THE EFFECT OF THE 2018 TARIFFS ON EUROPEAN WINE

by

Henry Johnson

A thesis

submitted in partial fulfillment

of the requirements for the degree of

Master of Science in Economics

Boise State University

December 2022

© 2022

Henry Johnson

ALL RIGHTS RESERVED

BOISE STATE UNIVERSITY GRADUATE COLLEGE

DEFENSE COMMITTEE AND FINAL READING APPROVALS

of the thesis submitted by

Henry Johnson

Thesis Title: The Effect of the 2018 Tariffs on European Wine

Date of Final Oral Examination: 19 April 2022

The following individuals read and discussed the thesis submitted by student Henry Johnson, and they evaluated the student's presentation and response to questions during the final oral examination. They found that the student passed the final oral examination.

Michail Fragkias Ph.D.	Chair, Supervisory Committee
Samia Islam Ph.D.	Member, Supervisory Committee
Kelly Chen Ph.D.	Member, Supervisory Committee
Geoffrey Black Ph.D.	Member, Supervisory Committee

The final reading approval of the thesis was granted by Michail Fragkias Ph.D., Chair of the Supervisory Committee. The thesis was approved by the Graduate College.

DEDICATION

To everyone out there who believes U.S. wine prices are too damn high.

ACKNOWLEDGMENT

I want to say thank you to Michail Fragkias for providing guidance in this research and making recommendations for looking at different models, and to Christopher Daigle for listening to me as I rant and complain about various bugs and difficulties in the modeling. I also want to thank Vanessa Lee and Terri Baston at the USITC for helping me get additional datapoints that aren't readily available online. And lastly, to Don Holley for telling me to go to graduate school.

ABSTRACT

This paper estimates a vector autoregression model for average wine prices across U.S. cities to assess the impact of tariff changes on the U.K., France, Germany, and Spain after they were enacted in October 2019. It uses impulse response functions to gauge how a one-unit impulse in the per-liter duty rate may effect the average wine price in the U.S. and the quantity of wine from various exporters to the U.S. It finds that a one-unit impulse in the duty rate levied against the bloc of countries impacted by the tariff results in a fall in the quantity of wine imported from those countries and that wine from the bloc of countries is substituted with wine from the top three exporters not included in the bloc.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1:
INTRODUCTION

Since the 18th century, the world has seen trade liberalization like never before. It is nearly beyond question to economists that this has had a significant impact on global prosperity. After all, the benefits of comparative advantage have been apparent since David Ricardo's seminal work, The Principles of Political Economy and Taxation (1821).

In the late 20th century, we saw a continuation of this trend with many new trade agreements. The U.S. created the North American Free Trade Agreement; the European Union created a single internal market; South America saw the Mercosur; countries in Asia organized the ASEAN Free Trade Agreement. It looked like international trade would be a force to be reckoned with into the 21st century.

As with many trends, there is variation around the trend line. In the U.S., the question of whether international trade is good for the economy has returned to take center stage. Politicians and pundits in both mainstream political parties proclaim that fair trade is more important than free trade.

The notion that international trade does not treat the United States fairly is so widespread that two of the U.S.'s 2016 presidential front-runners (Donald Trump and Bernie Sanders) held this belief—despite being on different sides of the political aisle. William Nordhaus pointed out that this embrace of protectionism was outside of the historical norm of Republican candidates (2018). Regardless of this, the Republican candidate took office.

During the Trump Administration, the U.S. saw several instances of trade restrictions. These restrictions came with various justifications, from national security to the resolution of World Trade Organization lawsuits. These justifications were not always convincing (Krugman, 2020).

In the fashion of Ricardo (1821), this paper looks at international trade of wine. The wine industry was simply an innocent bystander in these trade disputes. In the early 2000s, the U.S. sued the E.U. in the WTO over subsidies provided to their aircraft industry. And, after 15 years, in October 2019, the U.S. won the right to place an additional 25% tariff on $7.5 billion of goods from select European communities (U.S. v. France, Spain, Germany, U.K., 2019) and decided to place the tariff on wine. Wine is a large import from Germany, France, and Spain.

This paper intends to look at the effect of the additional 25% tariff on wine from these countries and uses a multivariate time series analysis to predict the average wine price in U.S. cities and try to weed out the effect of the tariff on domestic wine prices.

CHAPTER 2:

TRADE THEORY

In this chapter, we are going to discuss the theoretical foundations and prevailing views surrounding international trade. In the first section, we will quickly cover early justifications for trade liberalization in Adam Smith and David Ricardo, and then in the last two sections we will construct a general theory of trade with respect to tariffs. This more modern exposition is tailored from a presentation in Feenstra (2015).

2.1 Early Justifications for Free Trade

The theoretical underpinnings of international trade date back at least as far as Adam Smith's The Wealth of Nations (Smith, 1776). Smith wrote at a time dominated by mercantilism, a system aimed at increasing exports and hoarding precious metals, with adherents believing that by doing so a nation's wealth would increase. Smith asserted that the increase in wealth was instead by increases in trade expanding the division of labor and the quantity of goods. Following Smith's work, David Ricardo wrote more on the problematic effects of trade restrictions. His most influential contribution is his outlay of comparative advantage (pp. 90-91, 1821):

> England may be so circumstanced, that to produce the cloth may require the labour of 100 men for one year; and if she attempted to make the wine, it might require the labour of 120 men for the same time. England

> would therefore find it her interest to import wine, and to purchase it by the exportation of cloth.
>
> To produce the wine in Portugal, might require only the labour of eighty men for one year, and to produce the cloth in the same country, might require the labour of ninety men for the same time. It would therefore be advantageous for her to export wine in exchange for cloth. This exchange might even take place, notwithstanding that the commodity imported by Portugal could be produced there with less labour than in England.

In this example, despite Portugal being superior in producing both wine and cloth, it is in the interest of both Portugal and England for Portugal to produce wine and England to produce cloth and then exchange wine for cloth. This is because England is relatively superior at producing cloth than wine. Both of the countries in this example are made better off by this arrangement. This process, applied to the multiplicity of goods brought to market, increases not only the quantity of goods available, but also the diversity of desires that can be fulfilled.

2.2 Social Welfare of Trade

For our trade model, we will first assume that we have maximizing consumers with the utility function $\vec{c_0} + \vec{U}\vec{c}$ subject to $\vec{c_0} + p'\vec{c} \leq \vec{I}$, with $\vec{U}$ increasing and concave and $\vec{c}$ representing a vector of all consumers and the set available goods to each of them, other than the numeraire good, which is captured by $\vec{c_0}$. The numeraire good is the equivalent of one unit of labor. Also, let us define the optimal consumption bundle as $\vec{d}(p) = \vec{c}$, so that all leftover income for consumers is spent on the numeraire good; $\vec{c_0} = \vec{I} - p'\vec{d}(p)$. With this construction, social welfare is then the sum of each individual's welfare

$$W(p, I) \equiv \sum \left(\vec{I} - p'\vec{d}(p) + \vec{U}(\vec{d}(p))\right) \tag{2.1}$$

where $I = \sum \vec{I}$. The indirect utility function of the economy is then

$$\frac{\partial W}{\partial p} = \sum -\vec{d}(p) \equiv -d(p) \tag{2.2}$$

Let us assume only one good in the economy is effected by an import tariff. And the world price of this good is p_w. The domestic price is then $p = p_w + t$, where t is the amount of the tariff.

