market labour with unpaid household labour. The second is whether some responsibilities for unpaid household labour have been shifted from female to male members of the household.

The analysis in the paper shows that with the increase in labour market opportunities for women, women’s total time spent in the labour market remains the same, while men’s time in unpaid labour increases. Table 16 provides an overview of Newman’s paper.

<table>
<thead>
<tr>
<th>Table 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overview of Newman (2002)</td>
</tr>
</tbody>
</table>

**Objective**

- Examine the effects of women’s employment in an export industry on the allocation of paid and unpaid labour within the household.
- Separate the bargaining effects of wages from the substitution effect of wages on time use:
  - Substitution effect: following a wage increase, the individual substitutes housework for paid work;
  - Bargaining effect: higher wages increase the person’s ability to survive independently of the household and this reduces his/her domestic work supply.

**Methodology**

Three models are considered in the paper:

(a) **Heckman model**
- Two-step estimation model, initially chosen to correct for sample selection;
- Requires the identification of at least one exclusion restriction in the first-stage equation model;
- The results of the estimation greatly depend on the choice of the exclusion restriction; to avoid this problem, the author implements the Tobit model.

(b) **Tobit model**
- Single-equation model that takes into account the censoring of values at zero;
- Allows for all the variables to be tested as possible determinants of shares of participation;
- Requires the errors to be normally distributed, but the test for normality fails, so the author decides to use the censored least absolute deviation as an alternative methodology.

(c) **CLAD model**
- Does not require any assumptions about the distribution of the error term;
- The disadvantage of using this estimator compared to the Tobit is in precision (larger standard errors).

**Sample description**

- 562 households were surveyed, resulting in 2,567 individual observations.
- Quasi-experimental survey in two distinct regions of northern Ecuador: Cotocachi (control group – no export flower industry) and Cayambe (treatment group – export flower industry).
- The survey was modeled after the World Bank’s LSMS.
The sectoral approach

4
module

Objective
Equation estimated: Tobit vs. CLAD

\[ y_i^* = \beta_0 + \beta_1 x_i + \beta_2 w_i + \beta_3 h_i + \epsilon_i \]

\[ y_i = \begin{cases} 0 & \text{if } y_i^* \leq 0 \\ y_i^* & \text{if } y_i^* > 0 \end{cases} \]

The model is estimated first with Tobit and then with CLAD. Data are collected at the individual level \( i \) and the sample is split in two sub-samples: men and women.

Dependent and independent variables

- The dependent variables \( y \) are the share of time spent on unpaid and paid work of men and women above 10 years of age.
- The independent variables are:
  - Household characteristics \( h \): regional location, number of children, ratio of females to males in the household, household's assets, and urban location;
  - Individuals' characteristics \( x \): age, education, age difference between husband and wife, educational difference between husband and wife, migrant dummy, and marital status;
  - Hourly wage \( w \): own and husband’s/wife’s.

Note: In the CLAD model, \( x \) is a vector of household and individual characteristics, including hourly wages.

Results

Women’s employment in the non-traditional exporting industry (cut flower industry) affects the allocation of paid and unpaid labour within the household:

- Married men in Cayambe spent twice as much time on housework as did men in Cotocachi.
- This result is related to women’s increased participation in the labour market and their increased bargaining power.
- Regardless of the growth of the cut flower industry, the paper finds that women worked more than men when both paid work and housework were included.

4.2 Data sources

For the purpose of the analysis, a survey was designed and data collected in two regions of northern Ecuador – Cotocachi and Cayambe. These two regions are about 200 kilometres apart and are similar in their cultural and ecological characteristics. A total of 562 households were surveyed, resulting in 2,567 individual observations, covering all household members above the age of 10. The survey was modeled after the World Bank's LSMS, which we reviewed in Module 1. The survey includes detailed modules on expenditures, economic activity, health, education, fertility, and credit and savings, as well as a detailed accounting of time use. Two types of time-use data were collected to capture different time allocation habits and, accordingly, two measures were calculated. The first one is an accurate indicator for activities carried out in the last 24 hours. The problem with this measure is that it may miss unusual or irregular activities. Therefore, a second indicator was designed to capture general time dedicated to housework, rest, recreation, and work over the prior week. These data are often less precise and more subject to recall error, but they have the advantage of being less burdensome for the interviewee. The 24-hour recall data were collected only for the male and female heads of the household. The weekly data were collected for all household members interviewed.

