Title
Essays in Development Economics and Trade

Author
Vivalt, Eva

Publication Date
2011

Peer reviewed|Thesis/dissertation

Essays in Development Economics and Trade

by

Eva Vivalt

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Edward Miguel, Chair
Professor Pierre-Olivier Gourinchas
Professor Ernesto Dal Bó

Spring 2011

Essays in Development Economics and Trade

Abstract

Essays in Development Economics and Trade

by

Eva Vivalt

Doctor of Philosophy in Economics

University of California, Berkeley

Professor Edward Miguel, Chair

Economic development has the potential to improve lives. Three issues that directly affect economic development are conflict, trade, and innovation, the subjects of the three chapters in this dissertation.

Conflict causes enormous suffering, but the study of peacekeeping is plagued by endogeneity issues. The first chapter in this dissertation uses an instrumental variables approach to estimate the effectiveness of U.N. peacekeepers at ending episodes of conflict, maintaining the peace once peace has been obtained, and preventing another episode from ever re-occurring. I find that the likelihood of being sent U.N. peacekeepers varies with temporary membership in the U.N. Security Council and exploit this variation in my estimation. This variation also suggests that the leaders of countries in conflict often do not want their country to receive peacekeepers. The results indicate that even though peacekeepers are often unwanted, they help to maintain the peace after an episode of conflict has ended and reduce the likelihood that the conflict resumes.

After peace, trade is also considered crucial to development. In the standard trade literature, more productive firms should export over less productive firms, all else equal. This premise appears in Melitz (2003), Luttmer (2007), and Eaton *et al.* (2009), among others. However, we know that developing countries often suffer from market distortions (Hsieh and Klenow, 2009). The theory behind the second chapter of this dissertation combines Hsieh and Klenow-like distortions with a Melitz-like model that accords productivity a key role. Under this model, firm productivity matters less in the decision to export in sectors with greater distortions and firms facing greater distortions exhibit less of the productivity-based "churning" and re-allocation that Melitz predicts. The implication is that trade is less beneficial to productivity in developing countries. These predictions are tested using plant-level data from Colombia.

Finally, new products have been shown to increase welfare in a few studies. One branch of the literature has focused on estimating the welfare effects of very narrow and specific new products; another has estimated elasticities of substitution across a large number of varieties and then imputed gains from the new varieties that appear in the data. However,

one might suspect that the most important innovations occurred over a much longer period of time than has been studied to date. Thus, the final chapter of my dissertation focuses on a different question. It assigns an innovation date to each good and asks the question: how would welfare be affected if one were restricted to the set of goods available at an earlier time period? I modify the methodology in Feenstra (1994) and Broda and Weinstein (2006a) to answer this question. The estimates suggest that innovations are more important to welfare than previously thought. I also find that the price index that takes varieties into consideration favoured by the literature can yield deeply misleading results.

Contents

List of Figures

List of Tables

Acknowledgments

I am very grateful for the help that I received along the way to completing this dissertation.

This includes useful feedback from my primary advisor, Ted Miguel, as well as from Pierre-Olivier Gourinchas, Ernesto Dal Bó, David Card, Patrick Kline, and Frederico Finan.

I am also grateful for the support from friends, including Diana Chisholm, James Huggins and Vinci Chow. Finally, I thank Lady Gaga, without whose music writing this would have been more difficult.

Introduction

Economic development seems to improve human welfare, but it can be stymied in numerous ways. This dissertation focuses on three major issues that affect economic development: conflict, trade, and innovation.

Conflict is a leading cause of human suffering. Even apart from battle deaths and other traumas, conflict has many indirect effects, for example, on health and nutrition, and while it endures economic development is unlikely. For its part, trade has often been seen as more important than aid in encouraging economic development. The sheer volumes of money that enter a country through trade often swamp aid dollars, without causing the same concerns about incentives. Finally, there is reason to believe that innovations make a large contribution to human welfare. Apart from potentially resulting in cheaper inputs to production processes, new goods may be valuable simply by being of better quality or otherwise having characteristics that people desire.

The next three subsections provide an overview of my research in each of these areas.

Conflict: Peacekeepers Help, Governments Hinder

Much has been written on the initial causes of conflict (*e.g.* Fearon 1995; Collier and Hoeffler 1998; Powell 2002; Fearon and Laitin 2003). Yet the ending of conflicts or the post-conflict maintenance of peace is just as important, and relatively little has been written about these topics. Evaluating the effectiveness of peacekeeping has historically been difficult since peacekeepers are not randomly sent to episodes of conflict. Thus, there are concerns of endogeneity. This paper uses an instrumental variable to solve this problem and, in particular, answer three questions: whether peacekeeping helps to end episodes of conflict, whether peacekeeping helps to extend the duration of peace after an episode of conflict has ended, and whether peacekeeping helps prevent another episode of the same conflict from ever re-occuring.

Whether peacekeepers help or hurt the prospects for peace is theoretically ambiguous. Peacekeepers have been hypothesized to help maintain peace by a few mechanisms: increasing the cost of fighting through threatening the use of force and offering incentives to disarm; decreasing uncertainty about the actions and intentions of each party and making contracts more credible through monitoring and (more limited) enforcement; preventing iso-

lated or small groups of actors from acting as "spoilers" as well as preventing accidental re-engagements by providing a neutral physical buffer zone between parties; and decreasing political oppression or extraction from one side in those conflicts in which this is relevant (Doyle and Sambanis 2000; Fortna 2004; Mattes and Savun 2010). On the other hand, when one side has a more decisive victory, peace is more likely to last (Hensel, 1994); if peacekeeping resulted in conflicts being artificially cut short, there may also be more uncertainty about who would win were the conflict to resume, and uncertainty may lead to conflict (Fearon 1995; Slantchev 2004; Fey and Ramsay 2010); further, there is the question of to what extent the initial causes of conflict remain untouched by peacekeeping and, without addressing the root causes of the conflict, peacekeepers may only temporarily reduce violence.

My research focuses on U.N. peacekeeping. U.N. peacekeeping operations constitute the majority of the world's peacekeeping operations. The U.N. Security Council decides when an operation is to be approved; it is comprised of five permanent members and ten temporary members that serve for staggered two year terms, with five new temporary members rotating in each year. The timing of the assignment of countries to the U.N. Security Council provides plausibly exogenous variation in the likelihood of being sent peacekeepers. It is this assignment that I exploit to construct an instrument for receiving peacekeepers, which is otherwise endogenous to the conflicts. In the course of my research, I perform a few robustness checks to mitigate the concern that U.N. Security Council members are somehow special.

Although case studies suggest there is a lot of heterogeneity in the success of peacekeeping, I only estimate the overall effectiveness of U.N. peacekeeping along several dimensions, due to my identification strategy. While study of peacekeeping is hardly new (for older works, see *e.g.* Haas, Butterworth and Nye 1972; Wilkenfeld and Brecher 1984; Diehl, Reifschneider and Hensel 1996), few works have considered the potential endogeneity of where peacekeepers get sent. Doyle and Sambanis (2000) note the concern of endogeneity but do not find significant evidence of it in their dataset; Fortna (2004) does find evidence that endogeneity causes a problem and explicitly focuses on dealing with this issue with a study that adds characteristics of conflicts as controls to mitigate this problem; Gilligan and Sergenti (2008) use matching, though matching estimators tend to perform poorly in relatively small datasets (*e.g.* Abadie and Imbens 2006; Busso, DiNardo and McCrary 2009). My research on peacekeeping builds on this literature by using an instrumental variables and control function approach to address the endogeneity concerns. The control function approach, in particular, allows me to characterize selection. Through these methods I discover that leaders of a country in conflict often do not want peacekeepers, a new empirical finding. The results suggest that peacekeepers do often help prolong the peace once peace has been obtained and help lower the chance that another episode of the same conflict will ever re-occur. U.N. peacekeepers do not, however, shorten the duration of the episode of conflict itself. This is perhaps not surprising since peacekeepers are often sent only once there has been a break in the fighting. In sum, it seems peacekeepers can help countries escape from conflict - if governments will let them.

Trade Distortions in Developing Countries

In the standard trade literature, more productive firms should export over less productive firms, all else equal (*e.g.* Melitz (2003), Luttmer (2007), and Eaton *et al.* (2009)). However, we know that in developing countries, distortions often override the effects of productivity. Perhaps more famously, Hsieh and Klenow (2009) showed that the most productive firms in China and India were using relatively little labour and capital due to distortions affecting their profitability. My research combines Hsieh and Klenow-like distortions with a Melitz-like model that accords productivity a key role. The key implication of the model is that opening to trade does not always increase firms' productivity. Whether or not distortions do affect the decision to export is tested using data from Colombia.

The trade literature has largely focused on openness to trade as a factor that might enhance productivity. Yet results on this subject are mixed. Many have found that exporters are more efficient in general than non-exporters (*e.g.* Aw and Hwang, 1995; Griliches and Regev, 1995), but Clerides *et al.* (1998) and Bernard and Jensen (1999) suggested that this is largely because more productive firms tend to self-select into exporting, and other studies found varied evidence of the productivity-based "churning" of firms that Melitz predicts (*e.g.* Aw *et al.*, 2001; Aw *et al.*, 2000).

There is in fact a theoretical interaction between distortions and openness to trade that may affect productivity and could have caused these mixed early results. In Melitz, opening to trade results in productivity gains partially due to the re-allocation of resources from less to the more productive exporters. Exporters are assumed to be relatively productive, and when an economy opens to trade, these firms gain market share while some less productive firms exit, and overall productivity rises. But in the case in which there are distortions such that some relatively less productive firms export and some relatively more productive firms do not, this effect is weakened. Melitz-like "churning" of firms is no longer occuring based on productivity, but based on productivity and distortions. In extreme cases, the result from Melitz that opening to trade increases productivity may no longer hold.

I test the main premise of the model, that distortions affect the decision to export in a similar manner as productivity, using plant-level data from Colombia. I also construct some counterfactuals that suggest how much distortions affect the productivity gains from trade in Colombia.

The results suggest that distortions can indeed encourage or discourage firms from exporting. The implication is that trade may not improve productivity in particularly distorted sectors or countries as much as has historically been assumed and, conversely, estimates of the productivity-enhancing effects of trade that are based on more distorted cases may understate the true effects of trade on productivity in less distorted sectors or countries. Further, reducing distortions would seem to be able to cause particularly large productivity gains in sectors conducting trade, a result that has important implications for policy.

Welfare Gains from Innovation

It has historically been difficult to gauge the welfare gains from new products. Two main methods have been used in the literature to try to determine the benefits of innovations. One branch of the literature has focused on gathering very detailed data about demand for a set of similar goods and using this data in a discrete choice model to estimate the gains from new products (*e.g.* Hausman, 1997; Bresnahan and Gordon, 1997). The second method obtains data across a wide set of varieties and estimates elasticities of substitution between varieties using GMM and assumptions about the utility function. These elasticities are then used to obtain an estimate of the welfare effects of the new products via effects on a price index which is explicitly constructed to take changes in varieties into account.

Both approaches have their merits, but no paper has yet estimated gains from a comprehensive, historical set of innovations. It is precisely these earlier innovations that one might think more basic or for which there might be fewer substitutes. My research extends the methods used in the literature and applies them to a unique dataset to answer the question: how would welfare be affected if one were restricted to the set of goods available at an earlier time period? My research also suggests that the most commonly used price index which takes changes in varieties into account is flawed and should only be used under certain circumstances. I highlight these circumstances and adapt my methodology accordingly. The results suggest that the gains from more historical innovations appear to be much greater than would be suggested by literature that focuses on more recent innovations.

The following chapters detail each of these research projects in turn.

Chapter 1

Peacekeepers Help, Governments Hinder

1.1 Introduction

Conflict exacts a tremendous toll on the people in the areas in which it occurs. Apart from the direct impacts of conflict, such as deaths, conflict has many indirect effects and is implicated as a major cause of enduring poverty and lack of economic growth. A large literature exists focusing on the initial causes of conflict (*e.g.* Fearon 1995; Collier and Hoeffler 1998; Powell 2002; Fearon and Laitin 2003). However, there has been less study of either the ending of conflicts or the post-conflict maintenance of peace. The historical problem with evaluating the effectiveness of peacekeeping is that peacekeepers are not randomly sent to episodes of conflict and there is thus the concern of endogeneity. This paper for the first time identifies an instrumental variable that predicts which episodes of conflict receive U.N. peacekeepers and uses it to answer three questions: whether peacekeeping helps to end episodes of conflict, whether peacekeeping helps to extend the duration of peace after an episode of conflict has ended, and whether peacekeeping helps prevent another episode of the same conflict from ever re-occuring.

While the very name "peacekeepers" suggests an effective force, whether peacekeepers contribute to or lessen the odds of peace is theoretically ambiguous. Some ways in which peacekeepers have been predicted to help maintain peace are by increasing the cost of fighting through threatening the use of force and offering incentives to disarm; decreasing uncertainty about the actions and intentions of each party and making contracts more credible through monitoring and (more limited) enforcement; preventing isolated or small groups of actors from acting as "spoilers" as well as preventing accidental re-engagements by providing a neutral physical buffer zone between parties; and decreasing political oppression or extraction from one side in those conflicts in which this is relevant (Doyle and Sambanis 2000; Fortna 2004; Mattes and Savun 2010). In contrast, Hensel (1994) finds that when one side

has a more decisive victory, peace will be more likely to last, suggesting the unhappy conclusion that if peacekeepers help to artificially end conflicts sooner but less decisively, they may actually harm the long-term prospects of peace. If peacekeeping resulted in conflicts being artificially cut short, there may also be more uncertainty about who would win were the conflict to resume, and uncertainty may lead to conflict (Fearon 1995; Slantchev 2004; Fey and Ramsay 2010). Finally, there is the question of to what extent the initial causes of conflict remain untouched by peacekeeping. For example, externally imposed peace may leave a government in power that is not perceived as legitimate, again leading to an unstable future. If these theories are correct, peacekeepers may at best serve only to tampen down violence momentarily, with peace not lasting once the peacekeepers withdraw, if they did indeed manage to help maintain peace in the first place.

In this paper I focus on U.N. peacekeeping since U.N. peacekeeping constitutes the majority of the world's peacekeeping. U.N. peacekeeping operations are the purview of the U.N. Security Council, which has five permanent members and ten temporary members that serve for staggered two year terms, with five new temporary members rotating in each year. The timing of the assignment of countries to the U.N. Security Council provides plausibly exogenous variation in the likelihood of being sent peacekeepers. While we might think that Security Council members are in some way special, I perform a few robustness checks to mitigate this concern. I then exploit this variation to construct an instrument for receiving peacekeepers, which is otherwise endogenous to the conflicts.

Although case studies suggest there is a lot of heterogeneity in peacekeeping, I will only be able to estimate the general effectiveness of U.N. peacekeeping. These estimates still provide a significant contribution to the literature. While work has been done on peacekeeping for decades (*e.g.* Haas, Butterworth and Nye 1972; Wilkenfeld and Brecher 1984; Diehl, Reifschneider and Hensel 1996), little has addressed the potential endogeneity of where peacekeepers get sent. Doyle and Sambanis (2000) pay attention to the concern of endogeneity but do not find notable evidence of it in their dataset; Fortna (2004) does find evidence of it and explicitly focuses on dealing with the endogeneity of where peacekeepers get sent with a study that adds characteristics of conflicts as controls to mitigate this problem; Gilligan and Sergenti (2008) use matching, though matching estimators tend to perform poorly in relatively small datasets (*e.g.* Abadie and Imbens 2006; Busso, DiNardo and McCrary 2009). This work builds on this literature by addressing the endogeneity concerns with an instrumental variables and control function approach. The control function approach, in particular, allows me to characterize selection. Through these methods I discover that leaders of a country in conflict often do not want peacekeepers, a new empirical finding. I provide a very simple model to explain this in a subsequent section.

The rest of this paper proceeds as follows. The next section provides background information on U.N. peacekeeping, defining what I mean by U.N. peacekeeping operations and modelling the decision-making process that results in them. I then discuss why governments may not want their country to receive peacekeepers. Following this, I describe the data, detail the identification strategy, and present the results. Finally, I discuss the strengths and

limitations of the instrumental variable and control function approaches and provide some robustness checks before concluding.

1.2 Peacekeeping and the U.N. Security Council

1.2.1 U.N. Peacekeeping Operations

U.N. peacekeepers' role and what they are allowed to do has evolved over time. Traditionally, U.N. peacekeepers acted as a buffer force, physically positioned between combatants following a ceasefire. While not great enough in numbers to prevent determined parties from attacking their opponents, the peacekeepers could prevent isolated or small groups of actors from acting as "spoilers" as well as preventing accidental engagements.[1] In this traditional role they were also sent to observe the carrying out of ceasefire agreements (*e.g.* withdrawals of troops from a specified area) to help detect violations. At the same time as peacekeeping forces are sent, other U.N. personnel typically engage in diplomatic efforts such as trying to arrange meetings between the different sides, though this activity may also occur without a peacekeeping force in the country. Peacekeeping operations are predicated on a few requirements: the host state must consent to the forces; the forces must maintain impartiality; and the forces can only use minimal force, as defined in the resolutions that established the mission. Typically, this last stipulation allows peacekeepers to use military force only in self-defense.

After the end of the Cold War, U.N. peacekeeping operations began to expand in frequency and also in scope. In 1995, U.N. Secretary-General Boutros Boutros-Ghali described this new form of peacekeeping as "peacebuilding", the "creation of a new environment" that would contribute to lasting peace. Peacebuilding activities include: disarmament, human rights protection, humanitarian aid and programs to promote economic development, and election supervision. These peacebuilding activites, however, typically follow traditional peacekeeping operations rather than acting as substitutes for them. In this paper, I will define a peacekeeping operation as one in which military observers are sent to the country in conflict. Since there are relatively few cases of each type of peacebuilding, I will not try to evaluate the effectiveness of each type of peacebuilding separately; even apart from the small sample size, there is the concern that which type of peacebuilding activities are chosen may be endogenous to the type of conflict that had occurred, prohibiting estimation of the activities' independent effect on the outcome variables. If as rigorous a method could be used to disaggregate effects by the type of conflict or the type of peacebuilding used, that would be preferable, however, there is a trade-off between disaggregation and robustness, and while papers that do attempt to distinguish between different types of peacebuilding are valuable, so too is a rigorous treatment of the basic question of whether peacekeeping helps at all.

[1] For historical descriptions of spoilers, see Cochrane 2008.

1.2.2 The U.N. Security Council

U.N. peacekeeping operations come into existence when the U.N. Security Council passes a resolution authorizing them. It is in theory possible for the U.N. General Assembly to pass such resolutions, but the Security Council is the only body with the authority to make binding decisions. Since the Security Council determines whether peacekeepers are sent to a given episode of conflict, it is necessary to understand its structure and decision-making process.

The U.N. Security Council is comprised of five permanent members - the United States, Russia, the United Kingdom, France, and China - and ten non-permanent (or temporary) members elected for two year terms, with five of these ten seats contested each year. The temporary members are chosen by regional groups. The African Group chooses three members; the Group of Latin American and Caribbean States, the Asian Group, and the Western European States and Other States Group[2] each choose two members; and the Eastern European Group chooses one member. Before 1966, there were only six temporary members, and the groups were divided differently. The Group of Latin American and Caribbean States had two members, and the Commonwealth Group, Western European Group, Eastern European Group and Middle Eastern Group each had one member. A Security Council member must be nominated by its group and then receive a two-thirds vote in the U.N. General Assembly. The regional groups try to present a "clean slate" to the General Assembly, with one nominee per seat, however, on average, the U.N. General Assembly faces approximately seven strong candidates for the five seats up for election each year. Once a temporary member has served its two year term, it is ineligible for immediate re-election. Elections take place within the three months before the start of the term on January 1. Seats are at least weakly desirable, thus it is possible that larger, more influential states that can exert more pressure within their group are nominated more frequently (Malone 2000). Indeed, Japan and Brazil are disproportionately nominated. However, apart from these countries, which have particularly strong desires to serve on the Security Council, seats are assigned on more of a rotation system. For example, the African Group abides by a rotation system under which Northern Africa and Central Africa each receive 1 seat every 2 years in alternating succession; Eastern African and Southern Africa also rotate 1 seat every 2 years; and Western Africa receives one seat every 2 years. An Arab state is elected every 2 years, alternating between being from the Asian Group and from Northern Africa. One of Denmark, Norway, Finland and Sweden gets a seat every 4 years, as does one of Canada, Australia and New Zealand. Table A-1 in the Appendix provides a list of the years that states in the conflict dataset have served on the Security Council to date.