Now, let us add in the output of the good impacted by the tariff. Let y be the output of the good subject to the tariff and $C(y)$ and $C'y$ represent the total and marginal costs of production, respectively. We will also assume the only production input is labor and that total income equals labor supply. Lastly, we will assume all revenues from the tariff are returned to the consumers. The welfare function considering the tariff is then

$$W\left(p, I + tm + py + C(y)\right) \equiv W(t) \tag{2.3}$$

The change in welfare with respect to the tariff is then

$$\frac{\partial W}{\partial t} = -d(p)\frac{\partial p}{\partial t} + \left(t\frac{\partial m}{\partial p} + y\right)\frac{\partial p}{\partial t} + \left(p - c'(y)\right)\frac{\partial y}{\partial t} \tag{2.4}$$

Since imports, m, are the output of the good the tariff is levied against less the optimal domestic consumption bundle, we can rewrite the equation as

$$\frac{\partial W}{\partial t} = m\left(1 - \frac{\partial p}{\partial t}\right) + t\frac{\partial m}{\partial p}\frac{\partial p}{\partial t} + (p - c'(y))\frac{\partial y}{\partial t} \tag{2.5}$$

The domestic price is the world price plus the tariff, $p = p_w + t$, so we know that $1 - \frac{\partial p}{\partial t} = -\frac{\partial p_w}{\partial t}$. Under a perfect competition regime, price equals marginal cost, so the last term of equation 2.5 becomes 0.

$$\frac{\partial W}{\partial t} = t\frac{\partial m}{\partial p}\frac{\partial p}{\partial t} - m\frac{\partial p_w}{\partial t} \tag{2.6}$$

The change in welfare due to the tariff is captured by 2.6. The first term is the efficiency cost or deadweight loss to society and the second term is the effect of the tariff on the world price or the terms of trade faced by the country.

2.3 Implication of the Tariff

The effects of the tariff on social welfare mostly depend on the second term in equation 2.6. If the tariff has no impact on the world price, the derivative of $\frac{\partial p_w}{\partial t} = 0$ and the second term drops out. In this scenario, the country imposing the tariff has no impact on world price. And the resulting welfare for the country depends on the first term, which is negative due to concavity in the utility function ($d'(p) < 0$). However, if the imposition of the tariff does have an effect on the world price, then the effect on welfare varies with respect to the sign and magnitude of $\frac{\partial p_w}{\partial t}$.

Feenstra (2015) illustrates that for small changes in the tariff where a large country has the ability to impact world price, the welfare effect may be positive. This is because "a fall in the import price is an improvement in the terms of trade" (ch. 7,

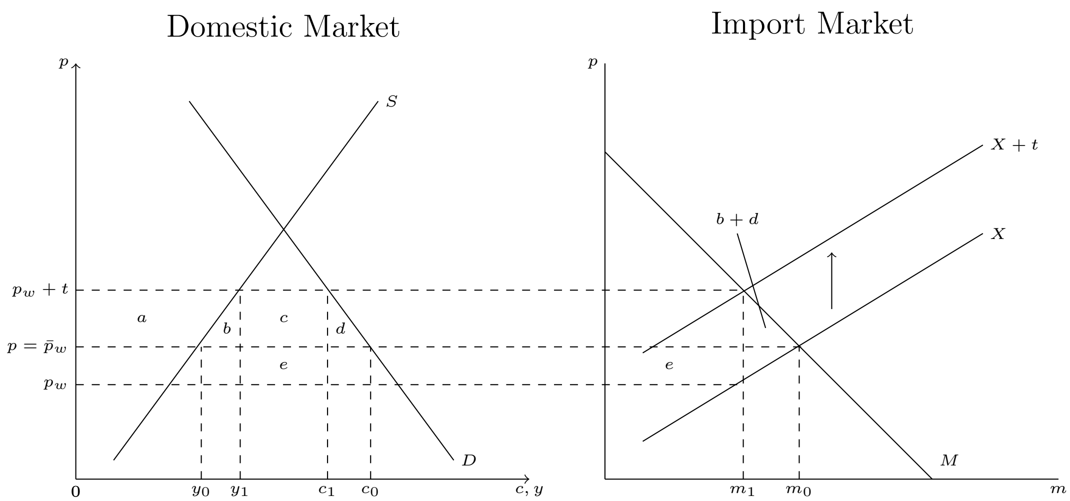

Figure 2.1: Domestic and Import Markets

2015). Figure 2.1 illustrates the domestic and import markets and the impact of a tariff on the export supply of the good. Here, we have an increasing export supply curve, and an original world price of $p = \bar{p}_w$, showing that without the tariff, the domestic price and world price are the same. After the imposition of the tariff, the world price with the tariff becomes $p_w + t$, which intersects the original export supply curve at a price of p_w. This change in domestic price is less than the tariff amount and therefore the effect is an improvement in the terms of trade. The resulting welfare is $e - (b + d)$, where e is a portion of the tariff revenue $(c + e)$ and $b + d$ is the deadweight loss. If the tariff is sufficiently large, any welfare gains are offset by the loss in consumer surplus $(a + b + c + d)$.

This illustration reveals that there may exist an optimal non-zero tariff under conditions of perfect competition if a tariff-imposing country has the ability to impact the world price of the good. The exact value of the optimal tariff depends on the elasticity of export supply. For more details on determining the optimal tariff, see chapter 7 in Feenstra (2015).

CHAPTER 3:
LITERATURE REVIEW

3.1 On Wine Flows

Zhang, Onel, and Seale evaluate the welfare effects of the additional 25% tariff placed on wine from the U.K., France, Spain, and Germany (Zhang *et al.*, 2021). This tariff is unique in that it only applies to select European countries. The authors estimate own- and cross-price elasticities of imported red and white wines. They use a source-differentiated import demand model to estimate the effects of the tariffs on a subset of European countries. Due to not having access to domestic sales price data, the authors used a multistage budgeting approach and maintain weak separability between domestic and imported goods. In this approach, they allocate expenditures between domestic and imported goods, and then between imported groups of red, white, and other wines, then between the different types of wines, and finally expenditures are broken out for each type of wine from source country. They find that consumer losses for the wines impacted by the tariffs were on par with the tariff, about 25%.

Greear and Muhammad investigate Japan's trade restrictions in the 21st century (2021); so far Japan has instituted bilateral trade agreements with major wine producing countries (Australia, Chile, France, Germany, Italy, Spain, and the U.S.) that have significantly dropped tariff rates. Prior to January 2020, U.S. imports were greeted by a 15% tariff. Wine tariffs on U.S. imports are currently at 8.5% and eventually will annually drop until they reach 0% over a 7-year period, in accordance with the United States-Japan Trade Agreement. Similarly, the bilateral agreements with European Union members, Chile, and Australia are all falling from 15% to 0%. The EU and Chilean wine tariffs became 0% in 2019 and Australian wine tariffs were eliminated in 2021. Using a source-differentiated dynamic framework, Greear and Muhammad (2021) estimate the effects of these tariff reductions on exports from each country and find that there isn't much substitution across wine-exporting countries.

Bianco, Boatto, Caracciolo, and Santeramo look at trade flows between 1997 and 2010 of 90% of bottled wine and claim that the recent trend for decreasing tariffs has been accompanied by increases in other trade frictions. Bianco et al. (2016) notes a variety of non-tariff barriers and mentions that the WTO classifies these into 6 broad categories: "food standards, labelling, conformity assessment, packaging, food containers and human health" (p. 11, 2016). They find that overall trade frictions for wine haven't changed much between countries in the 21st century as tariffs have been offset by other trade barriers.

Pinilla and Ayuda investigates the history of trade restrictions on Spanish wine in the early 20th century (2002). They look at difficulties Spanish producers had with exporting wine to Latin America. Latin American countries enacted tariffs to protect their nascent wine industries. Additionally, the French-Algerian open trade policy made France a less-willing importer of Spanish wine, charging tariff rates of over 80% on Spanish wine imports in the early 1930s. With the U.S.'s market completely closed, wine being unpopular in most of Europe, and France's excessively high tariff rates, the Spanish wine industry had a lot of trouble exporting.

Heien and Sims look at the impact of the 1989 Canada-United States Free Trade Agreement on wine exports to Canada from the United States (2000). They use export price index values from 1978 to 1994 of mostly Spanish and French wine in substitute of retail prices for estimating a demand function and then calculate the elasticities from the estimated model to estimate the impact of the counter-factual scenario that the trade agreement didn't take place. The original tariff was about 12% and fell by 25% per annum for the first couple of years and then 10% per annum for the last 5 years. Their research suggests that the primary source of the increase in wine flows from the U.S. to Canada was the reduction in non-tariff barriers, account for 90% of the increase in quantity over the 1989-1994 period.