4.3 Empirical methodology

The paper uses as a framework a model in which household decisions are derived from the maximization of the weighted sum of individuals' utilities. The survey was specially designed to calculate indicators that may influence the relative decision-making power of the individuals (Browning and Chiappori, 1998) and created utility weights based on them. In this model, wages affect the labour and domestic work supply functions directly as well as indirectly through the distribution function (through, for instance, a bargaining effect). As wage opportunities change among household members, the amount of labour supplied by each of them can be affected beyond the traditional substitution and income effects of a wage change. The bargaining effect of a higher relative wage would have an impact on domestic work supply as well as on unpaid work.

To capture this effect, the author used a quasi-experimental approach. Two groups were selected – the treatment group (households in Cayambe) and the control group (households in Cotocachi) – and the survey was detailed enough to capture many differences between these two groups. However, since the experiment was applied in a real economy, there are likely to be some unobservable differences that may affect the result. When carrying out this type of analysis, the main concern is the endogeneity problem. In this context, a problem would arise if the location of...
flower production was correlated with the qualities of the workforce that might also influence time allocation decisions of working individuals. However, this seems not to be the case here as the flower producers interviewed for the study reported that the characteristics of the workforce are irrelevant to their choice of location, which is strictly guided by the unique combination of ecological characteristics in Cayambe, as well as its proximity to a regional airport. If this assumption is true, then we have a relatively clean quasi-experimental setting to study the question of interest.

The paper uses different methodologies. To compare the different time allocations between the treatment (Cayambe) and control (Cotocachi) areas, the author uses average time spent on different economic activities, household tasks, and wages received by region, gender, marital status, and labour market participation (Tables 2–5 in the paper). She then compares them to check whether there are significant differences among those averages (easily performed with the command ttest in Stata). The paper then asks the question of which characteristics are the determinants of time allocation differences between individuals’ activity choices. The paper controls for exogenous factors that might influence individuals’ decisions by including variables such as human capital, age, marital status, location, and other household and regional characteristics.

The objective of the paper is to separate the wage effect from the bargaining effect. If men in the treatment group work more at home, is it because of wage differences or because women in that region have gained bargaining power through increased access to paid work opportunities in the export-oriented industry? To capture this effect, the model includes a Cayambe dummy that, under the assumption that the two samples are comparable, captures the additional effect of the presence of the flower industry. The author also controls for own individuals’ wages and spouses’ wages to help separate the substitution effect from the bargaining effect. Indeed, while the Cayambe dummy is more likely to capture the bargaining effect as a result of changing social norms, the wages are more likely to capture the substitution effect because theoretically they have the most direct impact on utility.

From a technical point of view, three models are used to test the paper’s hypothesis. The Heckman model was originally chosen to correct for sample-selection, but a Tobit model was estimated instead. As a matter of fact, the Tobit model is preferred to the Heckman model because it allows for all the variables to be tested as possible determinants of the share of time spent by men and women on unpaid and paid work, whereas in the Heckman model, some variables need to be excluded for identification purposes without a clear theoretical guideline for making this choice. The Tobit model, however, requires errors to be normally distributed and, in this paper, the normality test fails for the sample analysed. Therefore, the paper estimates a CLAD model because it has the advantage of not requiring any assumptions about the distribution of the errors. However, the CLAD model has the disadvantage of being less precise (larger standard deviation) than the Tobit model.