[2]The "other states" are Canada, Australia, and New Zealand.

1.2.3 The Decision-Making Process of the U.N. Security Council

What determines where peacekeepers are sent? I discuss three factors that affect the likelihood that episodes of conflict receive peacekeepers - temporary membership on the Security Council, the Cold War, and characteristics of the episodes of conflict - and the process under which decisions are made.

The temporary members of the Security Council influence the selection of conflicts to receive peacekeepers; in particular, when a country is a temporary member of the Security Council, it is rarely sent peacekeepers. Indeed, peacekeepers have never been authorized to be sent to a country with an active conflict for the first time while that country was a temporary member of the U.N. Security Council (with "active conflict" defined as having at least 25 battle deaths in that year). In about 11.5% of our 444 episodes of conflict, the state is a temporary member of the Security Council for at least one year, and among these episodes 9.8% are sent peacekeepers relative to 15.5% of other episodes. Permutation tests suggest the difference is significant at the 20% level on the whole dataset and at the 10% level after the end of the Cold War, when peacekeepers were more frequently sent.

Since the U.S. and the U.S.S.R. supported conflict in "proxy wars" during the Cold War, and since both the U.S. and the U.S.S.R. possessed vetoes on the U.N. Security Council, as permanent members, it is not surprising that U.N. peacekeepers were rarely sent before the end of the Cold War. A list of the years peacekeepers were sent to conflicts in my dataset is included in Appendix A. Other studies have also suggested that the Security Council's likelihood of sending peacekeepers is dependent on episode characteristics such as the number of deaths caused by the conflict (Fortna 2004).

When a security issue is raised, the Security Council has several tools at its disposal to try to get the sides to negotiate a diplomatic solution. The strongest action it can take is to pass a resolution, which requires nine affirmative votes and no permanent member veto. Resolutions can authorize a peacekeeping operation (PKO) if it is felt the situation calls for one. For such a resolution to be passed, however, the state(s) in conflict must agree to the peacekeeping operation.

In sum, the U.N. Security Council's decision-making process can be modelled as in Figure 1.1.[3]

In this paper, I will exploit the fact that a country is less likely to receive peacekeepers if it is serving on the Security Council at the time of the conflict to evaluate the effectiveness of U.N. peacekeeping. The validity of this instrument will be discussed in detail in a later section.

[3]Ideally, one would also be able to look at Security Council votes and use close votes to get further plausibly exogenous variation in the likelihood of being sent peacekeepers. I cannot do this here because no votes are close; resolutions will typically not be suggested unless it is known that they will pass or unless it is meant as a form of protest in expectation of a permanent member veto. Permanent member vetoes are themselves rare.

1.3 Incentives of Governments to Refuse Peacekeepers

Why might we expect governments to refuse peacekeepers?

To build the simplest of models, suppose that if a government refuses peacekeepers, it has some probability p of "winning" and suppressing the conflict on its own. We can normalize its payoffs to winning and losing to 1 and 0, respectively, so that a government expects to receive p if it refuses peacekeepers. If it accepts peacekeepers, its chances of winning are $p + q$, but it also bears some extra costs, c, the source of which will be discussed shortly. In this toy model, the government will refuse peacekeepers when $(p + q)(1 - c) < p$, *i.e.* when q is low or c is high.

Given the finding that peacekeepers help, governments should on average believe $q \geq 0$ if they know this and believe past peacekeeping performance predicts future peacekeeping performance in their situation. Of course, a government could still believe $q < 0$ for its particular situation.

As regards c, there are a few potential costs to allowing peacekeepers. First, it could be the case that country leaders prefer not to receive peacekeepers because they fear this would damage their reputation either domestically or internationally. Domestically, allowing peacekeepers could be seen as a sign of weakness and encourage rival groups to seek more power. Reputation has also been shown to have an impact on future dyadic conflicts (Crescenzi 2007). Further, even apart from any effects on the future likelihood of conflict, an international peacekeeping effort would draw negative attention to a country and potentially hurt it economically, such as by reducing its chances of attracting foreign investment.

Alternatively, if the government is in a powerful position, it may not want peace to be obtained and conflict ended, leaving anti-government forces in the country. These forces could be a threat to later stability or to the government should the peacekeepers leave. In particular, if the government envisions its position will only weaken or remain the same relative to the other side due to the peacekeepers' presence, it may believe the conflict would resume again afterwards with it at more of a disadvantage. This explanation would require distrust that peacekeepers were effective in the long term.

Finally, it should be noted that in all of this the incentives of government leaders do not necessarily align with those of the citizens. At minimum, government leaders may not bear all the costs of fighting; in the worst case, as in Chiozza and Goemans (2004), war could even be *ex post* efficient for leaders.

There are clearly many reasons why a government may refuse peacekeepers. While this paper cannot distinguish between alternative explanations, empirically establishing whether or not governments generally find the costs of peacekeepers high enough, relative to the potential help they can offer, that they do not want peacekeepers would be a significant contribution.

1.4 Data

For this analysis, I use the UCDP/PRIO Armed Conflict version 4 (2009) and Battle Deaths version 3 (2009) datasets. Conflict is defined in the dataset as "a contested incompatibility that concerns government and/or territory where the use of armed force between parties, of which at least one is the government of a state, results in at least 25 battle-related deaths" (Gleditsch *et al.* 2002). The dataset further divides the conflicts into episodes, where an episode is said to have ended if the number of battle deaths falls below 25 for at least one year. While this rule is artificial, some such rule must be chosen. This rule does appear to set a good bar since when it is used not many conflicts flicker in and out of existence, suggesting that episodes' entry or exit from the dataset is meaningful. I use these episodes as the unit of analysis and add a variable indicating whether or not a peacekeeping operation was sent.[4] It should be noted that since peacekeepers are occasionally sent to a situation that does not quality as an episode of conflict in my dataset, those peacekeeping operations are excluded. For each episode, I also code how many years the main country in which the episode took place was a temporary member of the U.N. Security Council during that episode. Additional variables code how many years in the preceding 5 and 10 years, respectively, the country was a temporary member of the U.N. Security Council. These will serve as controls, and the reason I code them separately is because one might expect them to have opposite effects on the likelihood of being on the Security Council during an episode and therefore on receiving peacekeepers. The logic is that some countries more frequently serve on the Security Council due to size and influence, and this would be picked up by how often the country served in the last 10 years.[5] On the other hand, a country is less likely to serve if it has recently been on the Security Council.

Finally, since I am interested in the effects of U.N. peacekeeping in the absence of other military interventions, I also code whether or not the CIA or KGB were involved in the conflict, using as sources Blum (2004), Andrew and Mitrokhin (2005), and Weiner (2007), following Easterly, Satyanath and Berger (2008). I also code whether or not a permanent member of the U.N. Security Council was directly involved in the conflict, following Fortna (2004). Since no U.N. peacekeepers are sent when a permanent member of the Security Council is involved in the conflict in my dataset, I exclude these cases to make my instrument stronger.

Data on U.S. economic assistence and military aid for each country and year was collected as in Kuziemko and Werker (2006) from the Greenbook compiled by USAID. Data on GDP is from the Penn World Tables, and trade data is adapted from Barbieri, Keshk and Pollins (2008).

[4]Conflicts involving Israel are excluded from all analyses as an outlier.

[5]The reason I do not go farther back is that once I control for a country's influence 10 years ago that country's influence 20 years ago is largely irrelevant; also, if I were to go back much further I would introduce bias since some countries did not exist that many years before the start of my analysis.

1.5 Identification Strategy

As mentioned, being a temporary member of the U.N. Security Council affects the likelihood of being sent peacekeepers. The basic empirical strategy of this paper is to use this plausibly exogenous variation in the likelihood of being sent peacekeepers to evaluate the effects of peacekeeping. In my main regressions I use both two stage least squares (2SLS) and a control function approach. In each, I estimate how the likelihood of receiving peacekeepers depends on other characteristics with the following first stage:

$$P_e = \alpha + \beta_1 SC_e + \beta_2 Z_e^p + \epsilon_e \tag{1.1}$$

where P a binary variable indicating whether U.N. peacekeepers were sent to the episode, e is the episode of conflict, SC is how many years the country was a temporary member of the Security Council during that episode of conflict, and Z^p are controls, including how many years the episode of conflict lasted, the year the episode of conflict began, the low estimate of how many deaths occurred in the last year of the episode from the PRIO Battle Deaths dataset[6], and other controls depending on the specification, and ϵ is an error term. Since U.N. peacekeepers were largely sent only after the end of the Cold War, I truncate my sample to those that ended after 1980 so as to have a reasonably strong instrument; I would restrict focus to even later cases but there is clearly a trade-off.[7]

In the 2SLS regressions, (1) has to be estimated by OLS to avoid the "forbidden regression" (Hausman 1983). With the control function approach, I can choose (1) to be estimated using a probit. The generalized residuals from this first stage are then included as a control in the second stage regression. The advantage of the control function approach is that if the model is correctly specified, it is more efficient than 2SLS (Wooldridge 2007). However, if it is misspecified, it will not be consistent, whereas 2SLS would be (Angrist 2001). Choosing to estimate the first stage with OLS and then including the residual as a control in the second stage will yield the same estimates as 2SLS. I will thus continue my exposition in terms of the control function approach though both methods are used.

The second stage regressions on the duration of peace are then represented by the following equation:

$$D_e^p = \alpha + \gamma_1 P_e + \gamma_2 Z_e^p + \gamma_3 Resid_e^p + \epsilon_e \tag{1.2}$$

where D^p is the duration of peace after peace has been obtained and $Resid^p$ are the residuals. Since the residuals are estimated from (1), the second stage is bootstrapped. Including the residuals as a control both removes the endogeneity and also tests whether the endogeneity was an issue. If the residual is significant, there was selection on the variables included in

[6]Other estimates of deaths could be used, but the best guess and high estimates of deaths have more noise and the deaths in the last year of the conflict seem to have the most explanatory power.

[7]Results are comparable, however, restricting the sample to those conflicts that ended after 1989, and they are available upon request.

the first stage.

We are also interested in whether peacekeeping helps prevent conflicts from ever re-occurring. Of course, we cannot know whether a conflict will ever re-occur, so instead I look at the distribution of how many years it takes conflicts that do re-occur to re-occur. The plots in Figures 1.2 and 1.3 show that if a conflict re-occurs, it will likely re-occur within the 10 years after the end of the conflict and usually within 5 years. If the estimates obtained by running regressions specified by equation (2) show that peacekeepers increase the duration of peace by many years, this would then provide suggestive evidence that peacekeepers can help keep the peace long enough to reduce the risk that countries fall back into conflict.[8]

Finally, we are interested in whether peacekeeping can help to end conflicts. *A priori*, we may think U.N. peacekeepers would have little effect, if any, on the duration of conflict since they are generally sent only once there has been some kind of peace agreement between the relevant sides. However, it is still possible that there will be an effect since the U.N. also helps to broker these agreements when it is interested in sending peacekeepers. The second stage equation estimated here is:

$$D_e^c = \alpha + \psi_1 P_e + \phi_2 Z_e^c + \phi_3 Resid_e^c + \epsilon_e \tag{1.3}$$

where D^c is the duration of the episode of conflict, Z^c are the controls, and $Resid^c$ is the residual from the first stage regression. In the first stage, equation (1) is now estimated without the control of the duration of the episode of the conflict. Without this control, to have a reasonably strong instrument I restrict the cases included to even closer to the end of the Cold War, when more peacekeepers were sent, requiring the episodes to have ended after 1985. One would also think that the number of years a country is a temporary member of the Security Council during an episode of conflict and the duration of that episode of conflict would be mechanically positively correlated, so instead of the number of years a country is a temporary member of the Security Council during the episode I now use the percent of years the country was a temporary member of the Security Council during the episode as a control in the first stage regression.

Since the outcome variable D_e^p represents durations which may be right-censored, I estimate (2) and (3) with a Tobit in the second stage. I also repeat the analysis using a Cox hazard model in the second stage. While D_e^c may also be censored in the sense that there are some conflicts which are still going on, peacekeepers are usually sent only once fighting has ended, so whether or not peacekeepers will even be sent to these on-going conflicts is as yet unknown. Thus, I only estimate the effects of peacekeepers on those conflicts which have already ended. While this is not ideal, it ultimately will not matter since it will turn out that there is no effect here no matter what is done.

[8]While this method is indirect, there are as yet few good ways to directly deal with problems that require a probit in both the first and second stage (Wooldridge 2002, 2007). Further, treating the second stage as a probit (with a value of 1 if a conflict has not yet re-occurred) would be discarding information relative to keeping the (albeit censored) lengths of time that peace has endured to date.

The standard errors for all regressions are clustered at the country level.

1.6 Estimating the Effects of U.N. Peacekeeping

In Table 1.1, we see that the proposed instrument, the number of years a country is on the Security Council during an episode of conflict, has the expected mechanical relationship with the duration of the episode. We also see that there is a kind of time trend; countries are less likely to have been on the Security Council during an episode of conflict recently. The number of times a country was on the Security Council during the last 10 years is also strongly correlated with how long the country is on the Security Council during an episode of conflict, as predicted. Finally, we can observe the same relationship found in Kuziemko and Werker (2006): non-military aid is correlated with temporary Security Council membership. No other covariates, notably ln trade with the U.S. or GDP per capita, are correlated with the number of years a country is on the Security Council during an episode of conflict.

When one does not consider the endogeneity of peacekeeping, U.N. peacekeeping does not seem to have a significant effect on the duration of peace, regardless of whether we use simple OLS, a Tobit model or perhaps most appropriately a Cox hazard model (Table 1.2, Table A-2 in the Appendix, and Table 1.3). Yet using the number of years a country is on the Security Council during an episode as an instrument, we see evidence that peacekeeping increases the duration of peace after an episode of conflict has ended (Table 1.4). Table 1.5 and, in the Appendix, Table A-3 estimate the effects of U.N. peacekeepers on the duration of peace using a control function approach. Since the Cox hazard model is estimating the effects on when peace fails and hazard ratios are reported, values below 1 represent a reduction in the likelihood that peace fails, whereas in the Tobit negative values represent fewer years of peace. In each regression, the generalized residual is always significant, again indicating that endogeneity is a concern.

The regressions in Tables 1.4, 1.5 and A-3 illustrate three things.

First, they reinforce the idea that endogeneity is a problem in peacekeeping, as found by Fortna (2004) but not by Doyle and Sambanis (2000).

Second, the control function approach in particular allows me to characterize the selection. A negative coefficient on the residual in the Tobit model suggests that the cases which are more likely to be selected to receive peacekeepers are also the cases that have lower values of the dependent variable, duration of peace; similarly, the coefficient on the residual in the Cox hazard model suggests that the cases which are more likely to be selected are those in which peace is more likely to fail. While it has previously been noted that cases which are "harder" along certain dimensions seem more likely to receive peacekeepers, the control function approach provides a direct test of this hypothesis. The magnitudes of the effects of peacekeepers are also notably larger than found in Fortna (2004).

Finally, and perhaps most interestingly, one implication of these results is that the leaders of countries that have fallen into conflict appear to want to block peace. This suggests

that the fact that conflict occurs in the first place is partially due to the leaders preferring conflict; in other words, going back to the model, they face a low q or high c. It should also be noted that while conflicts have many negative outcomes, the leaders usually do not fully bear all the costs, such as the deaths, themselves, and thus their incentives regarding allowing peacekeepers are not likely to be aligned with those of their citizens.

The empirical finding that governments do not want peacekeepers opens up many new avenues for further research. Are governments more willing to receive peacekeepers if they are losing the conflict? What are the true costs of peacekeeping to the recipient country, who bears those costs within the country, and how can they be lowered? Why might a government fear it will be in a worse position relative to the other side after the peacekeepers leave, and it is reasonable to be concerned about this if peacekeepers are in fact effective at creating lasting peace? Doyle and Sambanis (2000) found that peacekeepers also encourage democratization; is part of why governments want to avoid peacekeepers a lessened ability to extract rents?

Whether peacekeepers have an effect on the duration of the episode of conflict itself was checked, but no robust effects were found; tables showing the results of these regressions are included in an appendix (Tables A-4 and A-5 in the Appendix). It is possible that if we were to break the cases down into those which received peacekeepers earlier or later in their conflicts, we might be able to tease out an effect on a subset of the cases, but the main advantage of this paper is its use of an instrumental variable and control function approach to deal with the endogeneity of where peacekeepers are sent and breaking these cases into groups would reduce the sample size to the point where these methods could not be very useful. Further, there may be another hidden layer of endogeneity in terms of which conflicts get sent peacekeepers sooner as opposed to later.

Finally, while the emphasis of this paper is on the causal relationships, a couple of the control variables have interesting coefficients. Looking at Tables 1.4, 1.5 and A-3, military aid appears to have a strong positive relationship with the duration of peace after an episode has ended, while non-military aid has a negative correlation with it. It should be emphasized again that these results are not causal since it is possible, for example, that countries with worse conflicts received more aid and these worse conflicts were also more likely to resume independent of the aid received. It should also be observed that military aid has a negative relationship with the duration of an episode of conflict while non-military aid has a positive relationship with this duration (Tables A-4 and A-5). Although these are, again, only correlations, they accord with intuitions.

1.7 Limitations of the IV and Robustness Checks

It is possible that which countries get elected to the Security Council is dependent on country influence and this has an effect on peace itself. I thus use an alternative specification which includes not just how many years the country was a temporary member during the

episode of conflict but also how many years it was a temporary member leading up to the episode of conflict. The average episode in my dataset lasts 3.6 years, and I look at the 5 and 10 years prior to an episode rather than just the 1-2 years prior since only 5 countries are newly elected onto the Security Council each year and countries are generally ineligible for re-election immediately following a term on the Security Council.

This specification helps quell fears that the temporary members of the U.N. Security Council are special in some way that is directly correlated with the outcome variables. It further helps that while the assignment of seats is not random, states exogenously leave the Council since they cannot immediately be re-elected. Also, once a less prominent country has served, it is unlikely to serve again until most of the other less prominent countries in its group have served. Thus, if a country in conflict wanted to get on the U.N. Security Council to prevent peacekeepers from being sent, it would have difficulty in doing so if it had ever served before, and the possibility of later being in conflict and needing to be on the Security Council is unlikely to have entered into its previous decision to serve. It is unclear as to whether countries in conflict would foresee the possibility of being sent peacekeepers let alone foresee how Security Council membership could help it avoid receiving peacekeepers.

Finally, since an earlier study showed that U.N. Security Council membership is associated with an increase in aid (Kuziemko and Werker 2006), I include aid as an additional control to verify that the U.N. Security Council's effect on our outcome variables is not through aid rather than through the absence of peacekeepers.

As a robustness check, I regress the number of years a country serves as a member of the Security Council during the episode on a number of controls in Table 1.1 to show that being on the Security Council during an episode of conflict is uncorrelated with country characteristics that we might think matter to peace such as GDP or trade.

Apart from the concern that membership on the Security Council is proxying for country characteristics, which has been mostly mitigated by including years on the Security Council before the conflict as a control and by the results in Table 1.1, there is the more insidious concern that membership in the Security Council during the episode of conflict reflects something about the conflict. In particular, it could be that countries in conflict only make it onto the Security Council when the conflict is weak, with weaker conflicts both less likely to receive peacekeepers and more likely to last a long time. While I include battle deaths, the duration of the episode of conflict and aid as controls, all of which may reflect the nature of the conflicts, I cannot dismiss the possibility that the conflicts of countries that make it onto the Security Council are weaker in respects not captured by these variables. It should be noted, however, that the literature suggests that worse conflicts, rather than weaker conflicts, empirically have a shorter duration of peace (Doyle and Sambanis 2000; Fortna 2008). In this case, even if those countries that did make it onto the Security Council had weaker conflicts, the bias would be in the opposite direction of my findings.