3.2 On 2018 Trade War

Amiti *et al.* (2019) explore the impact of the 2018 trade policies in the U.S. and estimate that real income fell by $1.4 billion a month in the U.S. They estimate that the average tariff rates over 2018 rose from 1.5% to over 3%. "The trade war also caused dramatic adjustments in international supply chains, as approximately $164 billion dollars of trade [...] is lost or redirected in order to avoid the tariffs" (p. 22, 2019).

Flaaen *et al.* (2020) look at the effect of the 2018 tariffs on washing machines in the U.S. They find that the price of washers increased by 11.5% and the combined price of washers and dryers increased by 11.45% (despite dryers not being impacted by the tariffs). While U.S. prices increased, U.S. employment rose by 200 jobs in the industry and each job saved came with a price tag of over $800,000.

3.3 On Trade Restrictions

Kee *et al.* (2021) investigate elasticity estimates and constructs a trade restrictively index (TRI) to measure the impact of trade restrictions on trade. The authors show that the losses incurred from trade restrictions can be broken down "into three elements: import-weighted average tariffs, tariff variance, and the covariance between tariffs and import demand elasticities" (p. 681, 2021). The authors demonstrate that large tariff variances and high covariances makes the TRI and import-weighted tariffs diverge from one another. They find that in the U.S. and Canada, over 60% of the deadweight loss is due to high tariffs.

Irwin (2010) calculates a trade restrictiveness index, using a simplifying assumption of linear demand, to represent the U.S.'s trade policies from 1867 to 1961. This particular formulation yields a representative tariff that works as a substitute for the same amount of lost welfare as the country's trade policies. Irwin notes that imports were historically a relatively small part of the U.S. economy and so the costs felt by consumers weren't felt intensely. However, during times of greater imports (like during the Civil War), the cost of tariffs is much greater.

3.4 On Causality

Much econometric research attempts to show a causal relationship between the variables of interest. We typically define a relationship as causal if the model in question approximates a causal conditional expectation function (Angrist & Pischke, 2009). In this sense, we would say that the relationship between average price and the imposition of a tariff is causal if the absence of the tariff would not have resulted in an increase of average price, ceteris paribus. Uncovering what would have occurred in this counterfactual scenario is difficult to show with certainty. As a result, there are workarounds via statistical testing and clever thought processes that help us convince ourselves that the relationship observed is not simply due to chance. In Chapter 2, we discussed economic theory to provide a foundation for this belief. In Chapter 5, we will look at the statistical methods and modeling I use to test the structure of the model. One of the more popular methods used in this research is called Granger Causality.

3.4.1 Granger Causality

Granger-causality attempts to identify whether knowing x_{t-1} is useful in forecasting the value y_t. It does this by fitting a linear model to the endogenous variable with its lagged values and compares its fit to the fit of a model including the additional regressor and its lagged value. Afterward, it calculates an f-statistic. If the f-statistic is less than the critical value, we reject the null hypothesis that the input, x, does not granger-cause y.

CHAPTER 4:

DATA

The data used in this research is made publicly available by various governmental agencies. However, some of the data points had to be formally requested. The raw data is presented as a table in the Raw Data section of Appendix A and as files made available on GitHub (Johnson, 2022).

In this chapter, we will first go over the datapoints themselves and then look a bit deeper into data used in the study. For more information about the data or the transformations made on the data in order to get it structured for the study, see the data files in the data directory and the data-exploration-and-cleaning notebook in the notebooks directory in the accompanying GitHub repository (Johnson, 2022).

4.1 Data Descriptions

4.1.1 Imports and Exports

The trade data used in this study is collected and published by the U.S. International Trade Commission (USITC). The datapoints collected for this analysis are monthly, country-level aggregations of imports for consumption of wine and exports.

The datapoints collected for imports are as follows: Customs Value, First Unit of Quantity, Landed Duty-Paid Value, Dutiable Value, Imports Charges, Calculated Duty, and Charges, Insurance, and Freight (CIF). For exports, the datapoints collected were FAS Value and First Unit of Quantity of domestic exports.

Free Alongside Ship (FAS) Value is the value of exports at the U.S. port, including transport costs. First Unit of Quantity is the quantity used in assessing tariff charges (whereas Second Unit of Quantity pertains to quotas). The USITC measures quantities of wine in Liters. Imports Charges are aggregate transport costs (excluding duties). The Customs Value is the assessed value according to the U.S. Customs and Border Protection agency. Imports Charges is the sum of CIF and other charges (excluding duties). The Calculated Duty is the actual import duty as calculated by the Harmonized Trade Schedule (HTS).

4.1.2 Domestic Wine Production

The U.S Department of the Treasury's Alcohol and Tobacco Trade and Tax Bureau (TTB) publishes aggregations of domestic firms' filings of the Report of Wine Premises Operations Form (5120.17) within two months of the end of each period. Their monthly statistical release details various data points about wine production in the United States.

Their website provides the monthly publication back to 2008. The data went back to January 2000 by submitting a formal request to them. The datapoints used from their publication are bulk production of still wine and bottled production of still wine, cider, and effervescent wine. If available, the most present reports' prior year data was used since it contains the most up-to-date estimation of the various wine production data points. They publish their wine quantities in gallons instead of liters. Bulk wine is anything over 60 liters. For perspective, a barrel is 225 liters and a keg is 67 liters. A bottle is any container that is 4 liters or less.

4.1.3 Wine Prices

Average Wine Price and Producer Price Index data are collected by the Bureau of Labor Statistics (BLS) as monthly price data from 75 Urban Areas in the United States. The Average Wine Price series is a component in the calculation of the Consumer Price Index (Series ID 720311). Average Wine Price data is of sales prices of grape-based Red, White, and Rosé wines (excluding shipping costs for direct to consumer sales). And the Producer Price Index data is of costs for the Winery Industry. The Federal Reserve Bank of St. Louis (via its online database, FRED), makes these datasets available for the public in series APU0000720311 and PCU3121303121300 (U.S. Bureau of Labor Statistics, 2022), respectively.

Inflationary adjustments for production values use the producer price index for two reasons: 1) it may better-capture some of the supply-side issues related to the pandemic and 2) I couldn't find an industry-specific inflationary adjustment for the wine industry that made as much sense.

4.1.4 Disposable Income

Real Disposable Income by Month was collected from FRED from the U.S. Bureau of Economic Analysis (2022). This series is published in the Personal Incomes and Outlays release. The series uses the Consumer Price Index to adjust for inflation.

4.1.5 Population

Monthly population data is collected by the U.S. Bureau of Economic Analysis (BEA) in their Personal Income and Outlays release. This data is then presented by FRED in their series POPTHM. One shortcoming from this series is that it includes the Armed Forces overseas in the calculation.

If this was a large proportion of the population, this may result in per-capita rates of domestic wine quantities that are lower than they would otherwise be. But the percentage of military personnel overseas is small. For instance, there were 170 thousand U.S. military personnel overseas in 2020. The U.S. population was 300 million. To make the math easier, let us say there are 300 thousand overseas and 300 million people at home. That is 0.001% of the population overseas. This difference in population isn't going to have much of an impact in per-capita calculations.

4.2 Exploration

In this section, we are going to look into the data and see how the changes in the tariff rate on French, German, Spanish, and British wine may have affected trade. But first, let us look at the sort of magnitude the additional tariff had on the calculated duty for wine imports. In figure 4.1, we see that after the 25% tariff was levied, the effective tariff rate on the U.K., France, Spain and Germany went from about 1% to having a maximum rate of over 14%.

In figure 4.2, we can see that although the amount of world imports to the U.S. has had a steady upward trend in the 21^{st} century, the four countries impacted by the additional 25% tariff still represented over a fifth of total wine imports to the U.S. during the most popular times of the year for wine. After the tariffs, we see a drop in wine imports that isn't accompanied by the typical, cyclical increase.