### 4.4 Step-by-step explanation of how to do the estimations in Stata

The data file we will use for this application is called percsv.dta. We will use a separate do-file for each table in the paper, but you may also choose to write a single do-file that contains the commands to get all the tables and regressions included in the paper. In our case, each do-file is named after the table it generates.

#### Step 1: Generate Table 1 in the paper – Demographics of Cayambe and Cotocachi areas

To perform the first step, we use the do-file "Table 1 Final.do". First of all, we tell Stata to clear anything that was initially stored in the memory (clear). We then tell the output to scroll down without requiring user assistance (set more 1), and we define the line size in the Stata do-file to have 132 characters (set linesize 132).

We then define in which directory we will work (cd) and what data we will use for the analysis (use perscv).

To keep an output generated by the do-file, we will create a log file that captures all of the output from the time you open it until you close it, regardless of how much output is produced. You can later open this file with a text editor and edit the content to include it in your paper. You start a log file with the command:

```stata
log using name_of_file.log
```

and close it using the command:

```stata
log close
```

We want to create a table that compares variables across two populations, one being the treatment group (`treat==1`) and the other the
The sectoral approach

Control group (treat==0). One of the characteristics of Cayambe is that it has a larger number of migrants. For this reason, the author wants to see if Cayambe’s subpopulation has different characteristics from that of Cotocachi, and also show a separate column for the subpopulation who are migrants (migrant==1) in the treatment group. To be able to tabulate the information using the command by, we need to first sort the data (sort treat). Then the tabulation is performed using:

by treat: tab name_variable

where the name_variable is the age group (agegroup), the education level (edlevel), marital status (mar), whether the household head is a female (fhh), if they have children less than 15 years old (child), if they have children less than 6 years old (childlt6), and the relationship to the head of the household (rel2). The commands always have the same structure:

by treat: tab name_variable

The first line generates the columns “Cayambe All” and “Cotocachi”. The second/third line generates the column “Cayambe Migrants”.

Step 2: Generate Table 2 in the paper – Use of time, by gender and marital status

Next, we use the do-file “Table 2 Final.do”.101 The first command lines are as before. Now we have a new command svyset that declares a survey design for the dataset and the option strata(dominio) that indicates that dominio is the name of the variable identifying the strata:

svyset, strata(dominio)

Once Stata knows about the design of the survey via the command svyset, you can use the svy: prefix using syntax that is quite similar to the non-survey versions of the commands. For example, if you use the command svy: regress, it is like using a regular command regress, but it uses the information you have provided about the survey design and does the computations taking this information into consideration. In this application, we use svy: mean

svy: mean farm pdwork comwork hwork24 totwk recre percare tottime if treat==1, over(treat)

We use the command svy: mean with the option over (treat) to get the means for each variable (farm, pdwork, comwork, hwork24, totwk, recre, percare, and tottime) for each subgroup in the variable treat (here our two distinct populations in Ecuador). The variables are the ones in Table 2 of the paper and represent different activities performed by the individual. The analysis is first done for all men (sex==1), and then only for married men (if sex==1 & mar<=2) and non-married men (if sex==1 & mar>2). Replacing sex by sex==2, we obtain the same tables as before, but this time for women.

We want to know if the estimated means are different across the two populations. For this purpose, we use the command lincom (in this case for all men, time spent in farm work):

lincom [farm]1 - [farm]0

count if farm==. & sex==1 & treat==1

count if farm==. & sex==1 & treat==0

lincom computes point estimates, standard errors, t or z statistics, p-values, and confidence intervals for linear combinations of coefficients (here the difference in the mean time spent by all men in farm activities in the two populations) after any estimation command (in this case, lincom will consider the calculated coefficients with the command svy: mean).

The command count provides the number of observations in the category indicated by the expression in if.