Finally, to guard against weak instrument concerns, I report the Anderson-Rubin Wald

and the Stock-Wright LM test statistics that are robust in the presence of weak instruments.[9]

1.8 Conclusions

This paper estimates the effects of peacekeeping on the duration of conflict, the duration of peace after peace has been obtained, and the likelihood of the conflict ever re-occurring. Plausibly exogenous variation in the likelihood of being sent peacekeepers was exploited to estimate these effects. In particular, the longer countries in conflict serve as temporary members of the Security Council during the conflict itself, the less likely they are to be sent peacekeepers, even after controlling for previous service on the Security Council. While it is possible that this reflects something about the conflicts themselves, with only those countries in lighter conflicts able to serve on the Security Council, if we believe that worse conflicts lead to generally more fragile peace (Doyle and Sambanis 2000; Fortna 2004) this would create bias in the opposite direction of my findings. The two main methods used in this paper - 2SLS and a control function approach - each have their own advantages and disadvantages. While the control function approach can be inconsistent when the equations it estimates are misspecified, it has improvements in efficiency over 2SLS. That being said, in this paper we saw that the 2SLS results suggested a significant relationship between peacekeeping and the duration of peace even when using a linear first stage.

The results suggest that U.N. peacekeepers do indeed prolong the peace once peace has been obtained and lower the chance that another episode of the same conflict will ever re-occur. U.N. peacekeepers do not, however, shorten the duration of the episode of conflict itself. This is perhaps not surprising since peacekeepers are often sent only once there has been a break in the fighting.

The control function approach further told us that the greater the chance of being selected to receive peacekeepers, the shorter the duration of peace. Thus, when peacekeepers increase the duration of peace and lower the chance that another episode of the same conflict will ever re-occur, they do so for the worst conflicts. In fact, we cannot say much about whether peacekeepers would help if sent to the "easier" cases.

A further strength of this paper is that it looks at both civil and interstate conflicts. Focusing on only one type is more the norm, but this separation is arguably artificial when it comes to peacekeeping (Cunningham and Lemke 2009).

Finally, and perhaps most interestingly, this paper reveals that the leaders of countries in conflict often do not want to receive peacekeepers. This empirical finding adds to the literature and calls for further research.

[9]If we write the first stage regression as $X = Z\Pi + u$ and the second as $Y = X\beta + e$, where Z is an instrument for X, $Y = (Z\Pi + u)\beta + e$. Anderson-Rubin estimates $Y = Z\Gamma + \eta$ and tests that $\Gamma = 0$. If it fails to reject $\Gamma = 0$, it also fails to reject $\beta = 0$. The intuition is that it is robust to weak instruments because as instruments become weak Π becomes smaller and thus so does $\Pi\beta$, and the likelihood that $\Gamma = 0$ is rejected goes down. The Stock-Wright test follows similar reasoning.

In summary, these methods provide evidence from a completely new angle that supports the effectiveness of peacekeeping operations and that suggests that governments can be a hindrance to peace. A lot of heterogeneity in outcomes remain which can be modelled in future work. However, overall the results are encouraging because they suggest that even relatively small peacekeeping forces can often help countries avoid or escape from potential conflict traps - if governments will let them.

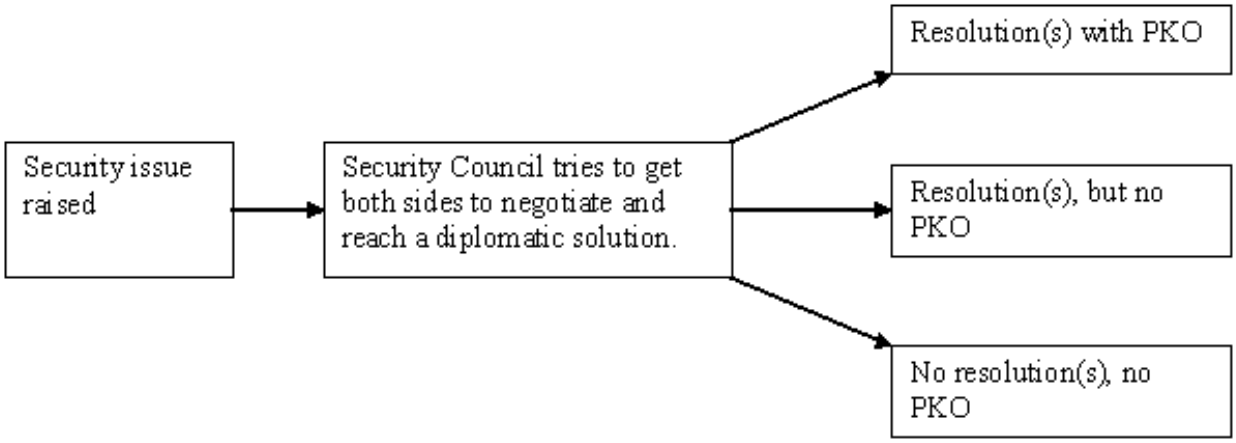


Figure 1.1: U.N. Security Council Decision-Making Process

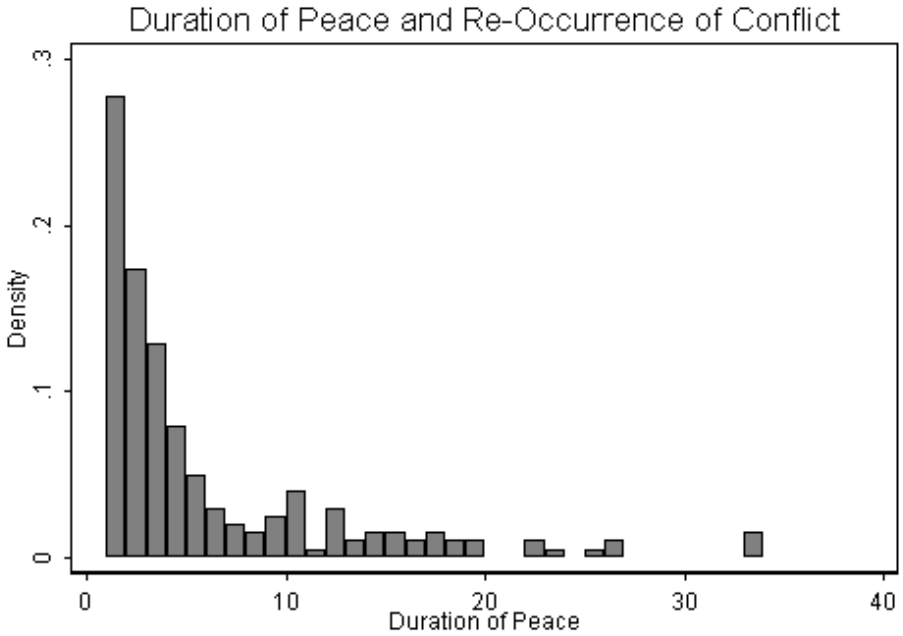


Figure 1.2: Re-Occurrence of Conflicts

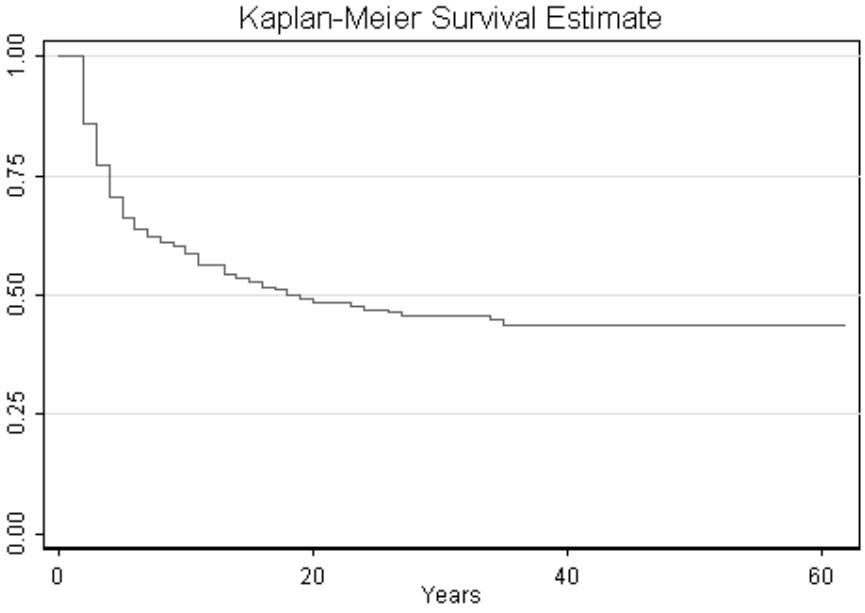


Figure 1.3: Kaplan-Meier Survival Estimate

Table 1.1: Poisson Regression of Years on Security Council During an Episode on Controls

	(1)	(2)
	b/se	b/se
Years on Security Council During Episode		
Battle Deaths	0.099	0.058
	(0.073)	(0.066)
Episode Duration	0.096***	0.101***
	(0.021)	(0.017)
Year Episode Ended	-0.062**	-0.048*
	(0.024)	(0.024)
Years on Security Council	-0.252	-0.208
in Last 5 Years	(0.194)	(0.195)
Years on Security Council	0.497***	0.445***
in Last 10 Years	(0.147)	(0.151)
Non-Military Aid	3.848***	3.431**
	(1.409)	(1.489)
Military Aid	-1.079	-0.688
	(0.841)	(0.868)
ln Trade with U.S.	-0.030	
	(0.089)	
GDP per Capita (current $, thousands)		-0.000
		(0.000)
Observations	184	165

In all tables in this chapter, battle deaths is in thousands, aid is in billions, and * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.2: OLS Estimation of the Effects of Being Sent Peacekeepers on the Duration of Peace once Peace is Obtained

	(1) b/se	(2) b/se	(3) b/se
Duration of Peace			
Peacekeepers Sent	0.552	0.279	0.716
	(1.169)	(1.301)	(1.178)
Episode Duration	-0.104	-0.111	-0.079
	(0.090)	(0.091)	(0.093)
Year Episode Ended	-0.590***	-0.597***	-0.569***
	(0.103)	(0.101)	(0.106)
Battle Deaths	1.590***	1.546***	1.489***
	(0.501)	(0.490)	(0.479)
Years on Security Council		-0.615	-0.527
in Last 5 Years		(0.907)	(0.893)
Years on Security Council		-0.373	-0.012
in Last 10 Years		(0.566)	(0.543)
Non-Military Aid			-9.138
			(5.965)
Military Aid			7.238**
			(3.187)
Observations	184	184	184
R^2	0.29	0.30	0.32

Table 1.3: Cox Hazard Model Estimation of Effects of Peacekeepers on the Duration of Peace

	(1) b/se	(2) b/se	(3) b/se
Duration of Peace (Hazard Ratio Reported)			
Peacekeepers Sent	0.770	0.792	0.742
	(0.280)	(0.298)	(0.263)
Episode Duration	1.022	1.023	1.017
	(0.021)	(0.021)	(0.023)
Year Episode Ended	1.017	1.018	1.014
	(0.018)	(0.018)	(0.020)
Battle Deaths	0.446	0.445	0.456
	(0.288)	(0.284)	(0.285)
Years on Security Council		1.067	1.053
in Last 5 Years		(0.172)	(0.172)
Years on Security Council		1.037	0.985
in Last 10 Years		(0.113)	(0.111)
Non-Military Aid			3.927
			(5.208)
Military Aid			0.249
			(0.270)
Observations	184	184	184

Table 1.4: 2SLS of the Effects of Being Sent Peacekeepers on the Duration of Peace

	(1) b/se	(2) b/se	(3) b/se
Peacekeepers Sent (First Stage)			
Years on Security Council	-0.134***	-0.100*	-0.145***
During Episode	(0.052)	(0.055)	(0.055)
Episode Duration	0.019***	0.018***	0.015**
	(0.007)	(0.007)	(0.006)
Battle Deaths	0.035	0.023	0.038
	(0.056)	(0.058)	(0.053)
Year Episode Ended	0.002	0.002	-0.002
	(0.005)	(0.005)	(0.005)
Years on Security Council		0.047	0.041*
in Last 5 Years		(0.030)	(0.022)
Years on Security Council		-0.079***	-0.118***
in Last 10 Years		(0.028)	(0.031)
Non-Military Aid			1.171**
			(0.481)
Military Aid			-0.541**
			(0.256)
Duration of Peace			
Peacekeepers Sent	12.342**	15.222**	9.914**
	(5.511)	(7.115)	(4.011)
Episode Duration	-0.247***	-0.293**	-0.154
	(0.093)	(0.119)	(0.100)
Battle Deaths	1.342**	1.401*	1.339***
	(0.560)	(0.721)	(0.396)
Year Episode Ended	-0.625***	-0.644***	-0.572***
	(0.127)	(0.131)	(0.122)
Years on Security Council		-1.582	-1.096
in Last 5 Years		(1.269)	(0.948)
Years on Security Council		1.192	1.300*
in Last 10 Years		(0.778)	(0.775)
Non-Military Aid			-17.725**
			(8.014)
Military Aid			12.115***
			(4.326)
Observations	184	184	184
First Stage F-stat	7.17	4.79	8.87
p-values of tests:			
Anderson-Rubin Wald test	0.00	0.01	0.05
Stock-Wright LM S	0.03	0.05	0.07

Table 1.5: Cox Hazard Model Estimation of Effects of Peacekeepers on the Duration of Peace, Control Function Approach

	(1) b/se	(2) b/se	(3) b/se
Duration of Peace (Hazard Ratio Reported)			
Peacekeepers Sent	0.144**	0.064**	0.134**
	(0.123)	(0.075)	(0.131)
Residual	2.717**	4.351**	2.793**
	(1.214)	(2.570)	(1.414)
Episode Duration	1.041**	1.052**	1.029
	(0.021)	(0.024)	(0.024)
Year Episode Ended	1.023	1.026	1.014
	(0.017)	(0.018)	(0.019)
Battle Deaths	0.526	0.537	0.543
	(0.305)	(0.301)	(0.302)
Years on Security Council		1.251	1.179
in Last 5 Years		(0.201)	(0.193)
Years on Security Council		0.802	0.776
in Last 10 Years		(0.123)	(0.131)
Non-Military Aid			17.486*
			(26.596)
Military Aid			0.108*
			(0.133)
Observations	184	184	184

Chapter 2

Distortions in the Decision to Export in Developing Countries

2.1 Introduction

In the standard trade literature, more productive firms should export over less productive firms, all else equal. This premise appears in Melitz (2003), Luttmer (2007), and Eaton *et al.* (2009), among others. The intuition is that more productive firms can obtain larger profits and so should be the first firms that choose to export as exporting a particular good becomes profitable. However, although productivity has historically been considered a very important determinant of a firm's choice to export, it has been shown in other fields that distortions can sometimes override the importance of a firm's productivity. Hsieh and Klenow (2009), in particular, showed that the most productive firms in China and India were using relatively little labour and capital due to distortions affecting their profitability. The theory behind this paper combines Hsieh and Klenow-like distortions with a Melitz-like model that accords productivity a key role. The model implies that opening to trade does not necessarily have to increase firms' productivity and may in fact decrease it if many distortions are present. Data from Colombia are used to provide support for the model.

To put this paper in the context of the literature, it will be helpful to think of a 2-by-2 matrix, with one dimension being constituted by the categories "distorted" and "not distorted" and the other consisting of the categories "closed" and "open". This matrix is represented in Figure 2.1. Melitz (2003) focused on the possible productivity gains of moving from the "closed, not distorted" cell to the "open, not distorted" cell; Hsieh and Klenow (2009), from the "closed, distorted" cell to the "closed, not distorted" cell.[1] This paper models the effects of distortions under an open economy, and hence considers the possible gains of moving from the "open, distorted" cell to the "open, not distorted" cell. We can also

[1] While Hsieh and Klenow's data were from open economies, the model did not consider trade.

to some extent consider the ramifications of moving in the three other directions illustrated by dashed arrows in the Figure, completing the matrix.

Yet this paper does more than apply Hsieh and Klenow to exporters. The difference stems from a hypothesized interaction between distortions and openness to trade. Both Hsieh and Klenow and Melitz focus on the productivity gains to re-allocations; in Hsieh and Klenow, reducing distortions results in productivity gains due to the re-allocation of resources from less to more productive firms, while in Melitz, opening to trade results in productivity gains partially due to the re-allocation of resources from less to the more productive exporters. In the upper-most part of Figure 2.2, we see a possible distribution of firms that would export under an open economy, where ϕ represents productivity. Under Melitz, if an economy with these potential exporters opened to trade, these firms, which are on average more productive, would gain market share and push up overall productivity in their sectors. But now consider the case in which there are distortions such that some relatively less productive firms export and some relatively more productive firms do not. The effect of first reducing these distributions is illustrated in the middle part of Figure 2.2, with the dashed black line representing the distribution of potential exporters under distortions and the solid black line representing the distribution of potential exporters when the distortions are reduced. The mean productivity of potential exporters rises, and so the mean productivity of firms that gain market share also rises. The overall effect on productivity gains from re-allocation is ambiguous without further assumptions, however, as it is possible that fewer firms now export.

It should be noted that in extreme cases, the result from Melitz that opening to trade increases productivity no longer holds. Consider the bottom part of Figure 2.2. If the firms under the black curve are the ones with the highest profits and the most likely to start exporting, due to distortions, then when the country opens to trade it is these firms that would get a boost and gain market share, lowering overall productivity.

The empirical component of this paper will test the main premise of the model. Namely, in the model, distortions affect the decision to export in a similar manner as productivity, with distortions allowed to either increase or decrease a firm's profitability. Thus, I look for evidence that Hsieh-Klenow-type distortions affect a firm's choice to export. I also ask whether distortions affect entry and exit decisions, separately. Finally, I will be able to construct some counterfactuals that suggest how much distortions affect the productivity gains from trade in Colombia.

The results suggest that distortions can indeed encourage or discourage firms from exporting. The implication is that trade may not improve productivity in particularly distorted sectors or countries as much as has historically been assumed and, conversely, estimates of the productivity-enhancing effects of trade that are based on more distorted cases may understate the true effects of trade on productivity in less distorted sectors or countries. This finding is robust to different methods and parameter choices, and it also persists when I focus on particularly homogeneous goods.

There is a large literature on trade reforms and on the potential productivity gains from

trade. It has long been established that exporters are more efficient in general than non-exporters (*e.g.* Aw and Hwang, 1995; Griliches and Regev, 1995). Clerides *et al.* (1998) and Bernard and Jensen (1999) found evidence that this is largely due to more productive firms tending to self-select into exporting. A related strand of the literature focused on the effects of reductions in tariffs and other trade reforms on exporting and the productivity of exporters (Fernandez, 2007; Pavcnik, 2002; Tybout and Westbrook, 1995; Haddad, 1993). Faced with heterogeneous results, it was sometimes posited that something other than productivity was driving selection into exporting (*e.g.* Aw *et al.*, 2001; Aw *et al.*, 2000). However, most of the earlier models of firms' selection into markets - not necessarily export markets - focused on productivity (Ericson and Pakes, 1995; Hopenhayn, 1992; Jovanovic, 1982). More recently, Foster *et al.* (2008) modeled selection in a domestic market to be a function of demand, productivity and factor prices that were producer-specific. This work is the closest, theoretically, to this paper, as factor prices could be hypothesized to vary across firms due to distortions.