Since 2010, the amount of wine produced for domestic consumption (calculated by subtracting U.S. exports from domestic production) has had a slight upward trend (figure 4.3). However, there is not visibly much deviation from the trend at the same time as the tariffs. This may be due to lags in wine production; most bottled wine is produced from grapes grown the prior year. Toward the tail end of the time series we

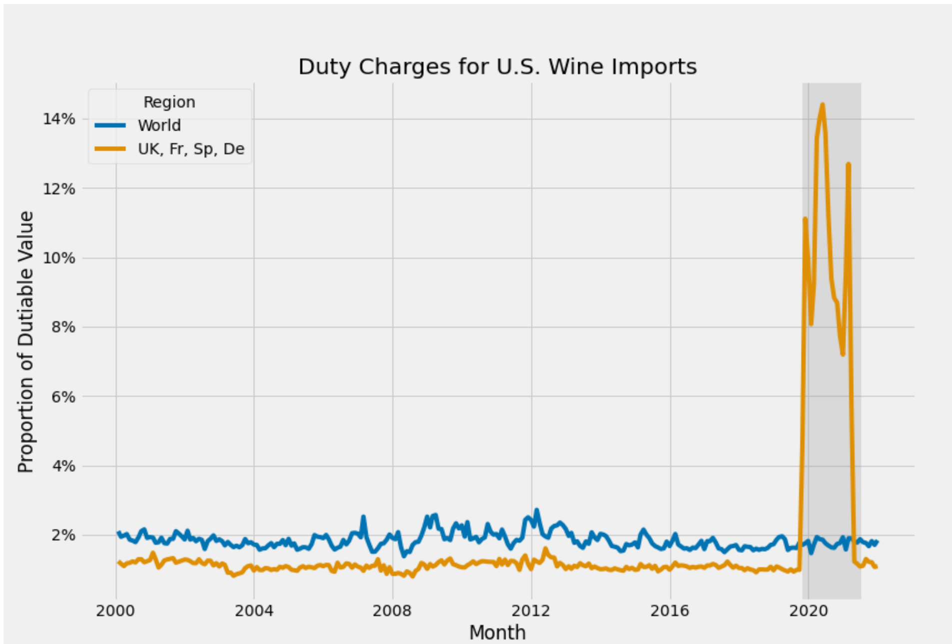

Figure 4.1: The proportion of wine duties per dutiable value in the U.K., France, Spain, and Germany against the rest of the world. Highlighted is the period when the additional 25% tariff was levied.

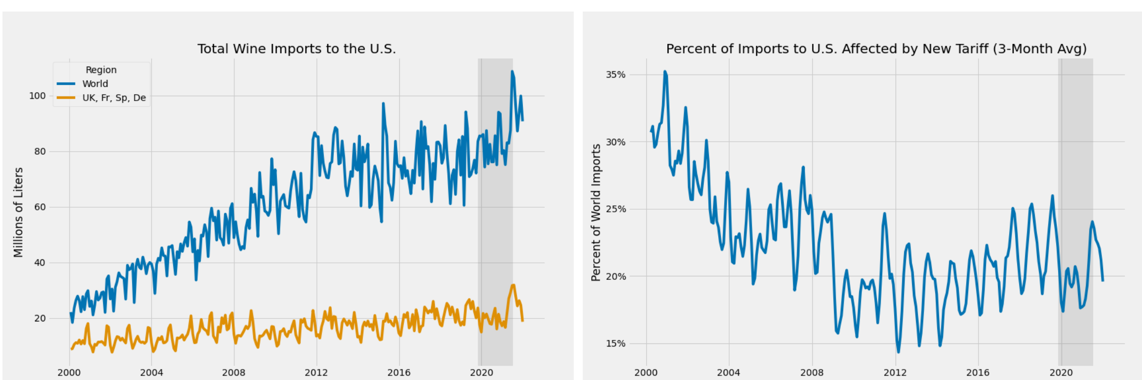

Figure 4.2: (Left) The total amount of wine imported to the U.S. for the U.K., France, Spain, and Germany against the rest of the world. (Right) U.S. imports from the U.K., France, Spain, and Germany as a percentage of world imports. The highlighted region represents the period with an additional 25% tariff levied.

see a lower spike in the production of wine during late 2021 when economic theory would have lead us to believe we would see an uptick in production as the protectionist policies increase the reward domestic producers receive from domestic sales.

This dip in production might be related to the heavy fire season in California during 2020. Wildfires not only have the ability to burn crops, but the smoke gets captured in the skins of the grapes and changes the tannins (particularly for red wines). The ash that rains down and coats the earth changes the composition of the soil which affects the following years' growing seasons. Each of these aspects may lead to a harvest being useless and the grapes getting trashed. In this situation, we would likely still see an attempt at producing wine from the harvest and the amount of wine produced would stay the same but the extent to which that wine is bottled would fall. All that being said, the variation in non-basic production (wine production for domestic consumption instead of exports, see figure 4.3) still follows the standard production cycle and so far does not look like it deviates far from the trend.

Between October 2019 and July 2021, the period that nearly a fifth of U.S. wine imports saw an over 10-fold increase in tariff charges, we see that the percent of imports from the countries the tariff was imposed on fell and imports from the rest of the world increased. In figure 4.4, we see that the average price of wine during this time ticked up about 5%.

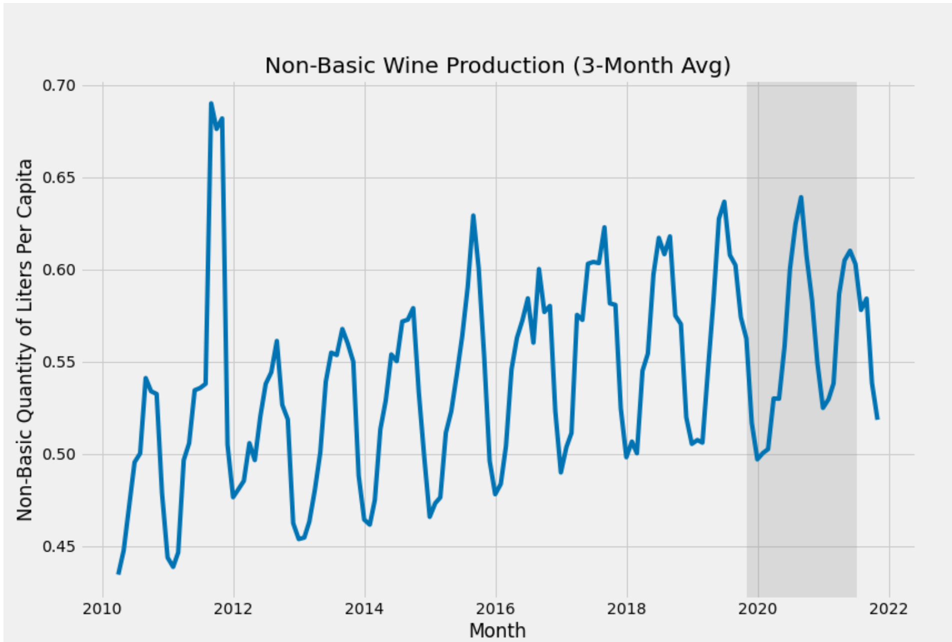

Figure 4.3: Monthly production of liters of wine per-capita for domestic consumption.

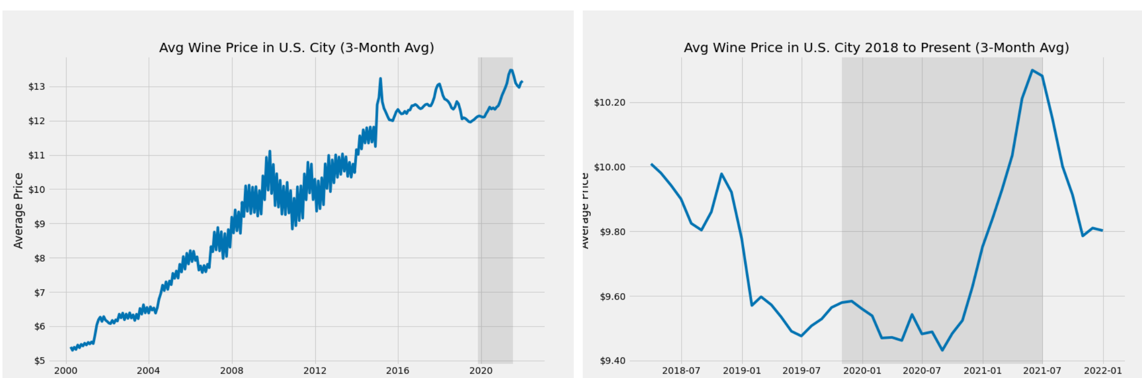

Figure 4.4: The average price (with a 3-month rolling average) of wine taken from a monthly survey of 75 urban areas across the U.S. This price is adjusted for inflation using the producer price index for the wine industry, which has an even lower price of wine than if the consumer price index was used.