A similar procedure is then implemented to see if there are differences in the mean across gender groups and not only across populations. For instance, to display the means for men and women living in Cayambe you need to run the following command:

svy: mean farm pdwork comwork hwork24 totwk recre percare tottime if treat==1, over(sex)

lincom [farm]Male - [farm]Female

The bottom of Table 2 in the paper shows the ratio of men’s to women’s time in each activity. These numbers can be obtained from the means displayed with the commands above.

Step 3: Generate Table 3 in the paper – Time spent performing household tasks, by gender, marital status, and labour market participation

Here we turn to the do-file “Table 3 Final.do”. We want to estimate if there is a significant difference (ttest) in the time spent on household tasks (hwork24) conditional on the spouse work-
The sectoral approach

The do-file first looks at households where the head is a man (sex==1) who is married or in a free union (mar<=2). The file first looks at men who work and men who do not work (not used in the table because of too few observations) as well as men who work in the flower sector in the treatment region compared to men who work in any other sector in the control region. For each of them, the do-file also shows the results based on whether the spouse works (spwks==1) or not (spwks==0).

The command `ttest` is then used to compare the means between the two regions.

*Men who work

```
ttest hwork24 if sex==1 & mar<=2 & rel==1 & wklast==1, by(treat)
ttest hwork24 if sex==1 & mar<=2 & rel==1 & wklast==1 & spwks==1, by(treat)
ttest hwork24 if sex==1 & mar<=2 & rel==1 & wklast==1 & spwks==0, by(treat)
```

*Men who don’t work -- not used in table because too few observations available

```
ttest hwork24 if sex==1 & mar<=2 & rel==1 & wklast==0, by(treat)
ttest hwork24 if sex==1 & mar<=2 & rel==1 & wklast==0 & spwks==1, by(treat)
ttest hwork24 if sex==1 & mar<=2 & rel==1 & wklast==0 & spwks==0, by(treat)
```

*Men who work in flowers in treat compared to men who work in other sectors in control

```
drop if wkflowrs==0 & treat==1
```

```
ttest hwork24 if sex==1 & mar<=2 & rel==1 & wklast==1, by(treat)
ttest hwork24 if sex==1 & mar<=2 & rel==1 & wklast==1 & spwks==0, by(treat)
```

The procedure is repeated for women to identify whether the same differences apply in the time spent on household chores across regions. Specifically, we look at households where the head is a woman as well as where the head is a man. The commands and syntax are similar to the previous ones and for this reason we do not specify them here.

**Step 4: Generate Table 4 in the paper – Wages by gender, marital status, and work type**

We use the do-file “Table 4 Final.do”. The command we use here is `summarize (sum)`. It provides mean, standard deviation, minimum, maximum, and number of observations for a variable. We can ask for more detailed information using `summarize, detail (sum, d)`, which provides us with additional statistics, including skewness, kurtosis, the four smallest and four largest values, and various percentiles.

The variable we are interested in describing is wage (jperhr) and we want to display it by region (treat), marital status (mar), gender (sex), and for those who work in the flower industry versus those who work in all other sectors (wkflowrs).

For example, if we want the summary statistics for wages in Cayambe for women who work in the flower sector and are married, the command line is:

```
sum jperhr if treat==1 & sex==2 & wkflowrs==1 & mar<=2, d
```

**Step 5: Generate Table 5 in the paper – Average hours per week spent performing main activities, by gender and marital status**

We now turn to the do-file “Table 5 Final.do”, which creates the table with the number of hours spent on different activities by gender and marital status for the two regions, using the time variable recorded on a weekly basis instead of the 24-hour measure of activities. The variables of interest here are paid work (hrsw), housework (hrsh), recreation (hrsr), and sleeping time (hrsl).

The commands used here (svy: mean and lincom) are the same as those in step 2 “Table 2 Final.do”.

\[ \text{ttest hwork24 if sex==1 & mar<=2 & rel==1 & wklast==1 & spwks==0, by(treat)} \]
Step 6: Generate Table 6 in the paper – CLAD and Tobit estimates of men’s share of time performing housework and paid work (dependent variable: individual’s share of housework)

We use the do-file “Table 6 Final.do”, where we estimate the Tobit and CLAD models of the determinants of time spent on household activities for men only.