This paper builds on the literature in a few ways. By extending Melitz (2003) to incorporate Hsieh and Klenow (2009) -type distortions and looking for evidence that these distortions affect the decision to export, this paper examines the extent to which opening to trade can cause productivity gains and highlights the effects of the interaction between distortions and opening to trade on productivity gains. Second, while many papers have focused on how narrower barriers to trade such as tariffs hinder firms from exporting, this paper looks at the effects of distortions other than those caused by tariffs and subsidies and gets a sense of what share of total distortions these represent. In particular, I am able to tie the Hsieh-Klenow type distortions that I impute to government ownership. The results suggest that countries considering trade reforms pay attention to domestic distortions, as well.

The rest of this paper proceeds as follows. In the next section of this paper I build a theoretical model extending Melitz (2003) to account for distortions. Later sections describe the data and detail the identification strategy to be used. Following this, results are presented and several robustness checks are conducted. Finally, I discuss the findings, before concluding.

2.2 A Simple Model Integrating Distortions with Melitz

Recall that in the Melitz model consumers have preferences over goods ω given by a CES utility function:

$$U = \left[\int_{\omega \in \Omega} q(\omega)^{\rho} d\omega \right]^{\frac{1}{\rho}}$$

where $q(\omega)$ is the quantity of good ω and ρ is defined by the elasticity of substitution, σ, as $(\sigma - 1)/\sigma$. The aggregate price for aggregate good U is given by:

$$P = \left[\int_{\omega \in \Omega} p(\omega)^{1-\sigma} d\omega\right]^{\frac{1}{1-\sigma}}$$

where $p(\omega)$ is the price of good ω. The optimal consumption of and expenditure on individual goods are then given by:

$$q(\omega) = Q\left[\frac{p(\omega)}{P}\right]^{-\sigma}$$
$$r(\omega) = R\left[\frac{p(\omega)}{P}\right]^{1-\sigma}$$

where Q and R are the aggregate consumption and expenditures or revenues, respectively, and $r(\omega)$ is the expenditure on good ω.

In the Melitz model, firms face a fixed cost to production as well as a constant marginal cost. Firms have the same fixed cost f but have different productivity levels ϕ. ϕ is drawn from a distribution $\mu(\phi)$. Given these assumptions, labour is a function of output q as follows: $l = f + \frac{q}{\phi}$. Each firm then faces a demand curve with constant elasticity σ and has the same profit-maximizing markup: $\frac{\sigma}{(\sigma-1)} = \frac{1}{\rho}$. The pricing rule is thus: $p(\phi) = \frac{w}{\rho\phi}$, where w is the wage rate, common across firms, which will be normalized to one.

To add distortions to Melitz, assume that firms face distortions, η, drawn from a distribution $\gamma(\eta)$ independent from ϕ. Assume for simplicity that $f = 0$ and that these distortions affect marginal costs so that: $l = \frac{q}{\phi(1-\eta)}$. The mark-up rule now implies $p(\phi, \eta) = \frac{1}{\rho\phi(1-\eta)}$. Profits π are given by $r(\phi, \eta) - l(\phi, \eta) = \frac{r(\phi,\eta)}{\sigma}$ where r is firm revenue.

Using the original equation for $r(\omega)$, we can see:

$$r(\phi, \eta) = R\left[\frac{p(\omega)}{P}\right]^{1-\sigma} = R(P(\rho\phi(1-\eta)))^{\sigma-1}$$

and consequently:

$$\pi(\phi, \eta) = \frac{R(P(\rho\phi(1-\eta)))^{\sigma-1}}{\sigma}$$

Now if firm 1 and firm 2 have productivity levels ϕ_1 and ϕ_2 and face distortions in their marginal costs η_1 and η_2, their outputs and revenues will differ by:

$$\frac{q(\phi_1, \eta_1)}{q(\phi_2, \eta_2)} = \left[\frac{\phi_1(1-\eta_1)}{\phi_2(1-\eta_2)}\right]^{\sigma}$$
$$\frac{r(\phi_1, \eta_1)}{r(\phi_2, \eta_2)} = \left[\frac{\phi_1(1-\eta_1)}{\phi_2(1-\eta_2)}\right]^{\sigma-1}$$

This tells us that in addition to the results from Melitz that a more productive firm, with higher ϕ, will be bigger, charge a lower price, and earn higher profits than a less productive firm, we can add the results that a firm that faces larger marginal costs from distortions, with higher η, will be smaller, charge a higher price, and earn lower profits.

We can also derive some results about the effects of trade on productivity under distortions. Melitz shows under his model that there are certain cut-off productivity levels ϕ^*_α and ϕ^* under autarky and under an open economy, respectively, below which firms will choose to exit the market, with $\phi^* > \phi^*_\alpha$. Also, only the firms with productivity levels above some threshold ϕ^*_x enter the export market, with $\phi^*_x > \phi^*$. If we define a new variable $V(\phi, \eta)$ such that it replaces ϕ in Melitz's model in determining firms' profit levels, then following the arguments in Melitz there will be certain cut-off levels V^*_α and V^* under autarky and under an open economy below which firms will choose to exit the market, with $V^* > V^*_\alpha$. Similarly, ϕ^*_x in Melitz is replaced by V^*_x. With no fixed costs, V takes the simple form $\phi(1-\eta)$.

Once a country opens to trade in Melitz, a firm's fortune is determined by its productivity. The most productive firms export and increase their market share and profits; some of the less productive firms also export and increase their market share but incur a loss in profits; still less productive firms continue to produce but do not export and lose both market share and profit; the least efficient firms exit the market. In my model, the same delineations apply, but with V rather than ϕ being the determining factor. Since V is a function of both ϕ and η, it is elementary to see that productivity is affected less by opening to trade in my model than in Melitz's. Depending on the distributions $\gamma(\eta)$ and $\mu(\phi)$, some $\lambda \geq 0$ fraction of firms with $\phi > \phi^*$ have $V < V^*$ due to a high, positive draw of η but exit the market, and some $\lambda' \geq 0$ fraction of less productive firms with $\phi < \phi^*$ have $V > V^*$ due to a low, negative draw of η and so stay in the market. The same arguments apply to the choice to enter or exit the export market, replacing ϕ^* with ϕ^*_x and V^* with V^*_x.

I will test the predictions of this model using data from Colombia. The next section describes the data; following this, the identification strategy will be detailed.

2.3 Data

This paper uses data from the Colombian annual survey of manufactures, the Encuesta Anual Manufacturera (EAM) from the Departamento Administrativo Nacional de Estadistica (DANE). These panel plant-level data cover the 1981-1991 period and include exports and variables that can be used to construct measures of productivity. Plants have been categorized into sectors represented by 4-digit SIC codes. 16,226 plants appear in the dataset, of which 2,442 export for at least one year over the time period. Roberts and Tybout (1997) describe the dataset in more detail.

As in Hsieh and Klenow (2009), I will make a distinction between revenue productivity (TFPR) and physical productivity (TFPQ). If we assume the standard Cobb-Douglas production function, TFPR is represented by:

$$TFPR_{ist} = \frac{p_{ist}Y_{ist}}{K_{ist}^{\alpha_s}(w_{ist}L_{ist})^{1-\alpha_s}}$$

where i is the plant, s is the sector, t is the period, Y is output, K is capital, w are wages, L is labour and α_s and $1-\alpha_s$ represent, respectively, the capital and labour shares, while TFPQ can be obtained if we know the plant-specific prices as:

$$TFPQ_{ist} \equiv \phi_{ist} = \frac{Y_{ist}}{K_{ist}^{\alpha_s}(w_{ist}L_{ist})^{1-\alpha_s}}$$

The main measure of TFPR that will be used in this paper is derived from the gross value of output, the total value of fixed assets, production and non-production labour, the value of intermediate consumption, and the value of investment, using Olley and Pakes' method (1996).

To impute ϕ_{ist}, I assume the same relationship between prices and quantities as Hsieh and Klenow (2009): $Y_{ist} = (p_{ist}Y_{ist})^{\frac{\sigma}{\sigma-1}}$. Figure 2.3 contains histograms of TFPR and TFPQ, respectively. I assume $\sigma = 3$ for the regressions in this paper.[2]

As in Hsieh and Klenow (2009), I will consider two types of potential distortions. First, there may be distortions that affect both capital and labour by the same proportion, which I will label η_y. For example, subsidies could artificially keep a firm's output high, which would result in lower marginal products of both capital and labour for the firms that receive them. Second, there may be distortions that affect the relative marginal products of capital and labour, which I will denote η_k. An example would be firms having differential access to

[2]It should be noted that the value chosen for σ does not affect much. η_y and η_k are normalized by their standard deviation, so σ's value, affecting all plants equally, has no effect on them. σ has an effect on ϕ, and for this reason I also conduct regressions using $\sigma = 5$, but σ does not have an effect on the rank of a plant's productivity relative to others in its sector, and thus the choice of σ is largely unimportant.

credit.

With these distortions defined, profits are given by:

$$\pi_{ist} = (1 - \eta_{y_{ist}}) p_{ist} Y_{ist} - w L_{ist} - (1 + \eta_{k_{ist}}) r K_{ist}$$

Each firm faces a demand curve with constant elasticity σ and has the same profit-maximizing markup $\frac{\sigma}{(\sigma-1)}$. The pricing rule is thus:

$$p_{ist} = \frac{\sigma}{\sigma - 1} \left(\frac{r}{\alpha_s} \right)^{\alpha_s} \left(\frac{w}{1 - \alpha_s} \right)^{1-\alpha_s} \frac{(1 + \eta_{k_{ist}})^{\alpha_s}}{\phi_{ist}(1 - \eta_{y_{ist}})}$$

The capital-labour ratio is given by:

$$\frac{K_{ist}}{L_{ist}} = \frac{\alpha_s}{1 - \alpha_s} \frac{w}{r} \frac{1}{1 + \eta_{k_{ist}}}$$

which in turn can be used to derive revenue per worker and the revenue-capital ratio:

$$\frac{p_{ist} Y_{ist}}{L_{ist}} = \frac{\sigma}{\sigma - 1} \frac{w}{1 - \alpha_s} \frac{1}{1 - \eta_{y_{ist}}} \tag{2.1}$$

$$\frac{p_{ist} Y_{ist}}{K_{ist}} = \frac{\sigma}{\sigma - 1} \frac{r}{\alpha_s} \frac{1 + \eta_{k_{ist}}}{1 - \eta_{y_{ist}}} \tag{2.2}$$

These should be proportional to the marginal revenue product of labour ($MRPL_{ist}$) and the marginal revenue product of capital ($MRPK_{ist}$), respectively. Equations (1) and (2) allow me to impute $\eta_{y_{ist}}$ and $\eta_{k_{ist}}$ directly from the data.

In the imputation of η_y and η_k, I follow Hsieh and Klenow (2009) by assuming a fixed $r = 0.1$ and a common w, using a plant's wage bill instead of the number of workers to measure L_{ist}, and assuming $\sigma = 3$. I use Colombian labour shares to compute the elasticity of output with respect to labour in each industry ($1 - \alpha_s$), but as a robustness check I set $1 - \alpha_s = 0.6$ in an alternative specification since, as Hsieh and Klenow note, any distortions may affect these values (2009). The issue is moot for most of the regressions anyway, since α_s varies at most by sector (using the Colombian labour shares) and my regressions are focused on intra-industry variation. Y, K, and L are each winsorized at 1% before all calculations to remove outliers; η_y and η_k are also winsorized after being calculated to again remove extreme values. It should be noted that due to the method of imputation, η_y is bounded above by 1 and η_k is bounded below by -1. The values that η_y and η_k take below or above

these bounds, respectively, are not particularly meaningful except in relative terms, as they depend on the values chosen for σ, α_s, and so on. Because of this, I normalize η_y and η_k by their standard deviations so that a 1 unit increase in their values represents a standard deviation. Histograms of normalized η_y and η_k are presented in Figure 2.4.[3]

There may be some concerns about what these imputed values of η_y and η_k are actually capturing. To address this, I consider a few factors that could plausibly cause variation in η_y and η_k and relate the imputed values to the data.

First, given the emphasis in the literature on the effects of tariffs on the decision to export, I correlate η_y and η_k with export taxes. Only 446 plants in my dataset pay export taxes, but among those that do η_y and η_k have a 0.5596 and 0.4212 correlation with the natural log of export taxes, respectively.[4]

Second, I summarize the average values of η_y and η_k for plants in each of nine categories to see if there is any relationship between a plant's distortions and whether it is a proprietorship, a limited partnership, a collective, a corporation, a *de facto* corporation, a joint partnership, a joint stock company, a cooperative, or a state enterprise.[5] As can be seen in Table 2.1, η_y and η_k vary across these categories seemingly at random, with one exception: the lowest mean value of both η_y and η_k is found among state enterprises. Low values of η_y and η_k, it should be recalled, suggest favourable conditions. These are the relatively profitable distortions. So there is some reason to believe η_y and η_k are capturing the kinds of distortions in the literature.

We may also suspect that η_y and η_k could be capturing geographic factors. When I summarize the average values of η_y and η_k for plants in each of nine regions identified in the data, however, there does not appear to be a clear trend that holds for both η_y and η_k.

There is still the possibility that η_y and η_k are partially capturing heterogeneity in the products that plants are producing within their sectors. I will address this concern later on by focusing a set of regressions on particularly homogeneous sectors.

Fitting a curve to the data using a locally weighted scatter plot smoother (loess), I find evidence supporting investigating the relationship between η_y, η_k and a plant's decision to export. Figure 2.5 plots loess curves conditional on controls, after centering the data to remove plant-level fixed effects.[6]

η_y and η_k here exhibit a negative relationship with exporting. The natural log of a plant's output and whether a plant exported the previous year both appear to have strong positive relationships with whether a plant exports, consistent with the literature (*e.g.* Roberts and Tybout, 1997).

[3]The bottom 5% and top 5% of η_y and η_k, respectively, are dropped in this figure.

[4]Because of this high correlation, I later try running my regressions excluding plants that paid export taxes, but the coefficients on η_y and η_k and their significance do not change much. These results are available upon request.

[5]There are also a few religious organizations, but these are not large producers or exporters and so are excluded for simplicity.

[6]5,000 nearest neighbours are used to obtain these plots.

These figures are only suggestive, but help to motivate further study of the relationship between distortions and exporting.

2.4 Identification Strategy

To test to what extent the decision to export depends on distortions, I run the following regression:

$$E_{ist} = \alpha + \beta_1 \eta_{y_{ist}} + \beta_2 \eta_{k_{ist}} + \sum_j \gamma_j Z_{jist} + \zeta X_t + \epsilon_{ist} \tag{2.3}$$

where E_{ist} is a 1/0 dummy indicating whether the plant exports, Z_{ist} is a vector of plant characteristics, including plant-specific fixed effects, X_t is a time trend, and ϵ_{ist} is an error term. Standard errors are bootstrapped. If higher η_y or η_k reduces the likelihood a plant exports, β_1 or β_2 should be negative.

In alternative specifications, I control for variables that reflect the kinds of distortions more typically discussed in the trade literature - tariffs and government subsidies - and find the distortions captured by η_y and η_k appear to have an effect beyond these.

For plant characteristics, I include the rank of a plant's productivity within its sector, normalized to run between 0 and 1. I also sometimes include a variable indicating whether the plant exported last year; the log of its output last year as a measure of size; the number of plants in the sector that exported last year; the log of its taxes last year; and the log of its subsidies last year.[7] I include productivity rank rather than raw productivity for the main regressions since it seems to be slightly more strongly related to whether a plant exports, possibly by reducing noise. Including productivity rank rather than raw productivity does not significantly change the results for η_y and η_k; results using raw productivity are available upon request.

I also run the same regression explicitly on entry into and exit from exporting, separately. These regressions are run on the set of of plants not exporting the previous year and those exporting the previous year, respectively. Since the main results from estimating (3) are based on intra-plant variation in exporting, the main results also capture entry and exit, jointly, and drop fewer observations than these latter regressions through the plant-specific fixed effects. Still, I run these secondary regressions on entry and exit separately for completeness and include results in the Appendix.

[7]Before taking the natural log of any variable, the small value 0.001 is added to it so as to avoid losing the observations that have a value of 0.

2.5 Results

In Table 2.2 we see that the higher η_y and η_k, the less likely a plant is to export, consistent with the model. Estimates retain significance when taxes are controlled for, suggesting distortions have effects beyond those captured by taxes. Plants with higher taxes (controlling for output) were themselves less likely to export; this despite the fact that export taxes comprised one portion of the total taxes variable. While subsidies and export taxes last period were included separately in alternative specifications, neither affected many plants and results were insignificant.[8]

Plants were also more likely to export this period if they exported last period, if they produced more, or if there was a larger number of exporters in the sector last year; all consistent with past research (*e.g.* Clerides *et al.*, 1998; Roberts and Tybout, 1997). Strangely, over the time period studied, plants were less likely to export as time went on, after controlling for the aforementioned variables. A plant's relative productivity rank within its sector-year appears as insignificant in these results, but has a weak positive association. All tables report the odds ratio, for easier interpretation; coefficients below 1 represent a negative relationship and coefficients above 1 a positive relationship.

Results from the regressions on only the entry or exit of plants from exporting are included in the Appendix. η_y and η_k exhibit weaker relationships when the sample is divided in this way, but still show some significant results, with all coefficients having the positive or negative relationship that would be predicted by the model. High η_y and η_k have opposite effects on entry and exit, as would be expected, discouraging entry and encouraging exit.

2.6 Robustness Checks

I conduct a series of robustness checks to support my main findings.

First, one could worry that η_y and η_k are best thought of as varying by plant rather than by year within plants. Then, one would wish to exploit the cross-sectional rather than the time-series dimension of the data. I thus re-run the regressions, collapsing the time-series dimension and using only average values of $\eta_{y_{ist}}$ and $\eta_{k_{ist}}$ for each plant. The dependent variable is now whether a plant ever exports and sectoral fixed effects are included, with standard errors again bootstrapped. Results are presented in Table 2.3 and appear comparable. Apart from the negative effects of $\eta_{y_{ist}}$ and $\eta_{k_{ist}}$, plants with higher average output are more likely to export.

As a second robustness check, I repeat the analyses using a subset of products that are plausibly fairly homogeneous, given that there could be unobserved heterogeneity in the products being manufactured within a sector. If goods within the same sector are heterogeneous by, for example, having different qualities, this could be reflected in η. In particular,

[8] These results are available upon request.

one might think that higher quality goods would have higher prices and these might also be the kinds of goods that would be exported. Of course, since η_y and η_k have a positive relationship with price, by construction, this should bias results in the opposite direction of my findings, making them a lower bound. Still, lest there are concerns, I run the same regressions on sets of seemingly homogeneous goods.

Specifically, I create a sub-category of goods which includes the sectors: sugar refining and sugar products; tobacco; sawmills; pulp mills; paper and cardboard boxes and containers; printing and publishing; paint, varnish and lacquer; petroleum refining; petroleum and coal products; tires; iron and steel; copper and aluminum; lead and zinc; tin and nickel; and precious metals. The results of the regressions run on this sub-category are presented in Table 2.4.

Of course, when restricting the sample to only these sectors, there are far fewer observations, and some significance is lost. Still, results that were previously positively correlated with the decision to export are still positively correlated, and results that were previously negatively correlated with the decision to export are still negatively correlated, with the exception of the number of exporters in the sector last year and the log of taxes last year.

Finally, I try changing parameter values to see if results change. In particular, I try setting $1-\alpha_s = 0.6$ for all sectors, instead of using Colombian labour shares, since the value of Colombian labour shares may themselves reflect distortions. As expected, these changes do not substantially affect my findings. The results of these regressions are included in the Appendix.

2.7 Discussion

Overall, the results seem to suggest that distortions affect the decision to export in substantial ways, causing firms' "churning" to depend less on productivity. This implies that firms in sectors or countries with particularly high distortions stand to improve their productivity less from trade than would otherwise be expected.

While I focused on the decision to enter or exit the export market, results would presumably be comparable regarding entering or exiting the domestic market.[9] The effects on entry and exit examined here involved distortions η that pushed some firms with $\phi > \phi_x^*$ below V_x^* and some firms with $\phi < \phi_x^*$ above V_x^*; obviously, if η generally has these effects that reduce the importance of productivity, it could also push some firms with $\phi > \phi^*$ below V^* and some firms with $\phi < \phi^*$ above V^*.