CHAPTER 5:
METHODS

In this chapter, we will look at the vector auto regression model used to analyze the economic data. Before getting started, we will first go over the implementation details of a vector autoregression model. And then we will quickly discuss how it is used. Afterward, we'll go into the methods I used in this research.

I started off with using a Granger causality test to identify which variables have a Granger-causal relationship. Afterward, I transformed the data to make it stationary by removing the seasonality and the trend in the data. Then I fitted a VAR model on the original dataset and identified a lag order to use. After this, I fit a model with the selected lag order and check for serial correlation in the residuals. In the Results section, we will use this model to predict the most recent periods' average price data and look into some of the impulse response functions. Due to an audit of the wine production data from the Alcohol and Tobacco Trade and Tax Bureau (TTB) that started in February 2022 and has yet to conclude (as of October 2022), recent data is for out-of-sample predictions.

5.1 Vector Auto Regression

In this analysis, I utilize the statsmodels library in python for a standard, stable vector autoregressive (VAR) model. This library implements the methods detailed in Lütkepohl (2005). The model is an extension of univariate time series' ARIMA models. Its assumptions are the following: zero conditional mean in the error term, the variables are stationary (their first and second moments are time-invariant), the variables are stable (the eigenvectors of the coefficient vectors are between 0 and 1), large outliers are unlikely, and there is no perfect multicollinearity. These assumptions allow us to utilize standard statistical tests in order to identify structure in the data and increase predictive power.

The VAR(p) model is as follows:

$$y_t = v + \sum_{i=1}^{p} A_i y_{t-p} + u_t, t = 0, \pm 1, \pm 2, ...$$
$$u_t \sim \mathcal{N}(\mu, \sigma^2)$$

where y_t is a $(Kx1)$ vector, v is a $(Kx1)$ vector, A_i are coefficient matrices (KxK), v is a $(Kx1)$ vector of constants, p is the lag order, and t is the period.

This model estimates the coefficients using least squares. The following exposition can also be found in Lütkepohl (2005, pp. 70-72).

First, we need to define a few variables in order to illustrate the model.

$$
\begin{aligned}
Y &:= (y_1, ..., y_n) && (K\text{x}n) \\
B &:= (v, A_1, ..., A_p) && (K\text{x}(K_p + 1)) \\
Z_n &:= \begin{bmatrix} 1 \\ y_n \\ \vdots \\ y_{n-p+1} \end{bmatrix} && ((K_p + 1)\text{x}1) \\
Z &:= (Z_0, ..., Z_{n-1}) && ((K_p + 1)\text{x}n) \\
U &:= (u_1, ..., u_n) && (K\text{x}n) \\
\vec{y} &:= \text{vec}(Y) && (Kn\text{x}1) \\
\vec{\beta} &:= \text{vec}(B) && ((K^2p + K)\text{x}1) \\
\vec{b} &:= \text{vec}(B^T) && ((K^2p + K)\text{x}1) \\
\vec{u} &:= \text{vec}(U) && (KT\text{x}1)
\end{aligned} \tag{5.1}
$$

The vec() operator takes a matrix and stacks it column by column in order by index ascending. The var(p) model can then be written more concisely as

$$Y = BZ + U \tag{5.2}$$

or

$$y = (Z^T \otimes I_K)\beta + u \tag{5.3}$$

and the covariance matrix of $\vec{u}$ is $\sum_{\vec{u}} = I_n \otimes \sum_u$. We are using $\otimes$ as the Kronecker product (Graham, 2018).

In standard least squares fashion, we're solving the following optimization problem:

$$\begin{aligned}
S(\vec{\beta}) &= \vec{u}^T(I_n \otimes \textstyle\sum_u^{-1})\vec{u} = \vec{u}^T(I_n \otimes \textstyle\sum_u^{-1})\vec{u} \\
&= [\vec{y} - (Z^T \otimes I_K)\vec{\beta}]^T(I_n \otimes \ )_u^{-1}[\vec{y} - (Z^T \otimes I_K)\vec{\beta}] \\
&= \mathrm{vec}(Y - BZ)^T(I_n \otimes \textstyle\sum_u^{-1})\mathrm{vec}(Y - BZ) \\
&= \mathrm{trace}\left((Y - BZ)^T \textstyle\sum_u^{-1}(Y - BZ)\right)
\end{aligned}$$

We also know:

$$\begin{aligned}
S(\vec{\beta}) = \vec{y}^T(I_n \otimes \textstyle\sum_u^{-1})\vec{y} + \vec{\beta}^T(Z \otimes I_K)(I_n \otimes \textstyle\sum_u^{-1})(Z^T \otimes I_K)\vec{\beta} \\
- 2\vec{\beta}^T(Z \otimes I_K)(I_n \otimes \textstyle\sum_u^{-1})\vec{y} \\
= \vec{y}^T(I_n \otimes \textstyle\sum_u^{-1})\vec{y} + \vec{\beta}^T(ZZ^T \otimes \textstyle\sum_u^{-1})\vec{\beta} - 2\vec{\beta}^T(Z \otimes \textstyle\sum_u^{-1})\vec{y}
\end{aligned}$$

So the minimization problem becomes:

$$\frac{\partial S(\vec{\beta})}{\partial \vec{\beta}} = 2(ZZ^T \otimes \textstyle\sum_u^{-1})\vec{\beta} - 2(Z \otimes \ )_u^{-1}\vec{y} = 0$$

Solving this problem yields the following:

$$\begin{gathered}
(ZZ^T \otimes \textstyle\sum_u^{-1})\hat{\vec{\beta}} = (Z \otimes \textstyle\sum_u^{-1})\vec{y} \\
\hat{\vec{\beta}} = ((ZZ^T)^{-1} \otimes \textstyle\sum_u)(Z \otimes \textstyle\sum_u^{-1})\vec{y} = ((ZZ^T)^{-1}Z \otimes I_K)\vec{y}
\end{gathered} \tag{5.4}$$

So $\hat{\vec{\beta}}$ is the matrix of estimated coefficients.

5.1.1 VAR Applications

This VAR(p) model is used primarily in forecasting and structural analyses (Lütkepohl, 2005). Its application is commonplace in macroeconomic policy-making, primarily in monetary economics. However, its use-case is to analyze multivariate timeseries wherein there's co-movements in the various series. For instance, Cudia (2012) used a VAR to investigate the effect of the 2008 global financial crisis on Philippine's export market. Though VAR models are used in structural analysis, they are not as useful for assessing causation as they are for data description and forecasting (Stock & Watson, 2001). However, it is still common for researchers to use impulse response functions to add additional pieces of evidence to claims about relationships between inputs.

5.2 Tests

5.2.1 Granger Causality Tests

At the onset, I used Granger causality tests to help identify which variables may potentially have an effect on the real price of wine. The following is the null hypothesis for the Granger Causality Test:

$$H_0 : \text{the series } x \text{ does not granger-cause } y$$
$$H_1 : \text{the series } x \text{ does granger-cause } y$$

The specification for the Granger causality test is:

$$y_t = \sum_{i=1}^{\infty} \alpha_i y_{t-i} + c_{1,t} + u_{1,t}$$
$$y_t = \sum_{i=1}^{\infty} \alpha_i y_{t-i} + \sum_{i=1}^{\infty} \beta_i x_{t-i} + c_{2,t} + u_{2,t}$$

the first equation is considered unrestricted (UR) and the second restricted (R).

The f-statistic is calculated by the following equation:

$$f = \frac{\frac{ESS_{UR} - ESS_R}{q}}{\frac{ESS_{UR}}{n-k}}$$

where $ESS = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$

I chose to use a maximum lag of 16 for the test. I ran these tests on all of the variables of interest and identified a set of variables that made sense theoretically, practically, and were statistically significant. After identifying variables that may be of help, I subsetted my data for those variables to use as inputs. Then I ran the Granger causality test on the subset of variables and have added the results from this test in the table B.1 in appendix B.

