Readme file for the MATLAB codes of A. Fernandez and F. Meza "Informal Employment and Business Cycles in Emerging Economies: The Case of Mexico", forthcoming in the Review of Economic Dynamics

July 2, 2014

This file contains the MATLAB (v. R2012a) codes for the paper in two folders: called "Empirics" and "Model". The former contains the codes and data used to generate the empirical results of the paper: Figures 1 and 2, and Tables 1 and 2. The latter contains the codes to replicate the model-related results of the paper: Figures 3 to 8, and Tables 4 to 6. In what follows we describe in detail the material in each of these two subfolders.

1 Folder "Empirics"

The folder "Empirics" has four subfolders with which the reader can replicate all the empirics-related results of the paper: Figures 1 and 2, and Tables 1 and 2. We describe next the content of each subfolder.

1.1 Figure 1

The codes in this subsection generate the two panels of Figure 1. The folder consists of a main script, "RollingCorrelations.m" which computes the rolling correlations between output and the measures of informal employment. The folder also contains two .mat databases, "RawData.mat" and "RawData2000_2010.mat" which are the data of the first and second sample, respectively (see also the EXCEL file with the raw data used in the paper). The folder also contains the auxiliary function "hpfilter.m" which filters the data using a Hodrick-Prescott filter.

1.2 Figure 2

This folder contains the two same '.mat' files that we mentioned in Figure 1, as well as the function "hpfilter.m". In addition to these files, the folder contains
two main scripts, "QuarterlyData.m" and "QuarterlyData2000_2010.m" which compute the rolling correlations, plot and format the figures and highlight the recession dates presented in Figure 2 of the paper. The auxiliary function "ini-ciofin2.m" computes the first and last observations of a dataset. The rolling correlations are computed using 21 points, and the interval confidence correspond to a 95% confidence bound.

1.3 Table 1

The folder Table 1 contains two subfolders, each of those subfolders compute the moments shown in the table, for each country. The structure of each subfolder is the same. Each folder contains a data file (i.e. DataCanada2.mat and DataReal-GMM2.mat), it also contains a general script "GMM_General_Script.m" which generates the GMM estimates of the moments shown in the table using all the other "m" files as auxiliary functions. Each folder also contains one .mat (GMM_results_Canada.mat, and Moms_Mexico_Table1 in each subfolder) containing the estimates and standard errors of the second moments.

1.4 Table 2

The folder Table 2 is very similar in structure to the folder Table 1. There is one subfolder for each sample. Each subfolder file has the files that reproduce the columns of the Table. And each one contains a .mat file with the raw data, a general script (each of them starts with the name GMM_GeneralScript) and many auxiliary functions. The third folder, called "Codes to generate conditional series.m" contains the scripts (FirstSampleConditionalSeries.m and SecondSampleConditionalSeries.m) with which the samples, conditioned on the phase of the cycle, are generated for columns 5 to 8 in the Table.

1.5 Data readme file

Next to the previous folders the user will find readmedatafernandezmeza11june2014.xls. This file contains the data used to generate the empirical results in the paper. The worsheet 'readme' describes the data used in Figure 1, Figure 2, Table 1 and Table 2, in that order. Most importantly, the data used in Figure 1 is located in worksheet 'Figure 1', the data used in Figure 2 is located in worksheet 'Figure 2', etc.

For each Figure/Table we specify which sample we are using, either 1987/01-2003/02 or 2000/02-2010/04. We work with a quarterly frequency.

For each variable used in the empirical part of the paper we specify its acronym in the worksheets, how each variable is measured (i.e "informal employment 4" is self-employment), the original data source -which is mostly INEGI, the official Mexican statistical agency-, and the name of the variable in Spanish, in the original data source. We provide detail on the empirical series loaded into the MATLAB codes to produce the Figures/Tables. We explain the steps to construct some of the series.
Most importantly, we specify which data file (a .mat file) is loaded into which code (an .m file) to produce each Figure/Table. The data presented in the worksheets is the data that appears in the .mat files.

2 Folded "Model"

The folder "Model" has four subfolders with which the reader can replicate all the model-related results of the paper: Figures 3 to 8, and Tables 4 to 6. We describe next the content of each subfolder.

2.1 Subfolder "Benchmark & Ext. 1"

The codes in this subfolder reproduce the main results of the paper using the benchmark model and those in the first extension considered in the robustness section. The user should open and run the code called "benchmark_model_run.m". The user must specify the options at the beginning that she wants in order to replicate the Tables/Figures of interest. These options allow the reader to replicate the second moments from the data and benchmark model in Table 4 (columns 1 and 3) and those in the first robustness presented in Table 5. It also permits replication of Figures 3 to 8 in the paper. The additional ".m" files are auxiliary routines that solve the model and compute the steady state. A header in each routine explains the purpose of it.

Since the model is solved using the routines by Schmitt-Grohe and Uribe (2003) some of these routines are original to that paper. Importantly, the codes use the Symbolic Math Toolbox of Matlab. Without this toolkit the routines will not run. The ".fig" files already contain the figures that the codes generate. Finally, the ".mat" files are used as input by some of the codes to run.\footnote{For some Matlab/computer configurations you need to make this change in line 375 of code benchmark_model.m: Instead of \( f = \text{subs}(f, [x,yy,yp,yp], \exp([x,yy,yp,yp])); \)
write \( f = \text{subs}(f, [x,yy,yp,yp]', \exp([x,yy,yp,yp]')); \)
You need to make a similar modification if you run the other experiments. For example, when running Noinformality_run.m, you need to make a similar change in Noinformality.m.}

2.2 Subfolder "Empirical Serial Correlation (Fig. 3)"

The code "RollingCorrelations.m" in this subfolder reproduces the empirical serial correlation (and its confidence band) between cyclical output and informal employment (using self-employment as a proxy). This is later used as input in the routine benchmark_model_run.m where the empirical correlation is compared to the model-based counterpart.

2.3 Subfolder "Extensions 2 - 5"

This subfolder contains the routines needed to replicate the results in the extensions 2-5 of the robustness Section of the paper presented in Table 6. To replicate
the results of each of these extensions the user should run the routine "Extension_X_run.m" where X = 2, 3, 4, 5. In order to speed up the computations, each routine calls the function "modelsolve.m" which already incorporates the (very large) matrices that solve the recursive formulation of the model. If the reader is interested in deriving these matrices a full version of the generic model used in all these extensions is given in the routine "modelsolve_extended_run.m". Finally, the EXCEL file "rwc.xlsx" presents the computations of the random walk component for each of the extensions considered in Table 6.

Warning: These routines will take several hours to run as they implement the grid search calibration. If the reader wants to see the results the folder also contains the workspaces with them for each extension.

2.4 Subfolder "No Informality Model"

This subfolder contains the routines needed to replicate the results of the model with no informal employment. In particular the routines generate the results reported in Table 4, last column, and some of the impulse response functions in Figure 7. The user should run the code "Noinformality_run.m".