

Estimating the effect of US Food Aid on Civil Conflict

Lawrence Chan

December 11th, 2017

Abstract

We duplicate the results of Nunn and Qian, who use lagged wheat production as an instrument for US food aid to conclude that US food aid increases the chance of civil (intrastate) conflict. In addition, we test for heteroscedasticity, and find that their assumption of homoscedasticity for reporting p -values is valid. Finally, we test for endogeneity and the validity of the overidentification assumption in their two-stage least squares regression, and find evidence of endogenous variables in their OLS regression and no evidence that the overidentifying assumption is invalid. This offers additional support for the validity of their 2SLS model and their results.

1 Introduction

There has been much debate about the effects of food aid from developed countries toward the developing world. Nunn and Qian's 2014 Paper, *US Food Aid and Civil Conflict*, uses a 2SLS regression (using shocks in the US agricultural market as an instrument to aid) and several robustness checks to conclude that US food aid increases the incidence and duration of civil conflicts.

In this work, we duplicate their analysis (with minor differences in control variables) and then perform additional tests for heteroscedasticity, endogeneity, and underidentification. We find no evidence for any of these, providing more support for the validity of their results.

2 Description of Dataset and Preliminary Analysis

Before jumping in to the discussion of Nunn and Qian's main results, we describe the datasets being used and provide some preliminary analysis.

The two main datasets are the UCDP/PRIO Armed Conflict Dataset and Food and Agriculture Organizations (FAO) FAOSTAT database. The dependent variables include the prevalence of war, defined as any armed conflict leading to more than 25 deaths, as well as civil and interstate war. The authors use US wheat aid as a proxy for US food aid, and wheat production from the previous year as an instrument for . We used the merged data set provided by the authors¹, which contains 4089 observations across 125 OECD countries over the period of 36 years from 1971 to 2006. (The actual dataset contains quite a few more rows, as their US wheat production data goes back further, to 1950.)

Summary statistics for the key variables are provided in table 2. Some of the results provided are different than those reported in Nunn and Qian's paper, as we report all observations in the dataset and not just those included in our final regression. Any conflict is an indicator variable for whether or not there has been a reported violent conflict with 25+ deaths during that year in that country, while the intrastate and interstate conflict variables are indicators for reported conflicts with 25+ deaths between the government and one or more rebel groups and between two states, respectively. Note that there are conflicts that do not fall into either category. The wheat aid and lagged US wheat production variables are reported in thousands of metric tonnes.

Before duplicating the analysis performed in the paper, it is helpful to

¹Available at the AER website here: <https://www.aeaweb.org/articles?id=10.1257/aer.104.6.1630>

Variable	Observations	Min	Max	Median	Mean	Standard Deviation
Any Conflict	5663	0	1	0	0.1898	0.3922
Intrastate Conflict	5663	0	1	0	0.1508	0.3579
Interstate Conflict	5663	0	1	0	0.0267	0.1611
US Wheat Aid (1000's of MT)	4389	0	1958	0	26.35	114.00
Lagged US Wheat Production (1000's of MT)	7112	25506	75813	51166	49529	14840.85
Received US Wheat Aid in Given Year	4389	0	1	0	0.3615	0.4805

Table 1: Summary statistics for the main variables of interest in the analysis.

3 Duplication of Regression Results

3.1 OLS

We start by performing an OLS regression with model

$$C_{it} = \beta F_{it} + \mathbf{x}_{it}\gamma + \delta_r t + \psi_i + \nu_{it},$$

where C_{it} is an indicator for the presence of conflict in country i during year t , F_{it} is the quantity of wheat aid shipped from the US to that country during that year, \mathbf{x}_{it} is a length K vector of containing information about the country during that year, $\delta_r t$ denotes region specific time trends for the region r the country is in, and ψ_i denotes country specific effects. In total, as there are 125 countries in six regions in the dataset, there are $132 + K$ variables to estimate.

, we consider

3.2 2SLS

In addition to the model above, we also perform a 2SLS regression with US wheat production during the previous year as an instrument for food aid. The model used is:

$$\begin{aligned} F_{it} &= \alpha P_{t-1} + \mathbf{x}_{it}\gamma + \delta_r t + \psi_i + \epsilon_{it} \\ C_{it} &= \beta F_{it} + \mathbf{x}_{it}\gamma + \delta_r t + \psi_i + \nu_{it} \end{aligned}$$

Where P_{t-1} is the US production of wheat in the preceding year.

4 Testing for Heteroscedasticity

5 Testing for Endogeneity and Overidentification Assumption

6 Conclusion and Further Work

Appendix A