5.2.2 Stationarity Tests

After identifying the columns of interest, I decomposed each of the datapoints to its trend, seasonality, and residual. Then I de-trended the data to remove its unit root and any seasonal changes. Much of the data had both an upward trend and a 12-month cycle in it due to the seasonality in consuming wine and harvesting grapes. This non-stationarity needed to be corrected so that the overall behavior of the time series to be time-invariant.

In order to help determine whether the data was stationary, I ran both Augmented Dickey-Fuller and Phillips-Perron tests. Both of these tests build on the Dickey-Fuller test for stationarity and have the following hypotheses:

$$H_0 : \text{the series has a unit root}$$

$$H_1 : \text{the series does note have a unit root}$$

The Augmented Dickey-Fuller test is specified as follows:

$$\delta y_t = c + \beta t + \alpha y_{t-1} + \phi_1 \delta Y_{t-1} + \delta Y_{t-2} + ... + \phi_p \delta Y_{t-p} + \epsilon_t$$

c is a constant, t is the time trend, p is the number of lags, and δ is the difference between values in the observations. If $\alpha = 0$, the series has a unit root and if it's less than 0, the series doesn't have a unit root. The Dickey-Fuller test statistic is calculated as $DF_\alpha = \frac{\hat{\alpha}}{\sigma_{\hat{\alpha}}}$, where σ is standard error.

The results from both the ADF and Phillips-Perron tests were consistently similar. Since the exposition of the Phillips-Perron test is really extensive, folks who are interested in the details of that test can check out Phillips (1987).

5.2.3 Johansen Cointegration Tests

The stationary datapoints are tested for cointegration using a Johansen Cointegration test. This test identifies the maximum likelihood estimators of the model by finding its series' eigenvalues. It determines the rank of the cointegrating series (r) using two likelihood ratio tests: the trace test and maximum eigenvalue test. The hypothesis for the test is:

$$H_0 : \text{rank}(A) \leq r$$

$$H_1 : \text{rank}(A) > r$$

This hypothesis is tested using the following likelihood ratio form:

$$-2ln(A) = -n \sum_{i=r+1}^{p} ln(1 - \lambda_i)$$

where r is the number of cointegrating vectors, n is the number of observations, p is the number of lags, and A is the model.

Both of the resultant critical values are presented in table B.2 in appendix B section 2.

5.2.4 Durbin Watson Tests

The Durbin Watson test identifies autocorrelation in the error term. Failing to reject the null hypothesis means that no autocorrelation was detected. The hypothesis is:

$$H_0 : \text{There is no first order autocorrelation.}$$

$$H_1 : \text{There is first order autocorrelation.}$$

The test specification is as follows. The nearer the Durbin-Watson statistic is to 2, the better.

$$\text{DW} = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2}$$

5.3 Modeling

5.3.1 Specification

The VAR(p) model used is a form of regression model with lagged variables that solves a system of equations for the endogenous variables. The specification I used is as follows:

$$
\begin{aligned}
\text{price} &\sim \text{bottled} + \text{exp\_q} + \text{income} + \text{imp\_q}_i + \text{duty\_r}_i + \text{cif\_r}_i + \text{prop\_imp}_i \\
\text{bottled} &\sim \text{price} + \text{exp\_q} + \text{income} + \text{imp\_q}_i + \text{duty\_r}_i + \text{cif\_r}_i + \text{prop\_imp}_i \\
\text{exp\_q} &\sim \text{price} + \text{bottled} + \text{income} + \text{imp\_q}_i + \text{duty\_r}_i + \text{cif\_r}_i + \text{prop\_imp}_i
\end{aligned}
$$

$$
\left.\begin{aligned}
\text{imp\_q}_j &\sim \text{price} + \text{bottled} + \text{exp\_q} + \text{income} + \text{imp\_q}_i + \text{duty\_r}_i + \text{cif\_r}_i + \text{prop\_imp}_i \\
\text{prop\_imp}_j &\sim \text{price} + \text{bottled} + \text{exp\_q} + \text{income} + \text{imp\_q}_i + \text{duty\_r}_i + \text{cif\_r}_i + \text{prop\_imp}_i
\end{aligned}\right\} i \neq j
$$

$$
\left.\begin{aligned}
\text{duty\_r}_{\text{bloc}} \sim \text{price} + \text{bottled} + \text{exp\_q} + \text{income} + \text{imp\_q}_i + \text{duty\_r}_i + \text{cif\_r}_i \\
+\text{prop\_imp}_i + \text{imp\_q}_{\text{bloc}} + \text{prop\_imp}_{\text{bloc}} + \text{cif\_r}_{\text{bloc}}
\end{aligned}\right\} i \neq \text{bloc}
$$

where all of the prices are adjusted for inflation using the wine industry-specific portion of the producer price index except *income*, which is adjusted using the consumer price index, $*\_r$ indicates price per liter, $*\_q$ indicates quantity of liters, $*\_w$ indicates a value for the world as a whole, *imp* represents imports, *exp* represents exports, and the duties and proportion of imports, *prop_imp*, variables are proportions for the whole world for each country or grouping of countries. The cif variable is for charges, insurance, and freight. The *bloc* subscript indicates the value is for the aggregate of the U.K., France, Germany, and Spain. The subscripts i and j indicate the country or groupings of countries. These groupings are either the values for Chile, Italy, Australia, the bloc of countries, or the rest of the world. All values are adjusted for stationarity and seasonality.

Though the duty rate for the bloc of countries impacted by the tariff is theoretically not an endogenous variable, the particular implementation of the vector autoregression model in python doesn't allow for estimating impulse response functions (IRF) on exogenous variables. I implemented the same model in R, as well, using a VARX package that allows for exogenous variables' IRF and the results from the model were nearly identical to this specification.

5.3.2 Autocorrelation

After estimating the model, I validated that the model is stable (that the eigenvalues of the model are between 0 and 1) and then used a Durbin Watson test to check for serial correlation in the residuals. The Durbin Watson statistic ranges from 0 to 4 with high values indicating a positive correlation and low values indicating a negative correlation. The nearer the values are to 2, the less autocorrelation is detected. The test statistic specification can be found in 5.2.4 and the results of the test can be found in Table B.3 in appendix B.

5.3.3 Stability

A VAR(p) model is considered stable if all of the eigenvalues of A are between 0 and 1. This means that y_t is stable if the following conditions are met:

$$\det(I_{K_p} - Az) \neq 0, for|z| \leq 1$$

$$\vec{u} := E(Y_n) = (I_{K_p} - A)^{-1}\vec{v}$$

$$\gamma_Y(h) = \sum_{i=0}^{\infty} A^{h+1} \sum\nolimits_U (A^i)^T$$

where $\sum_U := E(U_t U_t^T)$. So y_t is stable if $\det(I_K - A_1 z - ... - A_p z^p) \neq 0$, for $|z| \leq 1$.

Additionally, if a VAR(p) process is stable, we know that it is also stationary (p. 25, Lütkepohl, 2005). The stationarity condition requires that the first and second moments are time-invariant.

5.3.4 Impulse Response Analysis

Lastly, I conduct an impulse response analysis to see how a unit impulse in the endogenous variables may impact real prices. Impulse response analyses are a way of looking at the pairwise shock between variables in a model. These are computed using a moving average of all periods' representations of the model. Specifically,

$$\phi_{jk,i} = A_{jk,i} + \sum_{n=1}^{i-1} A_{jk,n}\gamma_{i-n}, \text{ for } i \in \{1, 2, ..., pK - p\}$$

where p is the lag order, k is the dimensions, and j is the shocked variable $j \neq k$.

CHAPTER 6:

RESULTS

Though the results of the impulse response functions weren't statistically significant, the root mean square error in for the predicted prices was about 0.4, which is really good. You can see the last three years of predicted values plotted against their actual values in figure 6.1. Economic theory suggests that an increase in the tariff rate would lead to an increase in the price of the good the tariff was imposed on. This model wouldn't be appropriate to use as evidence for a causal relationship between the tariff rate changes and average domestic wine price. That's because the impulse response function for the impulse of the duty rate on imports from the U.K., France, Germany, and Spain, a group of countries consisting of nearly a fifth of wine imports to the U.S., came back inconclusive. That being said, the expected directional changes in the impulses were decent.