We can compute some counterfactuals to gauge how much of a difference distortions

[9] I cannot observe the universe of possible plants that elect not to enter the domestic market, which is why I do not study it here. In contrast, while plants could plausibly enter the domestic and export markets simultaneously, or only export, it is likely that not much is lost by assuming that all would-be exporters produced for the domestic market - and hence appear in my sample - before exporting.

make. Using the coefficients from the first regression presented here to calculate the likelihood of exporting under various conditions, I try reducing the distortions to zero and taking draws of the plants, weighted to reflect their estimated likelihood of exporting. I find that reducing distortions to zero and holding the number of exporters constant would increase the productivity of exporters by an average of 4.9% by encouraging more productive exporters to select into exporting.[10]

We can also get a rough sense of the productivity gains from opening to trade under distortions compared to opening to trade under no distortions. While Colombia is quite an open economy, openness is a spectrum, and reducing trade costs can always make countries essentially more open to trade. It is in this sense that we can approach estimating the effects of moving between the various cells in Figure 2.1. Bernard *et al.* (2003), who focus specifically on a hypothetical 5% decrease in trade costs in the U.S., find that redistribution to the more productive exporting firms causes a 1% rise in productivity overall. 0.8% is attributed to the exit of less productive firms; 3% of all firms, in their simulation, exit the market entirely with the decrease in trade costs. If I predict which firms would be among the first 3% to exit in Colombia under current distortions and under zero distortions, I find that when distortions are eliminated the exiting firms are 17.4% less productive than the exiting firms under distortions. The theoretical re-allocation of resources from lower productivity, non-exporting firms to higher productivity, exporting firms also determines the extent to which openness causes productivity gains, and I cannot estimate this without making more assumptions. Still, the overall relationship between productivity gains from trade and distortions appears clear.

One may also wonder whether plant size varies in any systematic way with distortions. Theory would suggest that there are plants with low productivity that get a boost from distortions and so gain market share; there will also be plants which are hurt by their draw of η and, despite high ϕ, lose market share. By the manner that η was imputed, plant size would be mechanically negatively correlated with η, so I cannot independently examine this here. Similarly, plants with higher η would mechanically have higher prices and lower profits, as the model also predicts.

One may also ask how results would change without the assumption that η and ϕ were drawn independently from $\gamma(\eta)$ and $\mu(\phi)$. What if, instead, η were increasing in ϕ or decreasing in ϕ? The case in which η decreased in ϕ would be the case of most concern, because then a negative relationship between η and whether a plant exports could simply be picking up the unmodeled negative correlation. However, if anything, it would seem that η and ϕ would be positively correlated, since if plants had high ϕ they could more easily support high η; if plants with low ϕ had too high η, they would simply go out of business. Figure 2.6 shows the set of η and ϕ that would produce positive profits under the modified Melitz model in which $\phi(1-\eta)$ affects profits.[11]

[10]The set of exporters was drawn 100 times for this simulation.

[11]In this example, $\phi(1-\eta)$ has to be greater than 1 for producing to be profitable.

While it is possible that within this space of possible values of (η, ϕ), η and ϕ are still negatively correlated, it seems unlikely without a specific story. Other parameterizations of profitability would still obey the same logic: ϕ would increase profits, and η decrease them, so the maximum allowable η is increasing in ϕ.

2.8 Conclusions

This paper finds that distortions matter in a plant's decision to export and that they lead to less productivity-based "churning" of firms through exporting, and in so doing it makes three main contributions to the literature.

First, a model is developed in which general distortions can be added in a very natural manner. This model has the advantage of extending Melitz in a simple but useful way, while both encompassing and going beyond the common distortions considered in the literature such as taxes and subsidies.

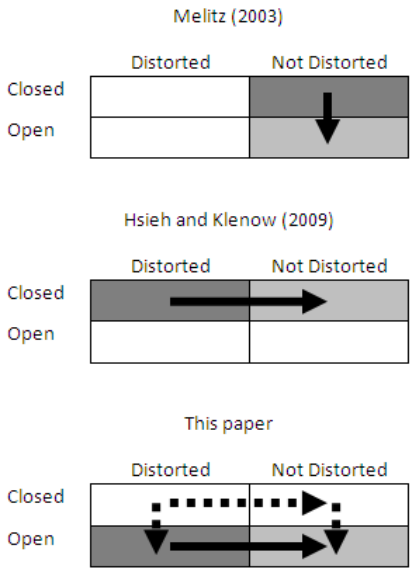
Second, the paper presents suggestive evidence that the main distortions affecting trade are not tariffs or subsidies.[12] It has long been believed that non-tariff barriers can be the greatest obstacle to trade; this paper emphasizes the possibly unintentional barriers to trade, distortions that pertain to plants producing for domestic markets, as well. Further, this paper suggests a way of estimating these other barriers to trade.

Finally, and perhaps most provocatively, the fact that these distortions lessen the importance of productivity for plants' profitability throws into doubt or weakens the results in the literature regarding the beneficial effects of trade on productivity in countries or sectors with particularly large market distortions. The flip side of this finding is that any estimates of the effects of trade on productivity that were made using data on firms facing significant distortions would understate the productivity gains from trade in less distorted markets.

In summary, this paper suggests revisiting the question of how much trade affects the productivity of firms in a country and explicitly taking distortions into account when making these kinds of estimations. The findings also imply that reducing distortions could cause particularly large productivity gains in sectors conducting trade, a result with clear policy implications.

[12] The reader will remember that subsidies, while not much discussed in the paper, were not much discussed for the very reason that they were uniformly found to be insignificant, though this could also be due to relatively few subsidies being observed.

Figure 2.1: Comparison of Models



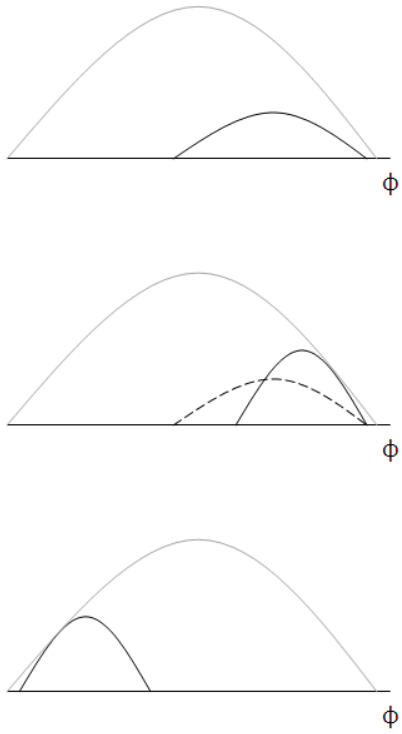


Figure 2.2: Exporter Productivity Distributions and Associated Gains from Trade
The gray lines in these diagrams represent the overall distribution of firms in the economy. In the upper-most diagram, the black line represents a plausible distribution of firms that would export in an open economy. In the middle diagram, we see that reducing distortions would shift this curve from the dashed black line to the solid black line. In the bottom diagram, the black line represents a possible set of would-be exporters under extreme distortions. If these firms gained market share by an opening to trade, productivity would decrease.

Figure 2.3: Histograms of TFPR and TFPQ

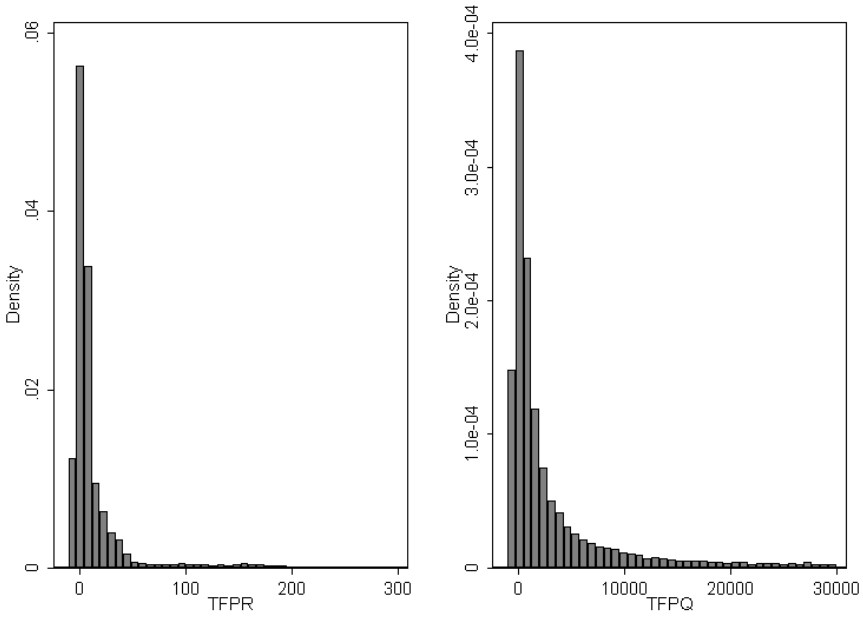


Figure 2.4: Histograms of Scaled η_y and η_k

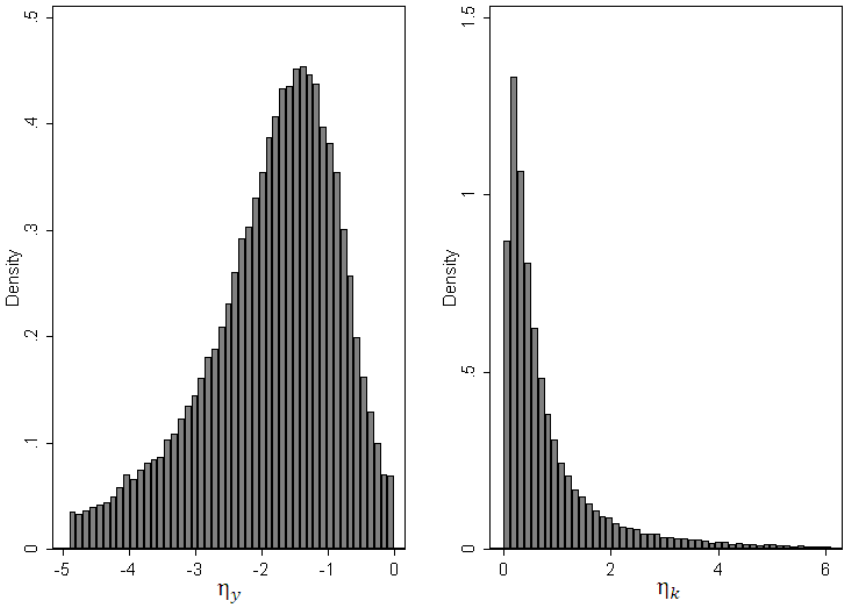


Table 2.1: Summary of η_y and η_k by Type of Plant

Type of Plant	Mean η_y	Mean η_k
State enterprises:	-1.65	0.52
Private enterprises (average):	-0.96	0.74
- Cooperative	-1.24	0.78
- De facto corporation	-1.20	1.11
- Limited partnership	-1.01	0.76
- Proprietorship	-0.98	0.72
- Joint partnership	-0.95	0.57
- Joint stock company	-0.89	1.20
- Collective	-0.87	0.96
- Corporation	-0.73	0.64

Figure 2.5: Loess Results

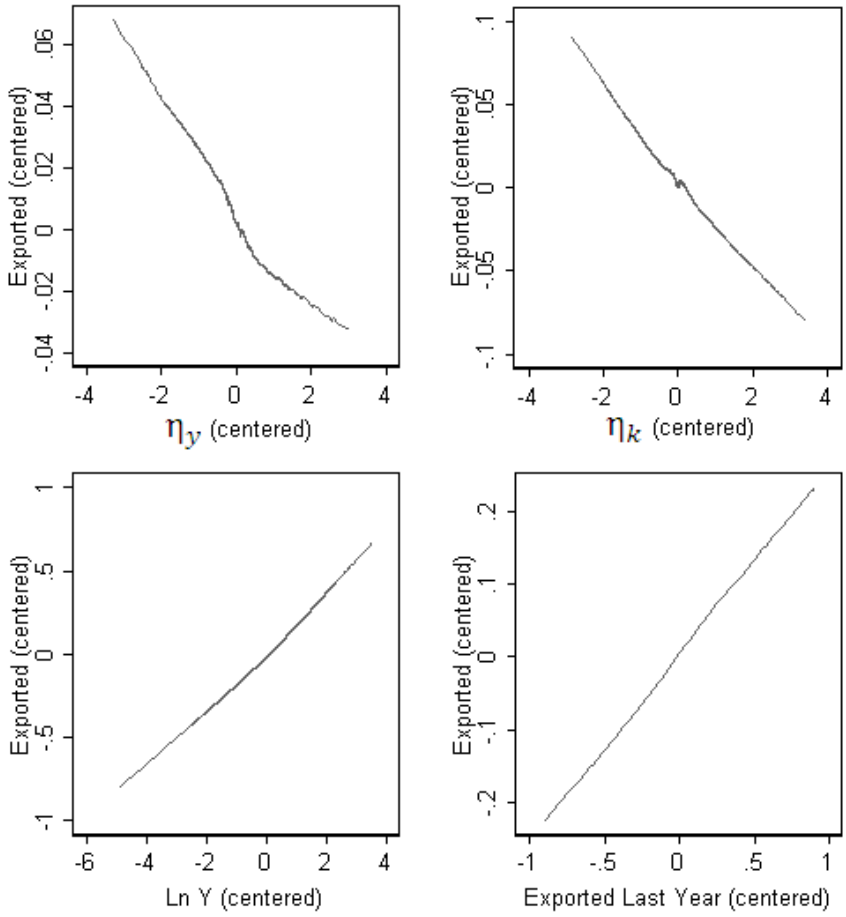


Table 2.2: Logit Regression on Whether a Plant Exports

	(1)	(2)	(3)	(4)
η_y	0.8190**	0.8478*	0.8306*	0.8604+
	(0.061)	(0.065)	(0.072)	(0.078)
η_k	0.7922***	0.7937***	0.7463***	0.7442***
	(0.051)	(0.041)	(0.058)	(0.048)
Productivity Rank	1.0578	1.0658	1.0680	1.1108
	(0.248)	(0.274)	(0.288)	(0.317)
Exported Last Year	3.2323***	2.9392***	3.1962***	2.9903***
	(0.194)	(0.147)	(0.223)	(0.145)
Ln Output	3.6935***	3.5012***	3.7061***	3.5748***
	(0.311)	(0.409)	(0.461)	(0.441)
Year	0.8835***	0.8675***	0.8838***	0.8716***
	(0.022)	(0.027)	(0.030)	(0.028)
Number of Exporters		1.0173***		1.0208***
in Sector Last Year		(0.003)		(0.003)
Ln Taxes Last Year			1.0004	0.9620***
			(0.012)	(0.010)
N	10852	10852	10429	10429

In all tables in this chapter, + p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001.

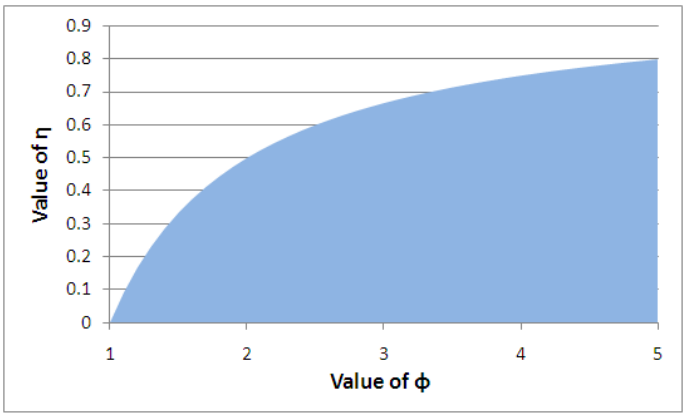
Table 2.3: Logit Regression on Whether a Plant Exports, Sector Fixed Effects

	(1)	(2)
Mean η_y	0.7811***	0.7674***
	(0.049)	(0.056)
Mean η_k	0.8454*	0.8115***
	(0.057)	(0.046)
Mean Productivity Rank	1.0482	1.1069
	(0.101)	(0.157)
Ln Mean Output	2.8211***	3.7310***
	(0.085)	(0.349)
Ln Mean Taxes		0.7779**
		(0.066)
N	13792	12974

Table 2.4: Logit Regression on Whether a Plant Exports, Particular Sectors

	(1)	(2)	(3)	(4)
η_y	0.5074*	0.4997*	0.5412*	0.5461*
	(0.139)	(0.164)	(0.137)	(0.159)
η_k	0.6795	0.6705	0.6877	0.6827
	(0.243)	(0.254)	(0.236)	(0.201)
Productivity Rank	3.3685	4.0288	3.0830	3.8122
	(6.067)	(7.493)	(3.926)	(6.742)
Exported Last Year	3.3813***	3.5571***	3.4007***	3.5276***
	(0.640)	(0.612)	(0.663)	(0.678)
Ln Output	4.4887**	4.7843***	4.3214**	4.4507***
	(2.516)	(2.048)	(1.957)	(1.831)
Year	0.8808	0.8802	0.8852	0.8830
	(0.125)	(0.102)	(0.103)	(0.090)
Number of Exporters		0.9797		0.9738
in Sector Last Year		(0.021)		(0.028)
Ln Taxes Last Year			1.0146	1.0486
			(0.053)	(0.079)
N	973	973	963	963

Figure 2.6: Example of permissible values of η for a given ϕ



Chapter 3

Welfare Gains from Innovation

3.1 Introduction

New products are believed to play an important role in increasing welfare. The beneficial effects of technological advances on productivity are obvious, but even apart from these benefits it has long been thought that people simply value some products more than others and an increase in the varieties available to them should increase welfare. Many papers have tried to estimate these welfare gains, but no study to date has provided an estimate of the gains from a comprehensive, historical set of innovations. This paper introduces a new method to conduct this exercise and presents results that indicate that more historical innovations resulted in much greater welfare gains than would be suggested by the literature to date.

Two main methods have been used in the literature to try to determine the benefits of innovations. In the first method, very detailed data is gathered about demand for a set of similar goods and this data, along with the response to the introduction of a new good, is used to estimate the gains from the new good with discrete choice models. Hausman (1997) and Bresnahan and Gordon (1997) are examples of classic works in this vein. In the second method, data across a wide set of varieties is gathered, but discrete choice models are not fit, since the goal is to estimate the returns from many new varieties, requiring far too much data. Instead, this second class of papers, notably including Broda and Weinstein (2006a, 2006b, 2010), estimates a set of elasticities of substitution between a vast array of goods using GMM and assumptions about the utility function. These are then used to obtain an estimate via effects on a price index which is explicitly constructed to take changes in varieties into account.

Both strands of the literature have made extremely important contributions. Yet no paper has yet provided an estimate that both covers a comprehensive set of goods and extends

back in time to consider more historical and basic innovations rather than smaller product tweaks and new brands. The papers examining the effects of a particular new good obviously could not estimate the welfare benefits of a broader set of innovations. And the second set of papers has focused on recent years' new varieties. Focusing on these varieties likely understates the welfare effects of new innovations since one might expect that the addition of a few minor sub-categories would not improve welfare by as much as more historical introductions of more basic categories of goods. For example, the 130th brand of paper towel probably does not have as much of an impact on human welfare as the introduction of the personal computer or satellite telecommunications. While it is true that recent work has specifically looked at different levels of categories of goods, with findings that accord to intuition[1], they do not go far enough back in time to get a comprehensive view of the effects of these more basic innovations.

The gap in the literature seems largely due to data limitations. No time-series data appears to be available for earlier time periods. This paper gets around this problem by creating a new dataset which, with several simplifying assumptions, allows us to obtain a back-of-the-envelope estimate.

A second contribution of this paper is its laying bare the mechanisms of the most commonly used price index which takes changes in varieties into account. I highlight the circumstances under which this index yields unreasonable results and adapt the methodology accordingly.