We first run the Tobit estimation for the share of time that the individuals spend on housework. To generate a Tobit model in Stata, we use the command `tobit`, followed by the outcome variable and the predictors. We also can specify the lower limit and/or upper limit of the outcome variable. While the lower limit is specified in parentheses after `ll`, the upper limit is specified in parentheses after `ul`. A Tobit model can be used to predict an outcome that is censored from above, from below, or both.

```
tobit hwsh2w age age2 educ married widdiv sucre1 hhsize numchil ratiofm assets urban treat migrant if sex==1 & age>=10, ll(0)
```

Here `hwsh2w` is the share of unpaid labour and is explained by a list of individual characteristics, household characteristics, the individual’s own wage, and the dummy variable for the Cayambe location (`treat`). The analysis is performed only for males who are at least 10 years old. The share cannot be lower than zero, and therefore the lower limit `ll(0)` is specified.

A result not shown in the paper is the test for normality for the residuals of the Tobit model. As mentioned before, the Tobit model requires residuals to be normally distributed. The do-file performs this test by running a Tobit regression, predicting the estimated values of the dependent variable and generating the residuals (`gen res=hwsh2w-yhat`). We can then use the command `sktest` to perform a test for normality based on skewness and another one based on kurtosis, and then combine the two tests into an overall test statistic. The residuals fail the normality test, suggesting that using a Tobit model may not be the best approach in this case.

The author then runs CLAD estimations for housework (`hwsh2w`) and paid work (`pdsh2w`) for all men and married household heads. This generates columns 2–5 in Table 6 of the paper. One possibility to run the estimation is to use the command `clad`. This is not an original Stata command and therefore you need to install it typing `net install sg153.pkg`. This will provide you with the programme for estimating Powell’s (1984) CLAD model and obtaining bootstrap estimates of its sampling variance. The CLAD estimator is a generalization of the least absolute deviations estimator, which is implemented in Stata in the command `qreg`. This programme sidesteps the issue of programming analytical standard errors and provides instead bootstrapped estimates of the sampling variance. See Newton et al. (2000) for details on how to write the command line.

At the time of writing the paper, the command `clad` did not exist; the author therefore had to write an alternative routine in which she reproduced the algorithm explained in Chay and Powell (2001). We have left that routine for the interested reader in the do-file. Fortunately, today we have the command `clad` that simplifies the implementation of the procedure. The command line is the following:

```
clad hwsh2w age age2 educ married widdiv sucre1 hhsize numchil ratiofm assets urban treat migrant if sex==1 & age>=10, ll(0)
```

The same procedure is then repeated to generate the other columns in Table 6 of the paper.

Step 7: Generate Table 7 in the paper – CLAD and Tobit estimates of women’s share of time performing housework and paid work (dependent variable: individual’s share of housework)

Finally, we use the do-file “Table 7 Final.do”. This estimates the Tobit and CLAD models of the determinants of time spent on household activities for women only.

The commands and syntax used here are the same as those used in the previous table, the only difference being that we focus on females (`sex==2`) instead of males (`sex==1`).

4.5 Discussion of findings and limitations of the analysis

The main findings of the paper analysed in this module show that increased participation of women in the labour market has a bearing on household labour allocation. Married men in the treatment group spend double the time on housework compared to men in the control group, and this is clearly related to women’s increased participation in the labour force because of the increase in employment opportunities in the non-traditional exporting sector. Women in the
treatment group, especially married women, do less housework than those in the control group. The author also controls for other determinants of time use, including household characteristics, such as the ratio of female to male members, and social characteristics, such as marital status and the attitude towards women in society. Overall, the gender impact of the growth of the cut flower industry is perceived as positive based on the author's idea that the increase in the participation of women in the labour market itself leads to cultural changes. One of the issues the study does not consider is the possibility that some of the housework is transferred to the children within the household, which might leave them with little time to go to school or to play.