In figure 6.2, I show the impulse response function for the duty rate per liter of wine imported to the U.S. on average domestic wine prices. The y-axis is in dollars and the x-axis is in periods. The dotted corridor is a 95% confidence interval. This impulse response function indicates that a one-unit impulse on the rate per liter of wine imports from the U.S., France, Germany, and Spain has a resulting $0.14 drop in the average wine price in urban U.S. cities in the next period. The second period, that drop is an additional differential is only $0.05. However, the confidence interval

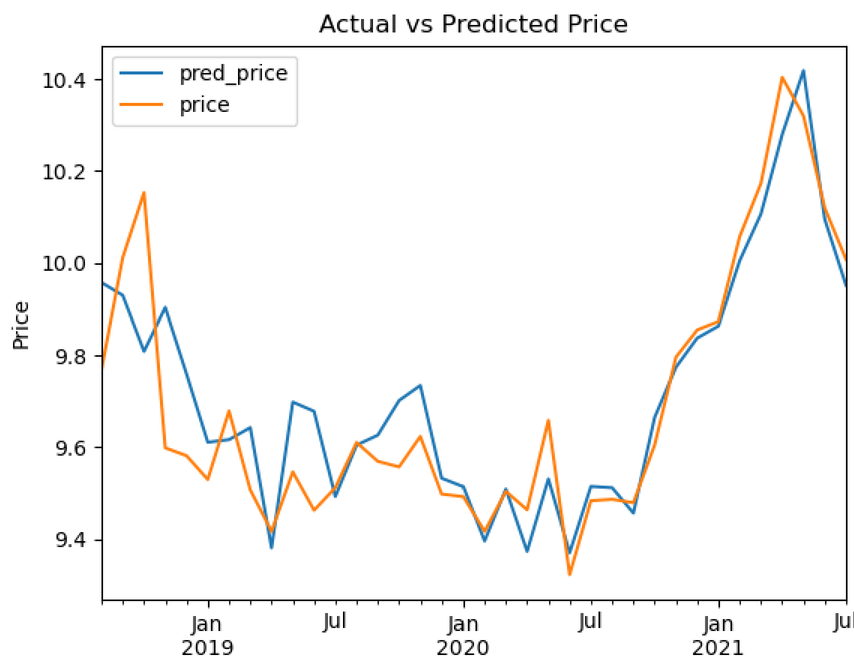

Figure 6.1: The predicted average wine price plotted against the actual average wine price.

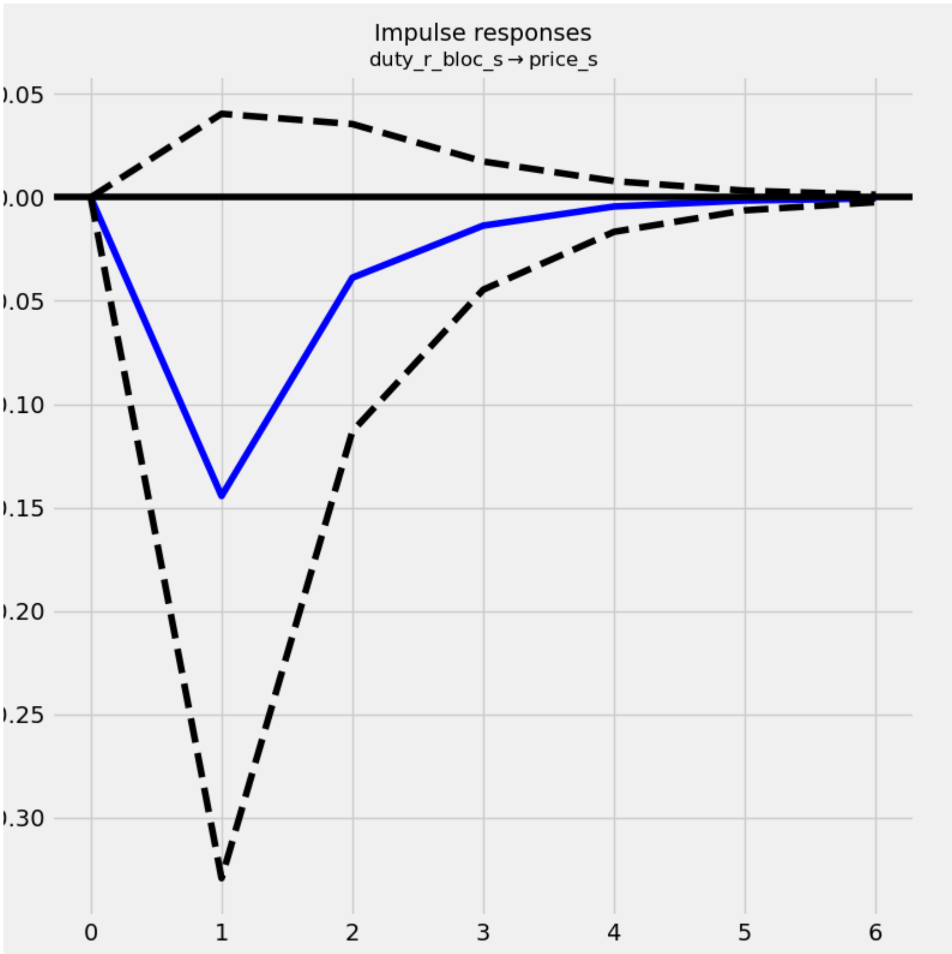

Figure 6.2: Response of average domestic wine prices from a one-unit impulse on the duty rate imposed on the U.K., France, Spain, and Germany

contains zero; this result isn't statistically significant. When we loosen this constraint to be at the 90% confidence level, we see statistical significance for the impulse of duty rate on domestic price as shown in figure 6.3.

The impulse response of the quantity of wine imports from the bloc of affected countries from the tariff makes a lot more economic sense. The impulse response suggests that a one-unit impulse in the duty rate on imports from the bloc has a 200,000 liter decrease of wine imported to the U.S. in the following period and around

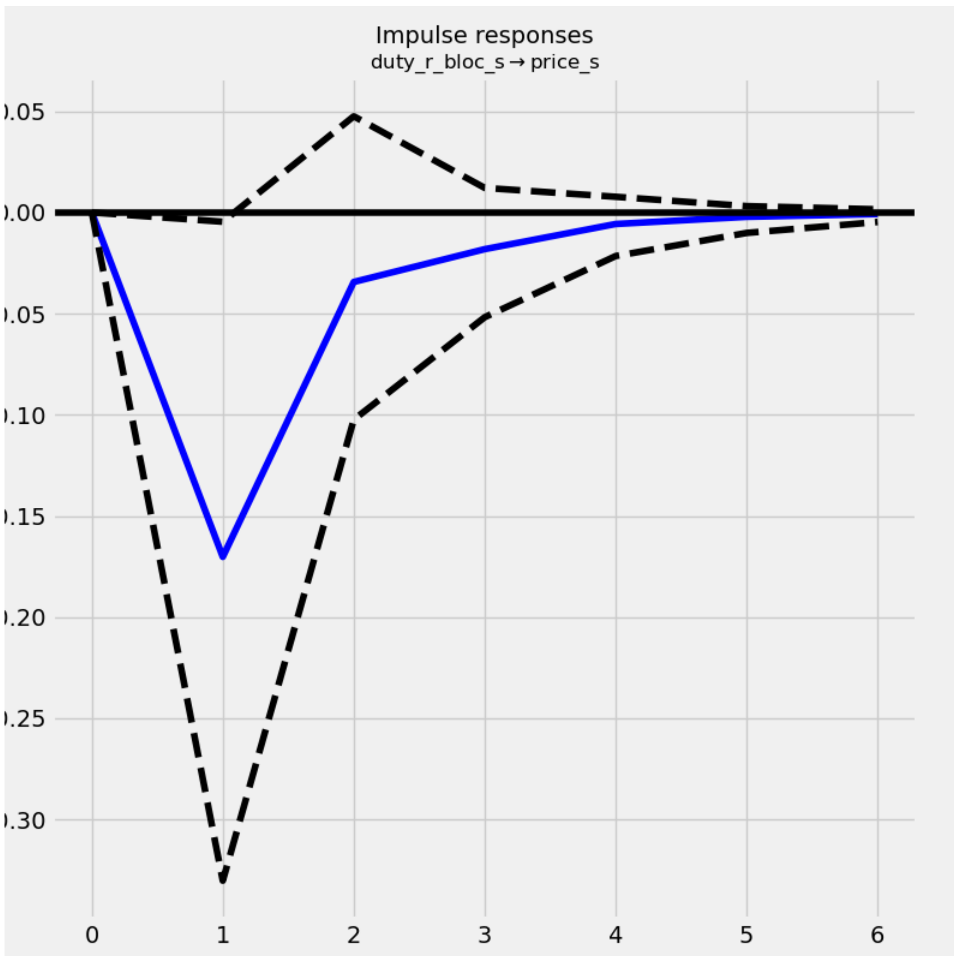

Figure 6.3: Response of average domestic wine prices from a one-unit impulse on the duty rate imposed on the U.K., France, Spain, and Germany with 90% confidence intervals