In the next section, the utility function and price index that will be used to evaluate the effects of new innovations are described. I then detail the methods that will be used to create this estimate. Following this, I describe the data and present and discuss results before concluding.

3.2 Utility Function and Price Index

I base my estimations on an adapted version of the methods used in Feenstra (1994) and Broda and Weinstein (2006a) for estimating the exact price index of a CES aggregate good. While Feenstra (1994) and Broda and Weinstein (2006a) focus on trade and differentiate between imported goods and domestically produced goods, my focus will be on all products consumed domestically within the U.S. Thus, in my model, there is no upper-level utility function aggregating imports and domestically produced goods - the utility function is simply

[1]For example, Broda and Weinstein (2006a) find that goods classified by X-digit product codes tend to be less substitutable amongst themselves than goods classified by more detailed X+Y-digit product codes, and hence the introduction of a new X-digit product would raise welfare by more than the introduction of a new X+Y-digit product.

given by:

$$M_t = \left(\sum_{g \in G} \left(\left(\sum_{v \in V} d_{gvt}^{1/\sigma_g} m_{gvt}^{(\sigma_g - 1)/\sigma_g} \right)^{\sigma_g/(\sigma_g - 1)} \right)^{(\gamma-1)/\gamma} \right)^{\gamma/(\gamma-1)} ; \gamma > 1, \sigma_g > 1 \quad \forall g \in G \quad (3.1)$$

where t indexes time, g indexes goods, v indexes varieties of goods, d is a parameter governing tastes and can also represent quality, m represents quantity, and σ_g represents the elasticity of substitution among varieties of good g. Varieties, here, represent different sub-categories within a broader product category. For example, within the good "apples", there could be a variety "Fuji apples". In the literature, varieties are defined as goods that come from a specific country (for example, "Argentinian apples" or "Argentinian Fuji apples", depending on the level of analysis). For me, however, varieties will simply be more narrowly defined sub-categories of goods, ignoring the country of origin; I use 10-digit Harmonized System codes, which are sub-categories of different 6-digit Harmonized System codes, which are in turn sub-categories of different 4-digit Harmonized System codes, and so on. While it is true that products with the same 10-digit Harmonized System code may also vary in quality or characteristics by country of origin, I ignore this additional possible layer of analysis for simplicity.

Ultimately, for my analyses I will be estimating an exact price index that takes product innovations into account. Sato (1976) and Vartia (1976) defined an exact price index that could be used with a CES utility function like the one in this paper as:

$$P_g = \Pi_{v \in I_g} \left(\frac{p_{gvt}}{p_{gvt-1}} \right)^{w_{gvt}} \quad (3.2)$$

where I_g represents the set of varieties that exist within in a good g and w represents ideal log-change weights. Cost shares $s_{gvt} \equiv \frac{p_{gvt} x_{gvt}}{\sum_{v \in I_g} p_{gvt} x_{gvt}}$ are needed in order to calculate w_{gvt}, where:

$$w_{gvt} \equiv \frac{(s_{gvt} - s_{gvt-1})/(\ln s_{gvt} - \ln s_{gvt-1})}{\sum_{v \in I_g} ((s_{gvt} - s_{gvt-1})/(\ln s_{gvt} - \ln s_{gvt-1}))} \quad (3.3)$$

Feenstra (1994) showed that this price index could be modified for use with sets of varieties that overlapped, but did not comprise the same set, across two periods. In particular, he introduced the following proposition, also used by Broda Weinstein (2006a, 2006b, 2010):

Proposition 1 *For $g \in G$, if $d_{gvt} = d_{gvt-1}$ for $v \in I_g = (I_{gt} \cap I_{gt-1}), I_g \neq \emptyset$, then the exact*

price index for good g with change in varieties is given by:

$$\pi_g = P_g \left(\frac{\lambda_{gt}}{\lambda_{gt-1}} \right)^{1/\sigma_g - 1} \tag{3.4}$$

where

$$\lambda_{gt} = \frac{\sum_{v \in I_g} p_{gvt} x_{gvt}}{\sum_{v \in I_{gt}} p_{gvt} x_{gvt}} \ , \ \lambda_{gt-1} = \frac{\sum_{v \in I_g} p_{gvt-1} x_{gvt-1}}{\sum_{v \in I_{gt-1}} p_{gvt-1} x_{gvt-1}} \tag{3.5}$$

and P_g is the conventional price index.

In words, assuming two periods, λ_{gt} represents the ratio of expenditures in period 2 on goods that are available in both periods ($I_{gt} \cap I_{gt-1}$) to the expenditures in period 2 on goods that are only available in period 2, while λ_{gt-1} represents the ratio of expenditures in period 1 on goods that are available in both periods to the expenditures in period 1 on goods that are only available in period 1. Therefore, the higher the expenditure share of new varieties, the lower λ_{gt} is and the smaller the exact price index is relative to the conventional price index.

An aggregate version of this price index would be given by:

$$\Pi = \left(\prod_{g \in G} P_g^{w_{gt}} \right) \prod_{g \in G} \left(\frac{\lambda_{gt}}{\lambda_{gt-1}} \right)^{w_{gt}/(\sigma_g - 1)} \tag{3.6}$$

Since I define varieties by product codes, I do not need an additional layer of aggregation as in Broda and Weinstein (2006a) to combine imports and domestically produced goods.

I will ultimately find that using the functional form of (6) can yield unreasonable results under certain conditions. However, I follow the literature and calculate it, before discussing the conditions under which it fails and possible solutions.

3.3 Identification Strategy

The overarching estimation method can be summarized as follows. First, elasticities of substitution are estimated using the methods outlined in subsection 3.1. I then depart from the literature and estimate the taste parameter d_{gvt} using p_{gvt}, m_{gvt} and the estimates of σ_g, in a manner I will describe in subsection 3.2. For each variety and good, I also record an approximate date of innovation. I can then construct sets of goods that were available at different historical time periods, such as in the year 1900. With these dates, and with d_{gvt} estimated, the question set out in the introduction can be answered: how would welfare be

affected if one were restricted to the set of goods available at an earlier time period? Rather than being interested in the change of varieties between some time periods $t-1$ and t, I consider the changes between t and a hypothetical $t+1$ in which the set of available varieties or goods is reduced to a subset of those available at time t; a subset determined by when the product was first created. In order to ultimately obtain an estimate of the exact price index in Equation (6) for this hypothetical time period, the share s_{gvt+1} of expenditures that would be spent on each variety in the restricted set needs to be estimated. In obtaining s_{gvt+1}, I assume that tastes d_{gvt} do not change between t and $t+1$; the evolution of p_{gct+1} is modelled as in Broda and Weinstein (2006a).

3.3.1 Estimating Elasticities of Substitution

I follow Feenstra (1994) and Broda and Weinstein (2006a) in my estimation of elasticities of substitution. The demand for each variety can be derived From Equation (1). Represented in terms of shares and changes over time, it is given by:

$$\Delta \ \ln \ s_{gvt} = \xi_{gt} - (\sigma_g - 1)\Delta \ \ln \ p_{gvt} + \epsilon_{gvt} \tag{3.7}$$

where $\xi_{gt} = (\sigma_g - 1) \ \ln \ [\phi_{gt}(d_t)/\phi_{gt-1}(d_{t-1})]$, $\phi_{gt} = \left(\sum_{v \in I_{gt}} d_{gvt}(p_{gvt})^{1-\sigma_g}\right)^{1/(1-\sigma_g)}$, and $\epsilon_{gvt} = \Delta \ \ln \ d_{gvt}$.

Still following Feenstra (1994) and Broda and Weinstein (2006a), the supply equation is given by:

$$\Delta \ \ln \ p_{gvt} = \psi_{gt} + \frac{\omega_g}{1+\omega_g}\Delta \ \ln \ s_{gvt} + \delta_{gvt} \tag{3.8}$$

where $\psi_{gt} = -\omega_g \Delta \ \ln \ E_{gt}/(1+\omega_g)$, $\omega_g \geq 0$ is the inverse supply elasticity, E_{gt} are total expenditures on good g at time t, and $\delta_{gvt} \equiv \Delta \ \ln \ \nu_{gvt}/(1+\omega_g)$ captures any random changes such as from technology. $E(\epsilon_{gvt}\delta_{gvt})$ is assumed to be 0. Similar to Feenstra (1994) and Broda and Weinstein (2006a), we can re-write Equations (7) and (8) so that ξ_{gt} and ψ_{gt} are eliminated by using first differences with respect to a reference variety k:

$$\Delta^k \ \ln \ s_{gvt} = -(\sigma_g - 1)\Delta^k \ \ln \ p_{gvt} + \epsilon^k_{gvt} \tag{3.9}$$

$$\Delta^k \ \ln \ p_{gvt} = \frac{\omega_g}{1+\omega_g}\Delta^k \ \ln \ s_{gvt} + \delta^k_{gvt} \tag{3.10}$$

where $\Delta^k x_{gvt} \equiv \Delta x_{gvt} - \Delta x_{gkt}$, $\epsilon^k_{gvt} = \epsilon_{gvt} - \epsilon_{gkt}$, and $\delta^k_{gvt} = \delta_{gvt} - \delta_{gkt}$. Multiplying Equations (9) and (10) together allow us to take advantage of the assumption that error terms are

independent across equations ($E(\epsilon^k_{gvt}\delta^k_{gvt}) = 0$):

$$(\Delta^k \ \ln \ p_{gvt})^2 = \theta_1(\Delta^k \ \ln \ s_{gvt})^2 + \theta_2(\Delta^k \ \ln \ p_{gvt}\Delta^k \ \ln \ s_{gvt}) + u_{gvt} \tag{3.11}$$

where $\theta_1 = \frac{\omega_g}{(1+\omega_g)(\sigma_g-1)}$, $\theta_2 = \frac{1-\omega_g(\sigma_g-2)}{(1+\omega_g)(\sigma_g-1)}$, and $u_{gvt} = \epsilon^k_{gvt}\delta^k_{gvt}$. σ_g and ω_g can then be estimated with the assumption that demand and supply elasticities are constant over varieties of the same good, exploiting the panel nature of the dataset. Feenstra (1994) and Broda and Weinstein (2006a) provide more details on estimating σ_g and ω_g under these assumptions using GMM.

3.3.2 Estimating the Price Index

To estimate the price index in (6) for the change between t and $t+1$, we need estimates of s_{gvt+1}, as discussed. I thus first estimate d_{gvt}, given p_{gvt} and m_{gvt}, and assume tastes do not change between t and $t+1$. I then estimate p_{gvt+1}, followed by m_{gvt+1} given d_{gvt+1} and p_{gvt+1}. This is where I depart from the literature, which does not estimate d_{gvt}, p_{gvt+1} and m_{gvt+1}.

To derive d_{gvt}, I use the fact that in equilibrium the ratio of marginal utility to prices should be constant across varieties. After much algebra, d_{gvt} can be found to be able to be obtained as:

$$d_{gvt} = \left(d_{11t}^{\left(\frac{\gamma-1}{\gamma}\right)\left(\frac{1}{\sigma_1-1}\right)} \frac{\left(\frac{\sum_{v\in g=1} p_{11t}m_{11t}}{p_{11t}(m_{11t})^{1/\sigma_1}}\right)^{\left(\frac{1}{\sigma_1-1}\right)\left(\frac{\gamma-1}{\gamma}\right)}}{\left(\frac{\sum_{v\in g} p_{gvt}m_{gvt}}{p_{gvt}(m_{gvt})^{1/\sigma_g}}\right)^{\left(\frac{1}{\sigma_g-1}\right)\left(\frac{\gamma-1}{\gamma}\right)}} \left(\frac{p_{gvt}(m_{gvt})^{1/\sigma_g}}{p_{11t}(m_{11t})^{1/\sigma_1}}\right) \right)^{(\sigma_g-1)\left(\frac{\gamma}{\gamma-1}\right)} \tag{3.12}$$

for any good g and variety v. Since utility is relative, and hence so is the taste parameter d, we can set d_{11t} (the taste parameter for the first variety of the first good) equal to 1 without loss of generality and find d_{gvt} relative to it. γ is a parameter that will have to be chosen; for the main estimation, I set $\gamma = 3$.

To estimate p_{gvt+1}, I use Equations (7) and (8), which were used to estimate the elasticities of substitution. With some work, they can be used to derive:

$$p_{gvt+1} = \left[\frac{\sum_{v_{t+1}} d_{gvt+1}(p_{gvt+1})^{1-\sigma_g}}{\sum_{v_t} d_{gvt}(p_{gvt})^{1-\sigma_g}} \right]^{\frac{1}{(\frac{1+\omega_g}{\omega_g}\frac{1}{\sigma_g-1}+1)(1-\sigma_g)}} p_{gvt}$$

With these d_{gvt+1} and p_{gvt+1}, and after more derivation, one can show that:

$$m_{gvt+1} = \frac{d_{gvt+1}}{(p_{gvt+1})^{\sigma_g}} \frac{(p_{11t+1})^{\sigma_1}}{d_{11t+1}} \frac{a_g}{a_1} m_{11t+1} \tag{3.13}$$

where $a_g = \left(\sum_{v \in g} (d_{gvt+1})^{1/\sigma_g} \left(d_{gvt+1} (p_{gvt+1})^{-\sigma_g} \right)^{(\sigma_g - 1)/\sigma_g} \right)^{(\gamma - \sigma_g)/(\sigma_g - 1)}$. m_{11t+1} can be set to 1 initially and then, once relative m_{gvt+1} are found, scaled so that the sum of all $p_{gvt+1} m_{gvt+1}$ equals total expenditures E.[2]

To summarize, I use the utility function, p_{gvt} and m_{gvt} to impute d_{gvt}. p_{gvt+1} is then imputed under assumptions about the equations governing supply and demand and that $d_{gvt+1} = d_{gvt}$. Then the utility function is again used, this time in conjunction with d_{gvt+1} and p_{gvt+1} to recover m_{gvt+1}. All the assumptions required for these imputations were made in Feenstra (1994) and Broda and Weinstein (2006a); the only new assumption is that $d_{gvt+1} = d_{gvt}$, or that tastes do not change instantaneously when we conduct the thought experiment of changing the set of available varieties.

Having used d_{gvt+1} and p_{gvt+1} to estimate the quantities m_{gvt+1} spent on the hypothetical restricted sets of goods, I can then go back and calculate s_{gvt+1} and w_{gt+1}. λ_{gt+1} and λ_{gt} are easily calculated given the assumptions and estimations,[3] and thus I can obtain the aggregate exact price index Π from equation (6) for period $t+1$.

3.4 Data

I use Feenstra's data on U.S. imports. These data represent the most comprehensive data on U.S. consumption across goods that include the price of each variety each year for a relatively long time series. Bar code data have been used elsewhere (Broda and Weinstein, 2010), but that dataset does not extend very far back in time and it contains so many highly similar goods that one would suspect the goods do not really fit the definition of a new innovation for this paper.

Instead, I use Feenstra's data, making the assumption that tastes for a variety that is imported are the same as tastes for the same variety that is domestically produced. One could imagine that within the same product category, the U.S. imports products that are slightly different, in unobserved ways, to the goods that it domestically produces. Still, I

[2]Total expenditures at the hypothetical $t+1$ are assumed to equal total expenditures at t. I keep expenditures constant so as to isolate the effect of the different sets of products. While it is true that people might choose to spend more or less depending on which products are available, the standard model from the literature does not allow savings and I will also make this simplification.

[3]λ_{gt+1} always equals 1, by assumption; this paper focuses only on the hypothetical loss of varieties from t to $t+1$.

consider, for example, demand for "tobacco refuse, except tobacco stems" or "woven glass fibers, not coloured", and so on, to be the same regardless of whether the product is imported or domestically produced. A sample of the products in my dataset is included in the Appendix. The dataset runs from 1972-2006, but due to changes in coding schemes between 1988 and 1989, I use only the portion of the data that begins in 1989, which uses 10-digit Harmonized System codes to identify products.

Innovation dates are coded using three sources: Bunch and Hellemans (1993); McNeil (1990); and Williams (1978). I use the first date that the product was commercially available. When there is doubt, the earliest possible date is used, so that later estimates are lower bounds.[4] I also assume that goods that fall into the following categories existed or had very close substitutes long before the beginning of the period of interest, to simplify coding: food products, clothing, beverages and tobacco, leather and shoes, wood products, paper products, printing, and mining. This clearly biases results downwards; there are some who focus on estimating the welfare gains from precisely the products that fall into these categories, such as coffee, tea and sugar (Hersh and Voth, 2009).

Two further methodological details associated with the use of this particular dataset will have to be discussed.

First, not every product is imported every year. In particular, some products are not imported in 2006, the last year in the dataset. This is problematic since data from the most recent year are used for some of the calculations - in particular, the quantities and prices in the last period (m_{gvt} and p_{gvt}) are used to calculate the most recent taste for the good (d_{gvt}) and impute future prices (p_{gvt+1}). For simplicity, only data on products that were available in 2006 are used, regardless of when else they were available. The full dataset, however, is first used to estimate elasticities of substitution, because the estimation of these elasticities depends on knowing the patterns of spending across all products.

The other peculiarity of the dataset is that the Harmonized System codes it uses occasionally change within it. They do not change at one time or in any pattern, and it is not trivial to match the codes since product descriptions can be ambiguous. Codes do not change to values previously used by another product, but are unique. Fortunately, using data on products that were available in 2006 eliminates any potential duplicate products. Some products' histories are cut short by this approach, but it should only result in more accurate estimates of elasticities of substitution and more conservative estimates of the welfare gains from new products as the number of new products with enough data for these estimations falls. In short, the results in this paper should truly be interpreted as lower bounds.

[4]For example, if I estimate the welfare gains from having the currently available set of goods as opposed to the set of goods that was available in 1900, and there are some goods that only became available in 1905 but are conservatively coded as having been available starting in 1895, the set of goods that would appear to have been available in 1900 in my dataset would be larger than true set was in actuality. In other words, I err on the side of over-estimating the set of goods available at any point in time.

3.5 Results

The estimates of σ_g that are obtained are summarized in Table 3.1. The 1st-99th percentile are also plotted in the histogram in Figure 3.1. The absolute highest elasticity of substitution among the 1,293 goods is 63.25, followed by 46.56.

Values of d_{gvt} can be simply thought of as scaling parameters that fit the model. Estimates of p_{gvt+1} tend not to differ much from p_{gvt}. Only about 21.8% of the prices differ at all from prices p_{gvt} when the set of varieties available in 1900 is used as the set available at a hypothetical $t+1$. This is because goods which do not see a change in varieties from t to $t+1$ do not exhibit changes in prices given that tastes d_{gvt} are assumed not to change instantaneously. Of those that do experience a change in prices, only the top 4.3% exhibit more than a 0.1% change in prices; the top 1.9% experience a price increase of at least 10%. These percents fall even lower when a later year than 1900 is used to obtain the hypothetical bundle of varieties.

λ_{gt+1} is always given by 1, as discussed. λ_{gt} ranges from essentially 0 to 1. Again, to give values with respect to the calculations involving the varieties available in 1900, very few goods exhibit $\lambda_{gt} \neq 1$, with only 3.7% having values below 0.9, 2.3% having values below 0.5, and 1.5% having values below 0.1. Despite the relatively low percent of goods having $\lambda_{gt} < 1$, the fact that λ_{gt} can take on particularly low values bodes ill for obtaining reasonable values for the price index in (6). This is not a flaw of the data but rather a flaw in the price index - it is problematic to use when λ_{gt} can approach 0, *i.e.* when a large share of varieties disappear. We will see this reflected in the results shortly, and a solution will be proposed.

Increases in the price index given by (6) can be thought of as stemming from two sources: increases in $P_g^{w_{gt+1}}$ or increases in $\left(\frac{\lambda_{gt+1}}{\lambda_{gt}}\right)^{w_{gt+1}/(\sigma_g-1)}$. Table 3.2 summarizes the values obtained for each component, for a variety of comparison years.