From the technical point of view, most of the problems with the quasi-experimental approach arise from self-selection and sample selection, as well as the comparability issues between the treatment and control group. Quasi-experimental estimates of effects are subject to contamination by confounding variables. The lack of random assignment in the quasi-experimental design method may allow studies to be more feasible, but this also poses many challenges for the researcher in terms of internal validity. Newman (2002) makes all possible efforts to find two comparable populations and exploits the fact that the location choice for flower firms seems to be unrelated to the characteristics of the labour force and only depends on agronomic conditions and available export infrastructure.

5 Conclusions

This module used the sectoral approach to examine the effects of trade on gender. In other words, we looked at the impact of trade on women engaged in particular sectors and industries of the market as well as the non-market economy instead of looking at the individual-level (microeconomic) and aggregate-level (macroeconomic) implications of trade on gender-related outcomes. When data are available, the sectoral approach allows us to look at the shifts in a country's production and export and import structure patterns and if these have translated into changes in women's economic and/or social status. For instance, the economic reforms implemented by Ecuador in the 1990s contributed to the development of non-traditional agricultural exporting sectors, most notably the cut flower industry. Being an industry that is traditionally female-intensive, there have been relevant gender repercussions, particularly changes in the intra-household allocation of time and tasks, which are the object of the paper reviewed in this module.

When deciding to adopt a sectoral approach to the study of trade and gender, it is thus important to conduct an ex-ante assessment of the concentration of women workers at the sectoral level to understand the gender patterns of employment as well as the relevance of horizontal gender segregation for the economy at hand. The impact of trade on gender at the sectoral level will depend both on the kind of structural transformations resulting from trade and on where women workers are most concentrated. For instance, if trade causes an expansion of a female-labour-intensive sector, women wage workers employed in that sector could potentially gain by, for instance, experiencing an increase in their wages; conversely, if trade causes a contraction of a female-intensive sector, women employed in that sector could be adversely affected by being displaced. Furthermore, women workers could be concentrated in non-tradable sectors, most notably non-tradable services such as health and education, and thus be unexposed to trade-related changes in the structural composition of the economy. In this case, women in the non-tradable sector could still be indirectly affected when, for example, trade results in an increasing specialization of the country towards the production of non-tradable services. One of the causes of changes in the structural composition of the economy could be a trade shock. For example, in the paper by Newman (2002), the shock is caused by economic reforms, including trade reforms, that have led to a boom in the growth of non-traditional agricultural exports, specifically cut flowers, and thus to the growth of the cut-flower industry that is traditionally a female-labour-intensive sector.

As regards the method of analysis, this module presented a series of censored and truncated regression models that are often useful to estimate empirical specifications of women's participation in the labour market. In the hands-on application, we looked at the difference in the allocation of housework responsibilities between men and women, comparing one area where there are work opportunities in the exporting sector with another where those opportunities do not exist. The technique adopted in the paper analysed is the CLAD, which offers many statistical advantages compared to the most commonly used Heckman and Tobit models. However, as for the choice of data, the most appropriate econometric tool you should use depends on the purpose of your investigation. As already
mentioned, the CLAD is one of many models that have been applied in sectoral studies. Other authors have also looked at the effect of trade on gender outcomes at the sectoral level using different methodologies. For instance, we have seen how the analysis of global value chains has been implemented to explore the relationship between trade and gender. Moreover, Nicita and Razzaz (2003) explore whether an increase in textile and apparel exports in Madagascar benefits the poor and examine its effect across gender groups by first using a propensity score. Depetris Chauvin and Porto (2011) use a methodology akin to the microeconomic approach. All these studies provide an ex-post assessment of the impact of trade on gender, but you can also decide to use the sectoral approach to carry out an ex-ante analysis. In this case, you can employ the simulation tools described in Module 1 of this volume.
REFERENCES