The values obtained for the $P_g^{w_{gt+1}}$ component are largely reasonable; however, since λ_{gt} can be so low, $\left(\frac{\lambda_{gt+1}}{\lambda_{gt}}\right)^{w_{gt+1}/(\sigma_g-1)}$ yields some very high results. If one were to aggregate across all goods using the price index specified in (6), the price index would be said to increase by a staggering 1.40 x 10^{25}% when limited to the set of varieties available in 1900, even if the most extreme 1% of the values for individual goods were winsorized. Table 3.3 provides these values for the other years considered.

Given the lengths to which I go in order to make the estimates lower bounds, the large results obtained suggest that the functional form of (6) is not appropriate. Defining the price index differently could yield more reasonable results. In particular, one should notice that results appear more reasonable for more recent time periods, when λ_{gt} would not approach zero. Yet there is also a parallel between more historical periods of time and more finely disaggregated data: both would tend to encourage low λ_{gt}. The solution would then seem to be that for more historical periods of time, one could redefine the level of aggregation so

that values of λ_{gt} remained well above 0. Tables 3.4 and 3.5 show how results would change if we defined goods by the first 4 digits of the Harmonized System code rather than by the first 6 digits.[5]

Of course, this adjustment method is *ad hoc*, but it exposes the mechanics of the price index used in the literature. I would suggest that any work that uses a variation of (6) use only the most modest estimates attainable by setting the distinction between groups to be at a low level of disaggregation, in order to have more confidence in the results. This is, of course, disappointing, as it is largely the high amounts of disaggregation that make Feenstra (1994) and Broda and Weinstein (2006a) appealing. However, higher levels of disaggregation mechanically inflate results.

Even when using the much more aggregated 4-digit product levels, it should be noted that the smallest of my estimates still dwarf the estimates in the literature, suggesting that something other than aggregation is at work given that the literature focuses on more disaggregated product levels.

The earlier innovations studied in this paper have a theoretical reason to be more important to welfare; one may think they would represent more basic innovations, have fewer substitutes, or represent whole new classes of goods. This appears to be born out in the data, with substitutability positively correlated with innovation date. Figure 3.2 shows a plot of σ_g against innovation dates for those products created post-1900 using 4-digit product codes. Since σ_g varies at the goods level, I plot them against average innovation dates for each good. The raw correlation is 0.336. A regression of innovation dates on σ_g makes the same point in Table 3.6, clustering standard errors at the goods level.[6]

Combined with the relatively large smallest estimates of Π, it does seem that the welfare gains from more historical innovations are larger than would be inferred from the literature to date which considers more recent innovations.

3.6 Conclusions

To summarize, I find evidence suggesting that the welfare gains from innovations over time are substantial, but I also find evidence that the current price index favoured by the literature can yield deeply misleading results.

The results suggest that gains from more historical innovations may be larger than gains from more recent innovations. In particular, we may think that more important innovations occurred earlier in time. The data support this intuition, with more recent innovations proving to be more substitutable.

---

[5]In Table 3.5, the single most extreme value for each of $P_g^{w_{gt+1}}$ and $\left(\frac{\lambda_{gt+1}}{\lambda_{gt}}\right)^{w_{gt+1}/(\sigma_g-1)}$ was dropped before the calculation of Π for the year 1900 results.

[6]Extreme values of σ_g, *i.e.* $\sigma_g > 50$, are dropped for these analyses. The relationship is only stronger when the values are included, but one might think these σ_g unreasonably large.

Regarding the finding that the price index yields unreasonably large results under certain circumstances, it could be argued that this problem is particular to the use of historical data - Equation (6) only poses problems through λ_{gt} when a large share of varieties "disappears", and this is more likely to happen the farther back one goes in time. However, the size of λ_{gt} is also a function of the level of aggregation of the data. The more disaggregated the data, the more of an impact the disappearance of a variety will have on λ_{gt}, all else equal. This is an unpleasant consequence of the functional form of (6); one would prefer if the gains from new innovations did not depend on the level of aggregation examined.

Future work will explore the appropriateness of different price indices in more detail, focusing on ensuring that the level of aggregation does not affect results. In the interim, one should exercise caution in applying this type of price index to data. Despite this problem, it still seems clear that innovations do have a very large effect on welfare, as results are much larger than in the literature even when using more aggregated data. Thus, the gains from more historical innovations appear to be much greater than would be suggested by estimates that focus on more recent innovations.

Table 3.1: Summary Values of σ_g

Percentile	σ_g
1	1.08
5	1.25
25	1.82
50	2.67
75	4.57
95	14.48
99	28.44

Figure 3.1: Histogram of the 1st-99th percentiles of σ_g

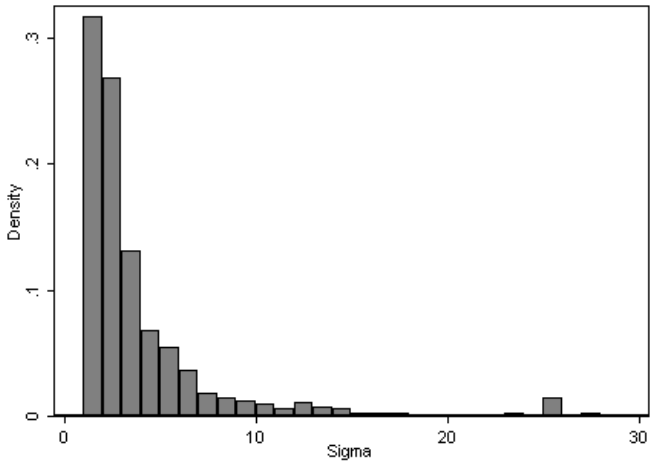


Table 3.2: Summary Values of (1) $P_g^{w_{gt+1}}$ and (2) $\left(\frac{\lambda_{gt+1}}{\lambda_{gt}}\right)^{w_{gt+1}/(\sigma_g-1)}$

	1900		1910		1920	
Percentile	(1)	(2)	(1)	(2)	(1)	(2)
1	1.01	4.38	1.00	1.25	1.00	1.12
2	1.01	1.25	1.00	1.05	1.00	1.01
3	1.00	1.06	1.00	1.00	1.00	1.00

	1930		1940		1950	
Percentile	(1)	(2)	(1)	(2)	(1)	(2)
1	1.00	1.07	1.00	1.02	1.00	1.02
2	1.00	1.00	1.00	1.00	1.00	1.00

Table 3.3: Summary Values of Π, Winsorizing $P_g^{w_{gt+1}}$ and $\left(\frac{\lambda_{gt+1}}{\lambda_{gt}}\right)^{w_{gt+1}/(\sigma_g-1)}$ at 1%

Year	Π
1900	1.40 x 10^{23}
1910	498.92
1920	24.71
1930	5.60
1940	1.51
1950	1.48

Table 3.4: Summary Values of (1) $P_g^{w_{gt+1}}$ and (2) $\left(\frac{\lambda_{gt+1}}{\lambda_{gt}}\right)^{w_{gt+1}/(\sigma_g-1)}$ at 4-digit Product Levels

	1900		1910		1920	
Percentile	(1)	(2)	(1)	(2)	(1)	(2)
1	1.00	1.13	1.00	1.02	1.00	1.12
2	1.00	1.02	1.00	1.00	1.00	1.01

	1930		1940		1950	
Percentile	(1)	(2)	(1)	(2)	(1)	(2)
1	1.00	1.02	1.00	1.00	1.00	1.00

Table 3.5: Summary Values of Π at 4-digit Product Levels

Year	Π
1900	9.23
1910	3.27
1920	1.58
1930	1.14
1940	1.05
1950	1.05

Figure 3.2: Plot of 4-digit σ_g Against Mean Innovation Date

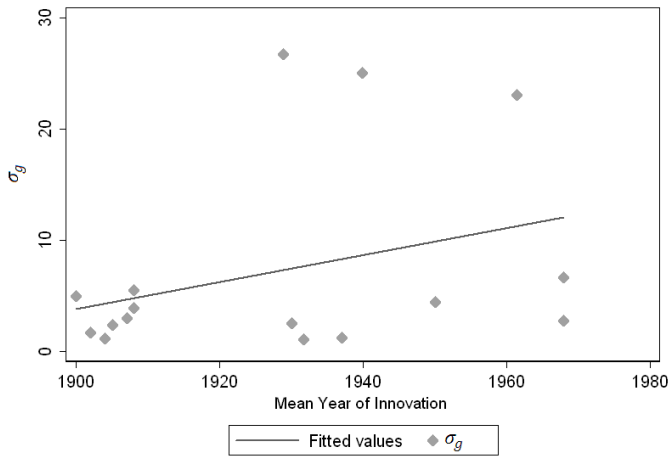


Table 3.6: Regression of σ_g on Innovation Date, Post-1900

	b/se
Innovation Date	0.348***
	(0.08)
Constant	-655.136***
	(151.78)
N	247

* p<0.05, ** p<0.01, *** p<0.001

References

Abadie, Alberto and Guido Imbens (2008). "On the Failure of the Bootstrap for Matching Estimators", Econometrica, vol. 76(6).

Abraham, Arpad and Kirk White (2006). "The Dynamics of Plant-Level Productivity in U.S. Manufacturing", mimeo.

Ackerberg, Daniel *et al.* (2006). "Structural Identification of Production Functions", mimeo.

Andrew, Christopher and Vasili Mitrokhin (2005). The World Was Going Our Way: The KGB and the Battle for the Third World. New York: Basic Books.

Angrist, Joshua (2001). "Estimation of Limited Dependent Variable Models With Dummy Endogenous Regressors: Simple Strategies for Empirical Practice", Journal of Business and Economic Statistics, vol. 19 (1).

Armington, Paul (1969). "A Theory of Demand for Products Distinguished by Place of Production", International Monetary Fund Staff Papers, vol. 16.

Aw, Bee Yan, *et al.* (2001). "Firm-Level Evidence on Productivity Differentials and Turnover in Taiwanese Manufacturing", Journal of Development Economics, vol. 66.

Aw, Bee Yan, *et al.* (2000). "Productivity and Turnover Patterns in the Export Market: Firm-Level Evidence from Taiwan and South Korea", World Bank Economic Review.

Aw, Bee Yan and Amy Hwang (1995). "Productivity and the Export Market: A Firm-Level Analysis", Journal of Development Economics, vol. 47.

Barbieri, Katherine, Omar Keshk and Brian Pollins (2008). "Correlates of War Project Trade Data Set Codebook, Version 2.0".

Bernard, Andrew and J. Bradford Jensen (2004). "Why Some Firms Export", The Review of Economics and Statistics, vol. 86 (2).

Bernard, Andrew and J. Bradford Jensen (1999a). "Exceptional Exporter Performance: Cause, Effect or Both?", Journal of International Economics, vol. 47 (1).

Bernard, Andrew and J. Bradford Jensen (1999b). "Exporting and Productivity", NBER Working Papers, no. 7135.

Bernard, Andrew *et al.* (2003). "Plants and Productivity in International Trade", American Economic Review, vol. 93 (4).

Blum, William (2004). Killing Hope: U.S. Military and CIA Interventions Since World War II. Monroe: Common Courage Press.

Boutros-Ghali, Boutros (1995). An Agenda for Peace. New York: United Nations.

Bresnahan, Timothy and Robert James Gordon (*eds.*) (1997). The Economics of New Goods. Chicago: University of Chicago Press.

Broda, Christian and David Weinstein (2010). "Product Creation and Destruction: Evidence and Price Implications", American Economic Review, vol. 100 (3).

Broda, Christian and David Weinstein (2006a). "Globalization and the Gains from Variety", Quarterly Journal of Economics, vol. 2.

Broda, Christian and David Weinstein (2006b). "From Groundnuts to Globalization: A Structural Estimate of Trade and Growth", NBER Working Papers, no. 12512.

Bunch, Bryan and Alexander Hellemans (1993). The Timetables of Technology. New York: Simon and Schuster.

Busso, Matias, DiNardo, John and Justin McCrary (2009). "New Evidence on the Finite Sample Properties of Propensity Score Matching and Reweighting Estimators", IZA Discussion Papers, #3998.

Chiozza, Giacomo and H.E. Goemans (2004). "International Conflict and the Tenure of Leaders: Is War Still *Ex Post* Inefficient?", American Journal of Political Science, vol. 48 (3).

Clerides, Sofronis, *et al.* (1998). "Is Learning-by-Exporting Important? Micro-dynamic Evidence from Colombia, Mexico and Morocco", Quarterly Journal of Economics, vol. 113.

Cochrane, Feargal (2008). Ending Wars. Cambridge: Polity Press.

Collier, Paul and Anke Hoeffler (2000). "Greed and Grievance in Civil War", World Bank Policy Research Paper, no. 2355.

Collier, Paul and Anke Hoeffler (1998). "On Economic Causes of Civil War", Oxford Economic Papers, vol. 50 (4).

Crescenzi, Mark (2007). "Reputation and Interstate Conflict", American Journal of Political Science, vol. 51 (2).

Cummingham, David and Douglas Lemke (2009). "Distinctions Without Differences? Comparing Civil and Interstate Wars", presented at the 105th Annual Meeting of the American Political Science Association, Toronto, CA. September 3-6.

De Loecker, Jan (2007). "Product Differentiation, Multi-Product Firms and Estimating the Impact of Trade Liberalization on Productivity", NBER Working Papers, no. 13155.

Diehl, Paul (2008). Peace Operations. Cambridge: Polity Press.

Diehl, Paul, Jennifer Reifschneider and Paul Hensel (1996). "United Nations Intervention and Recurring Conflict", International Organization, vol. 50 (4).

Doyle, Michael and Nicholas Sambanis (2000). "International Peacebuilding: A Theoretical and Quantitative Analysis", American Political Science Review, vol. 94 (4).

Easterly, William, Satyanath, Shanker, and Daniel Berger (2008). "Superpower Interventions and Their Consequences for Democracy: An Empirical Inquiry", NBER Working Paper Series, Working Paper 13992.

Eaton, Jonathan *et al.* (2009). "A Search and Learning Model of Export Dynamics", mimeo.

Ericson, Richard and Ariel Pakes (1995). "Markov-Perfect Industry Dynamics: A Framework for Empirical Work", Review of Economic Studies, vol. 62.

Fearon, James (1995). "Rationalist Explanations for War", International Organization, vol. 49 (3).

Fearon, James and David Laitin (2003). "Ethnicity, Insurgency and Civil War", American Political Science Review, vol. 97 (1).

Feenstra, Robert (1994). "New Product Varieties and the Measurement of International Prices", American Economic Review, vol. 84.

Fernandes, Ana (2007). "Trade Policy, Trade Volumes and Plant-Level Productivity in Colombian Manufacturing Industries", Journal of International Economics, vol. 71 (1).

Fernandes, Ana and Isgut Alberto (2005). "Learning-By-Doing, Learning-By-Exporting, and Productivity: Evidence from Colombia", Policy Research Working Paper Series, World Bank, no. 3544.

Fey, Mark and Kristopher Ramsay (2010). "Uncertainty and Incentives in Crisis Bargaining: Game-Free Analysis of International Conflict", American Journal of Political Science, forthcoming.

Fisher, Irving (1922). The Making of Index Numbers. Boston: Houghton-Mifflin.

Fortna, Virginia Page (2004). "Interstate Peacekeeping: Causal Mechanisms and Empirical Effects", World Politics, vol. 56 (4).

Fortna, Virginia Page (2008). Does Peacekeeping Work?. Princeton: Princeton University Press.

Fortna, Virginia Page and Lise Morjé Howard (2008). "Pitfalls and Prospects in the Peacekeeping Future", Annual Review of Political Science, vol. 11.

Foster, Lucia *et al.* (2008). "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?", American Economic Review, vol. 98 (1).

Gilligan, Michael and Ernest Sergenti (2008). "Do UN Interventions Cause Peace? Using Matching to Improve Causal Inference", Quarterly Journal of Political Science, vol. 3.

Gleditch, Nils *et al.* (2002). "Armed Conflict 1946-2001: A New Dataset", Journal of Peace Research, vol. 39 (5).

Griliches, Zvi and Haim Regev (1995). "Firm Productivity in Israeli Industry, 1979-1988", Journal of Econometrics, vol. 65.

Haas, Ernst, Butterworth, Lyle, and Joseph Nye (1972). Conflict Management by International Organizations. Morristown: General Learning Press.

Haddad, Mona (1993). "How Trade Liberalization Affected Productivity in Morocco", Policy Research Working Paper Series, World Bank, no. 1096.

Hausman, Jerry (1997). "Valuation of New Goods Under Perfect and Imperfect Competition", in Bresnahan and Gordon, *eds.*, The Economics of New Goods. Chicago: University of Chicago Press.

Hausman, Jerry (1983). "Specification and Estimation of Simultaneous Equation Models", in Z. Grichiles and M.D. Intriligator (*eds.*), Handbook of Econometrics, vol. 1. Amsterdam: Elsevier Science.

Hensel, Paul (1994). "One Thing Leads to Another: Recurrent Militarized Disputes in Latin America, 1816-1986", Journal of Peace Research, vol. 31 (3).

Hersh, Jonathan and Hans-Joachim Voth (2009). "New Goods in a Malthusian World: The Welfare Gains from Coffee, Tea and Sugar", http://www.voxeu.org/index.php?q=node/3936. Last accessed: April 8, 2011.

Hopenhayn, Hugo (1992). "Entry, Exit and Firm Dynamics in Long-Run Equilibrium", Econometrica, vol. 60.

Hsieh Chang-Tai and Peter Klenow (2009). "Misallocation and Manufacturing TFP in China and India", The Quarterly Journal of Economics, vol. 124 (4).

Jovanovic, Boyan (1982). "Selection and the Evolution of Industry", Econometrica, vol. 50.

Katayama, Hijime *et al.* (2009). "Firm-Level Productivity Studies: Illusions and a Solution", International Journal of Industrial Organization, vol. 27.

Kuziemko, Ilyana and Eric Werker (2006). "How Much is a Seat on the Security Council Worth? Foreign Aid and Bribery at the United Nations", Journal of Political Economy, vol. 114 (5).

Levinsohn, James and Amil Petrin (2003). "Estimating Production Functions Using Inputs to Control for Unobservables", The Review of Economic Studies, vol. 70 (2).

Luttmer, Erzo (2007). "Selection, Growth and the Size Distribution of Firms", Quarterly Journal of Economics, vol. 122 (3).

Malone, David (2000). "Eyes on the Prize: The Quest for Nonpermanent Seats on the U.N. Security Council", Global Governance, vol. 6 (1).

Mattes, Michaela and Burcu Savun (2010). "Information, Agreement Design, and the Durability of Civil War Settlements", American Journal of Political Science, vol. 52 (2).

McNeil, Ian (*ed.*) (1990). An Encyclopaedia of the History of Technology. New York: Routledge.

Melitz, Marc (2003). "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity", Econometrica, vol. 71 (6).

Miguel, Edward, Shanker Satyanath and Ernest Sergenti (2004). "Economic Shocks and Civil Conflict: An Instrumental Variables Approach", Journal of Political Economy, vol. 112 (4).

Olley, G. Steven and Ariel Pakes (1996). "The Dynamics of Productivity in the Telecommunications Equipment Industry", Econometrica, vol. 64 (6).

Pavcnik, Nina (2002). "Trade Liberalization, Exit and Productivity Improvements: Evidence from Chilean Plants", Review of Economic Studies, vol. 69.

Powell, Robert (2002). "Bargaining Theory and International Conflict", Annual Review of Political Science, vol. 5.

Rauch, James and Joel Watson (2003). "Starting Small in an Unfamiliar Environment", International Journal of Industrial Organization, vol. 21 (7).

Roberts, Mark and James Tybout (1997). "The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs", American Economic Review, vol. 87 (4).

Sato, Kazuo (1976). "The Ideal Log-Change Index Number", Review of Economics and Statistics, vol. 66.

Slantchev, Branislav (2004). "How Initiators End Their Wars: The Duration of Warfare and the Terms of Peace", American Journal of Political Science, vol. 48 (4).

Tybout, James and M. Daniel Westbrook (1995). "Trade Liberalization and Dimensions of Efficiency Change in Mexican Manufacturing Industries", Journal of International Economics, vol. 31.

UCDP/PRIO (2009). "UCDP/PRIO Armed Conflict Dataset Codebook, Version 4-2009". http://www.pcr.uu.se/research/UCDP/data_and_publications/datasets.htm.

Van Beveren, Ilke (2010). "Total Factor Productivity Estimation: A Practical Review", Journal of Economic Surveys.

Vartia, Yrjo (1976). "Ideal Log-Change Index Numbers", Scandinavian Journal of Statistics, vol. 3.

Weiner, Tim (2007). Legacy of Ashes: The History of the CIA. New York: Doubleday.

Werner, Suzanne (1999). "The Precarious Nature of Peace: Resolving the Issues, Enforcing the Settlement and Renegotiating the Terms", American Journal of Political Science, vol. 43.

Williams, Trevor (1978). A History of Technology, vol. 1-6. Oxford: Clarendon Press.

Wilkenfeld, Jonathan and Michael Brecher (1984). "International Crises, 1945-1975: The UN Dimension", International Studies Quarterly, vol. 28 (1).

Wooldridge, Jeffrey (2009). "On Estimating Firm-Level Production Functions Using Proxy Variables to Control for Unobservables", Economic Letters, vol. 104 (3).

Wooldridge, Jeffrey (2007). "What's New in Econometrics?", NBER Summer Institute.

Wooldridge, Jeffrey (2002). Econometric Analysis of Cross Section and Panel Data. Cambridge: The MIT Press.

Appendix

Table A-1: Data Summary

Country	Years Peacekeepers Were Sent	Years in Conflict	Years as Temporary Members of the S.C.
Afghanistan	1988-1990	1978-2001, 2003-	
Albania		1946	
Algeria		1954-1963, 1991-	1968-1969, 1988-1989, 2004-2005
Angola	1988-1999	1961-2002, 2004, 2007-	2003-2004
Argentina		1955, 1963, 1974-1977, 1982	1948-1949, 1959-1960, 1966-1967, 1971-1972, 1987-1988, 1994-1995, 1999-2000, 2005-2006
Azerbaijan		1992-1995, 2005	
Bangladesh		1975-1992	1979-1980, 2000-2001
Bolivia		1946, 1952, 1967	1964-1965, 1978-1979
Bosnia and Herzegovina	1992-2002	1992-1995	
Brunei		1962	
Burkina Faso		1985, 1987	1984-1985, 2008
Burundi	2004-2007	1965, 1991-1992, 1994-2006, 2008-	1970-1971
Cambodia		1946-1953, 1966-1998	
Cameroon		1957-1961, 1984, 1996	1974-1975, 2002-2003
Central African Republic	1998-2000, 2007-	1996-1997, 2001-2002, 2006	
Chad	1994, 2007-	1966-1972, 1976-1984, 1986-1987, 1989-1994, 1997-2002, 2005-	
Chile		1973	1952-1953, 1961-1962, 1996-1997, 2003-2004
China		1946-1950, 1954, 1956, 1958-1959, 1962, 1967, 1969, 1978-1981, 1983-1984, 1986-1988	
Colombia		1964-	1947-1948, 1953-1954, 1957-1958, 1969-1970, 1989-1990, 2001-2002
Comoros		1989, 1997	
Congo		1993-1994, 1997-1999 2002	1986-1987
Costa Rica		1948	1974-1975, 1997-1998, 2008
Cote D'Ivoire	2004-	2002-2004	
Croatia	1992-1998, 1996-2002	1992-1993, 1995	2008

Cuba		1953, 1956-1958, 1961	1949-1950, 1956-1957, 1990-1991
Cyprus	1964-	1955-1959, 1974	
Democratic Republic of Congo	1960-1964, 1999	1960-1962, 1964-1965, 1967, 1977-1978, 1996-2001, 2006-	
Djibouti		1991-1994, 1999, 2008	1993-1994
Dominican Republic	1965-1966	1965	
Ecuador		1995	1950-1951, 1960-1961, 1991-1992
Egypt	1956-1967	1951-1952, 1956, 1993-1998	1946, 1949-1950, 1984-1985, 1996-1997
El Salvador	1991-1995	1969, 1972, 1979-1991	
Equatorial Guinea		1979	
Eritrea	2000-2008	1997-2000, 2003	
Ethiopia		1960, 1964-1992, 1994-1996, 1998-	1967-1968, 1989-1990
France		1961-1962	
Gabon		1964	1978-1979, 1998-1999
Gambia		1981	1998-1999
Georgia	1993-2009	1991-1993, 2004, 2008-	
Ghana		1966, 1981, 1983	1962-1963, 1986-1987, 2006-2007
Greece		1946-1949	1952-1953, 2005-2006
Grenada		1983	
Guatemala	1997	1949, 1954, 1965-1995	
Guinea		2000-2001	1972-1973, 2002-2003
Guinea-Bissau		1963-1973, 1998-1999	1996-1997
Haiti	1993-2001, 2004-	1989, 1991, 2004	
Honduras		1957	1995-1996
India	1949-	1947-1951, 1956-1959, 1961-1971, 1978-	1950-1951, 1967-1968, 1972-1973, 1977-1978, 1984-1985, 1991-1992
Indonesia	1999-2005	1946-1950, 1953, 1958-1962, 1965, 1967-1969, 1975-1992, 1997-2005	1973-1974, 1995-1996, 2007-2008
Iran	1988-1991	1946, 1966-1968, 1974, 1979-1988, 1990-1993, 1996-1997, 1999-2001, 2005-	1955-1956
Iraq	1991-2003	1958-1959, 1961-1970, 1973-1996, 2003-	1957-1958, 1974-1975
Kenya		1952-1956, 1982	1973-1974, 1997-1998

Laos		1946-1953, 1959-1961, 1963-1973, 1986-1990	
Lebanon	1958, 1978-	1958, 1975-1976, 1982-1986, 1989-1990	1953-1954
Lesotho		1998	
Liberia	1993-1997, 2003-	1980, 1989-1995, 2000-2003	1961
Macedonia		2001	
Madagascar		1947, 1971	1985-1986,
Malaysia		1948-1960, 1963-1966, 1974-1975, 1981	1965, 1989-1990, 1999-2000
Mali		1990, 1994, 2007-	1966-1967, 2000-2001
Mauritania		1957-1958, 1975-1978	1974-1975, 1977-1978
Mexico		1994, 1996	1946, 1980-1981 2002-2003
Moldova		1992	
Morocco	1991-	1953-1958, 1971, 1975-1989	1963-1964, 1992-1993
Mozambique	1992-1994	1964-1974, 1977-1992	
Myanmar		1948-2003, 2005-	
Nepal		1960-1962, 1996-2006	1969-1970, 1988-1989
Nicaragua	1989-1992	1978-1979, 1981-1989	1970-1971, 1983-1984
Niger		1991-1992, 1994, 1996-1997, 2007-	1980-1981
Nigeria		1966-1970, 2004	1966-1967, 1978-1979, 1994-1995
North Korea		1949-1953	
North Yemen	1963-1964	1948, 1962-1970, 1972, 1978-1982	
Oman		1957, 1972-1975	1994-1995
Pakistan		1971, 1974-1977, 1990, 1995-1996, 2004-	1952-1953, 1968-1969, 1976-1977, 1983-1984, 1993-1994, 2003-2004
Panama		1989	1958-1959, 1972-1973, 1976-1977, 1981-1982, 2007-2008
Papua New Guinea		1989-1990, 1992-1996	
Paraguay		1947, 1954, 1989	1968-1969
Peru		1965, 1982-1999, 2007-	1955-1956, 1973-1974, 1984-1985, 2006-2007
Philippines		1946-1954, 1969-	1957-1963, 1980-1981, 2004-2005
Romania		1989	1962, 1976-1977, 1990-1991, 2004-2005
Russia (Soviet Union)		1946-1950, 1956, 1990-1991, 1993-1996, 1999-	

Rwanda	1993-1996	1990-1994, 1997-2002	1994-1995
Saudi Arabia		1979	
Senegal		1990, 1992-1993, 1995, 1997, 2000-2001, 2003	1968-1969, 1988-1989
Sierra Leone	1998-2005	1991-2000	1970-1971
Somalia	1992-1995	1978, 1982-1984, 1986-1996, 2001-2002, 2006-	1971-1972
South Africa		1966-1988	2007-2008
South Vietnam		1955-1964	
South Yemen		1964-1967, 1986	
Spain		1980-1981, 1987, 1991-1992	1969-1970, 1981-1982, 1993-1994, 2003-2004
Sri Lanka (Ceylon)		1971, 1984-2001, 2003, 2005-	
Sudan	2005-	1963-1972, 1976, 1983-	1972-1973
Surinam		1986-1988	
Syria		1966, 1979-1982	1947-1948, 1970-1971, 2002-2003
Tajikistan	1994-2000	1992-1996, 1998	
Tanzania		1978	1975-1976, 2005-2006
Thailand		1946, 1951, 1974-1982, 2003-	1985-1986
Togo		1986, 1991	1982-1983
Trinidad and Tobago		1990	1985-1986
Tunisia		1953-1956, 1961, 1980	1959-1960, 1980-1981, 2000-2001
Turkey		1984-	1951-1952, 1954-1955, 1961
Uganda	1993-1994	1971-1972, 1974, 1978-1992, 1994-2007	1966, 1981-1982
United Kingdom		1971-1991, 1998	
United States of America		1950, 2001-2002, 2004-	
Uruguay		1972	1965-1966
Uzbekistan		1999-2000, 2004-	
Venezuela		1962, 1982, 1992	1962-1963, 1977-1978, 1986-1987, 1992-1993
Vietnam		1946-1954, 1965-	2008
Yemen		1994	1990-1991
Yugoslavia		1991	1950-1951, 1956, 1972-1973, 1988-1989
Zimbabwe		1967-1968, 1973-1979	1983-1984, 1991-1992

"Years in Conflict" is not identical to the episodes of conflict. For example, if a country has one episode of conflict from 1980-1984 and a different conflict occurs from 1983-1990, that country will be listed as being in conflict from 1980-1990 under "Years in Conflict".

Table A-2: Tobit Estimation of the Effects of Being Sent Peacekeepers on the Duration of Peace once Peace is Obtained

	(1) b/se	(2) b/se	(3) b/se
Duration of Peace			
Peacekeepers Sent	1.710	1.383	1.963
	(2.584)	(2.710)	(2.568)
Episode Duration	-0.207	-0.213	-0.173
	(0.154)	(0.156)	(0.162)
Year Episode Ended	-0.419**	-0.432***	-0.401**
	(0.165)	(0.161)	(0.164)
Battle Deaths	5.919*	5.727	5.589*
	(3.546)	(3.485)	(3.353)
Years on Security Council		-0.859	-0.703
in Last 5 Years		(1.554)	(1.539)
Years on Security Council		-0.469	0.001
in Last 10 Years		(1.003)	(1.015)
Non-Military Aid			-11.492
			(10.731)
Military Aid			11.822*
			(7.118)
Observations	184	184	184

Table A-3: Tobit Estimation of the Effects of Being Sent Peacekeepers on the Duration of Peace once Peace is Obtained, Control Function Approach

	(1) b/se	(2) b/se	(3) b/se
Duration of Peace			
Peacekeepers Sent	20.737**	26.158**	17.770**
	(8.384)	(10.094)	(8.830)
Residual	-11.237**	-14.480***	-9.496*
	(4.688)	(5.372)	(4.841)
Episode Duration	-0.417***	-0.493***	-0.291*
	(0.151)	(0.166)	(0.162)
Year Episode Ended	-0.482***	-0.516***	-0.407**
	(0.153)	(0.153)	(0.159)
Battle Deaths	3.841	3.770	3.737
	(2.686)	(2.487)	(2.507)
Years on Security Council in Last 5 Years		-2.382	-1.614
		(1.517)	(1.517)
Years on Security Council in Last 10 Years		2.059	2.149
		(1.351)	(1.530)
Non-Military Aid			-24.962**
			(11.865)
Military Aid			19.043**
			(8.476)
Observations	184	184	184

Table A-4: OLS Estimation of the Effects of Being Sent Peacekeepers on Episode Duration

	(1) b/se	(2) b/se	(3) b/se
Episode Duration			
Peacekeepers Sent	1.893	1.962*	1.315
	(1.225)	(1.162)	(1.208)
Year Episode Ended	-0.100**	-0.109**	-0.140***
	(0.044)	(0.046)	(0.052)
Battle Deaths	2.647***	2.662***	2.640***
	(0.497)	(0.487)	(0.452)
Years on Security Council		0.562	-0.079
in Last 10 Years		(0.553)	(0.524)
Years on Security Council		-1.249	-1.105*
in Last 5 Years		(0.770)	(0.655)
Non-Military Aid			12.341**
			(5.226)
Military Aid			-5.026*
			(3.010)
Observations	184	184	184
R^2	0.16	0.17	0.21

Table A-5: 2SLS Estimation of the Effects of Being Sent Peacekeepers on the Duration of the Episode of Conflict

	(1) b/se	(2) b/se	(3) b/se
Peacekeepers Sent (First Stage)			
Percent Years on Security Council	-0.250***	-0.170**	-0.242***
During Episode	(0.081)	(0.084)	(0.087)
Battle Deaths	0.053	0.043	0.037
	(0.061)	(0.061)	(0.059)
Year Episode Ended	-0.002	-0.002	-0.005
	(0.006)	(0.006)	(0.006)
Years on Security Council		0.038	0.042*
in Last 5 Years		(0.033)	(0.024)
Years on Security Council		-0.093***	-0.138***
in Last 10 Years		(0.033)	(0.036)
Non-Military Aid			1.031**
			(0.450)
Military Aid			-0.512**
			(0.220)
Duration of Episode			
Peacekeepers Sent	-5.308	-8.493	-0.903
	(6.140)	(10.294)	(3.845)
Battle Deaths	3.004***	3.063***	2.645***
	(0.484)	(0.671)	(0.367)
Year Episode Ended	-0.189***	-0.223***	-0.267***
	(0.069)	(0.081)	(0.063)
Years on Security Council		-1.213	-1.402*
in Last 5 Years		(1.000)	(0.733)
Years on Security Council		-0.353	-0.405
in Last 10 Years		(1.057)	(0.679)
Non-Military Aid			16.985***
			(4.862)
Military Aid			-7.765**
			(3.275)
Observations	174	174	174
First Stage F-stat	7.49	4.89	7.83
p-values of tests:			
Anderson-Rubin Wald test	0.33	0.27	0.81
Stock-Wright LM S	0.21	0.16	0.81

Table A-6: Logit Regression on the Entry of Plants

	(1)	(2)	(3)	(4)
η_y	0.8946	0.9316	0.8967	0.9268
	(0.087)	(0.072)	(0.076)	(0.092)
η_k	0.7894**	0.7888**	0.7219**	0.7136***
	(0.064)	(0.072)	(0.072)	(0.064)
Productivity Rank	1.1515	1.1155	1.1060	1.1064
	(0.416)	(0.309)	(0.334)	(0.294)
Ln Output	3.3259***	3.1355***	3.2292***	3.1093***
	(0.433)	(0.445)	(0.463)	(0.457)
Year	0.9916	0.9683	0.9946	0.9781
	(0.043)	(0.041)	(0.037)	(0.042)
Number of Exporters		1.0174***		1.0202***
in Sector Last Year		(0.003)		(0.004)
Ln Taxes Last Year			1.0145	0.9738+
			(0.012)	(0.014)
N	6976	6976	6699	6699

Table A-7: Logit Regression on the Exit of Plants

	(1)	(2)	(3)	(4)
η_y	1.0720	1.0499	1.0499	1.0278
	(0.135)	(0.116)	(0.121)	(0.111)
η_k	1.0818	1.0772+	1.0768+	1.0720
	(0.063)	(0.046)	(0.046)	(0.064)
Productivity Rank	0.7849	0.8235	0.6444	0.6813
	(0.478)	(0.526)	(0.452)	(0.463)
Ln Output	0.3192***	0.3330***	0.3011***	0.3163***
	(0.062)	(0.065)	(0.059)	(0.070)
Year	1.2719***	1.2914***	1.2816***	1.3019***
	(0.068)	(0.071)	(0.074)	(0.086)
Number of Exporters		0.9811**		0.9823**
in Sector Last Year		(0.007)		(0.006)
Ln Taxes Last Year			1.0094	0.9970
			(0.098)	(0.110)
N	2098	2098	2013	2013

Table A-8: Logit Regression on Whether a Plant Exports, $1 - \alpha_s = 0.6$

	(1)	(2)	(3)	(4)
η_y	0.8254*	0.8582*	0.8274*	0.8605+
	(0.062)	(0.061)	(0.070)	(0.071)
η_k	0.8039***	0.8068**	0.7690***	0.7695***
	(0.043)	(0.057)	(0.059)	(0.044)
Productivity Rank	1.0283	1.0359	1.0473	1.0863
	(0.280)	(0.234)	(0.266)	(0.263)
Exported Last Year	3.2306***	2.9379***	3.1928***	2.9885***
	(0.160)	(0.216)	(0.194)	(0.208)
Ln Output	3.6501***	3.4506***	3.7050***	3.5606***
	(0.427)	(0.413)	(0.458)	(0.532)
Year	0.8856***	0.8700***	0.8834***	0.8719***
	(0.027)	(0.030)	(0.031)	(0.032)
Number of Exporters		1.0173***		1.0207***
in Sector Last Year		(0.002)		(0.003)
Ln Taxes Last Year			1.0004	0.9624**
			(0.011)	(0.012)
N	10852	10852	10429	10429

Table A-9: Data Excerpt Showing Harmonized System Product Definitions

H.S. Code	Description
0601101500	TULIP BULBS, DORMANT
0601103000	HYACINTH BULBS, DORMANT
0601104500	LILY BULBS, DORMANT
0601106000	NARCISSUS BULBS, DORMANT
0601107500	CROCUS CORMS, DORMANT
0601108500	LILY OF THE VALLEY PIPS, DORMANT
0601109020	IRIS BULBS, DORMANT
0601109040	GLADIOLUS CORMS, DORMANT
0601109060	BEGONIA (TUBEROUS) TUBERS, DORMANT
0602902000	ORCHID PLANTS, LIVE
0602903010	CHRYSANTHEMUMS WITH SOIL ATTACHED TO ROOT
0602903090	HERBACEOUS PERENNIALS NESOI SOIL ATTACHED TO ROOTS
0602904000	HERBACEOUS PERENNIALS, NESOI
0602905000	MUSHROOM SPAWN
0602906010	TREES AND SHRUBS WITH SOIL ATTACHED TO ROOTS NESOI
0602906020	POINSETTIAS, LIVE, WITH SOIL ATTACHED TO ROOTS
0602906090	PLANTS, LIVE, WITH SOIL ATTACHED TO ROOTS, NESOI
0602909010	TREES AND SHRUBS LIVE WITHOUT SOIL ATTACHED, NESOI
0602909090	PLANTS, LIVE, WITHOUT SOIL ATTACHED TO ROOTS NESOI
0603103000	MINIATURE (SPRAY) CARNATIONS, FRESH
0603106010	SWEETHEART ROSES, FRESH
0603106030	SPRAY ROSES, FRESH
0603106060	ROSES, NESOI, FRESH
0603107010	POM POM CHRYSANTHEMUMS, FRESH
0603107020	CHRYSANTHEMUMS EXCEPT POM POM, FRESH
0603107040	ANTHURIUMS, FRESH
0603107050	DENDROBIUM ORCHIDS, FRESH
0603107060	ORCHIDS EXCEPT DENDROBIUM, FRESH
0603108010	ALSTROEMERIA, FRESH
0603108020	GYPSOPHILA, FRESH
0603108030	LIILIES, FRESH
0603108040	SNAPDRAGONS, FRESH

NESOI = Not Elsewhere Specified or Included