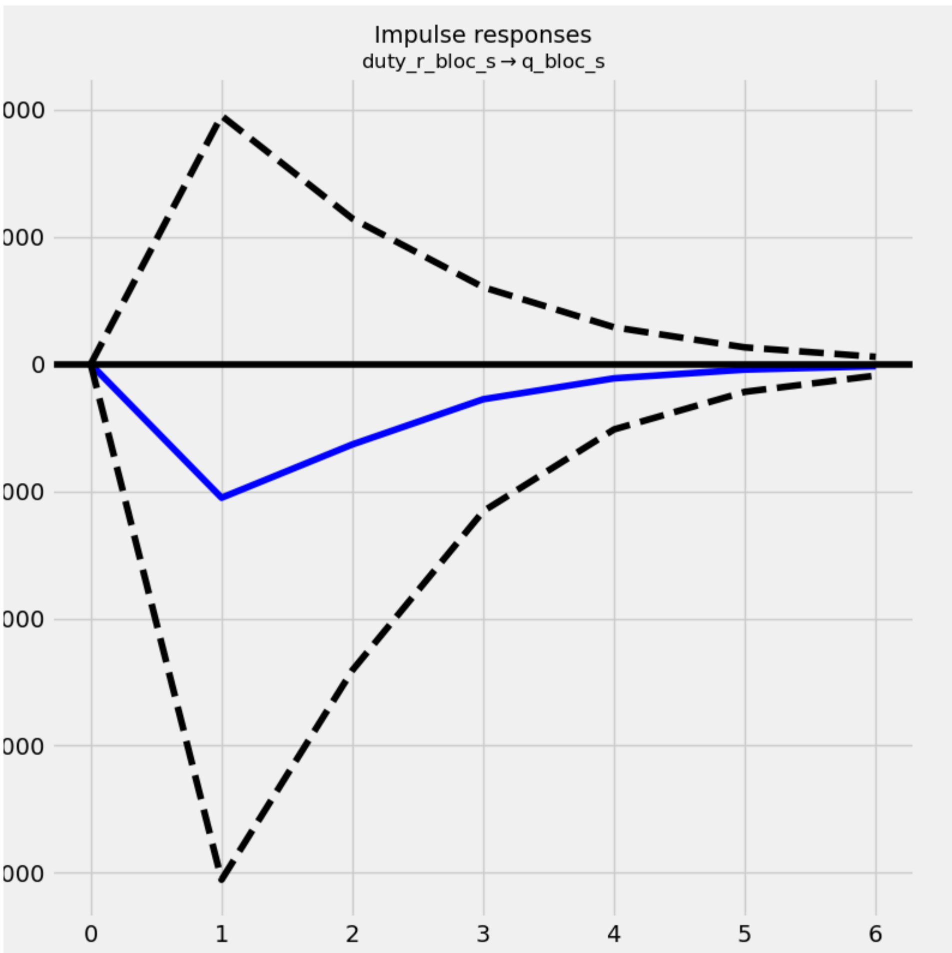

Figure 6.4: Response of average domestic wine prices from a one-unit impulse on the duty rate imposed on the U.K., France, Spain, and Germany

160,000 liter additional decrease in the second period, as shown in figure 6.4. The confidence interval shows that this finding is also not statistically significant. The impulse response of quantity of liters from the bloc on price was positive but near zero ($1e^{-8}) and not statistically significant (figure 6.5).

This may be due to wine from the U.K., France, Spain, and Germany being more expensive than the substituted wine American consumers are opting to consume. Figure 6.6 shows that there may be an inverse relationship between quantity imported from the U.K., France, Germany, and Spain and the top three importers to the U.S.—Italy, Australia, and Chile. This impulse response function shows slight statistical significance.

I suspect that the primary reason for the lack of statistical significance in many of these impulse response functions is that the model uses about 20 regressors with 2 lags for a total of 60 variables while there are only 260 observations in the dataset. However, VAR models are typically used for forecasting and the model appears to perform relatively-well for forecasting the average wine price.

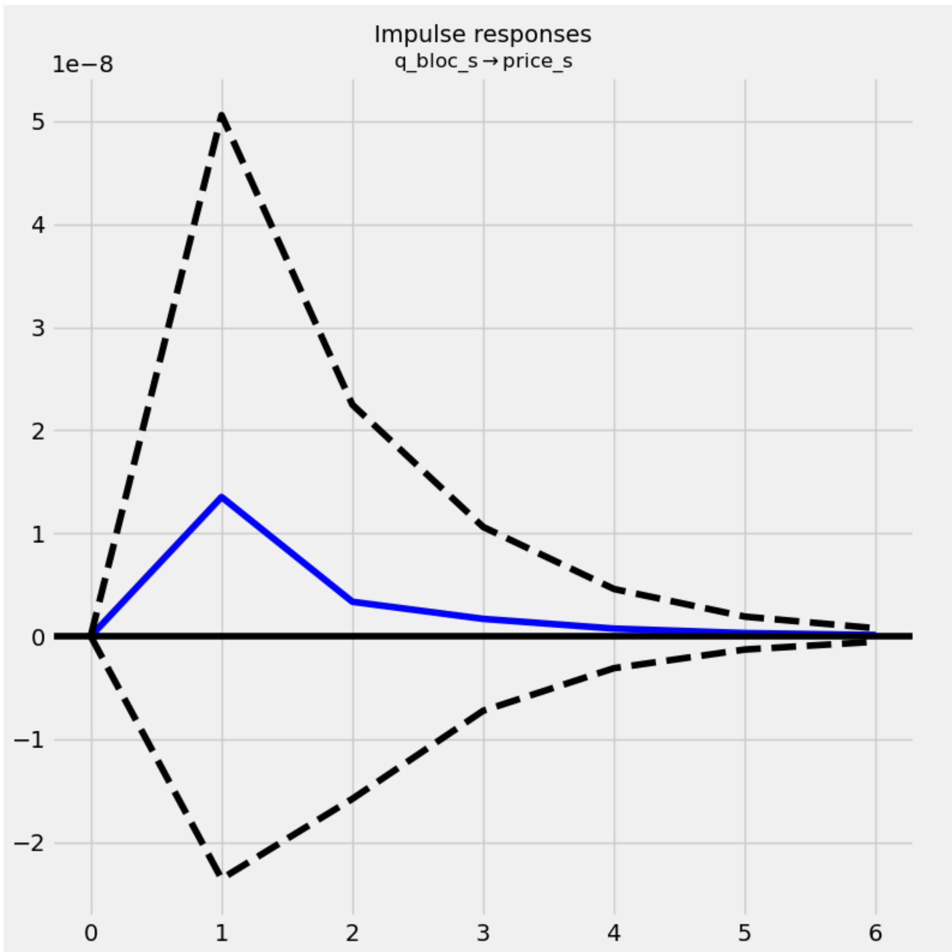

Figure 6.5: Response of average domestic wine prices from a one-unit impulse on the quantity imported of wine from U.K., France, Spain, and Germany

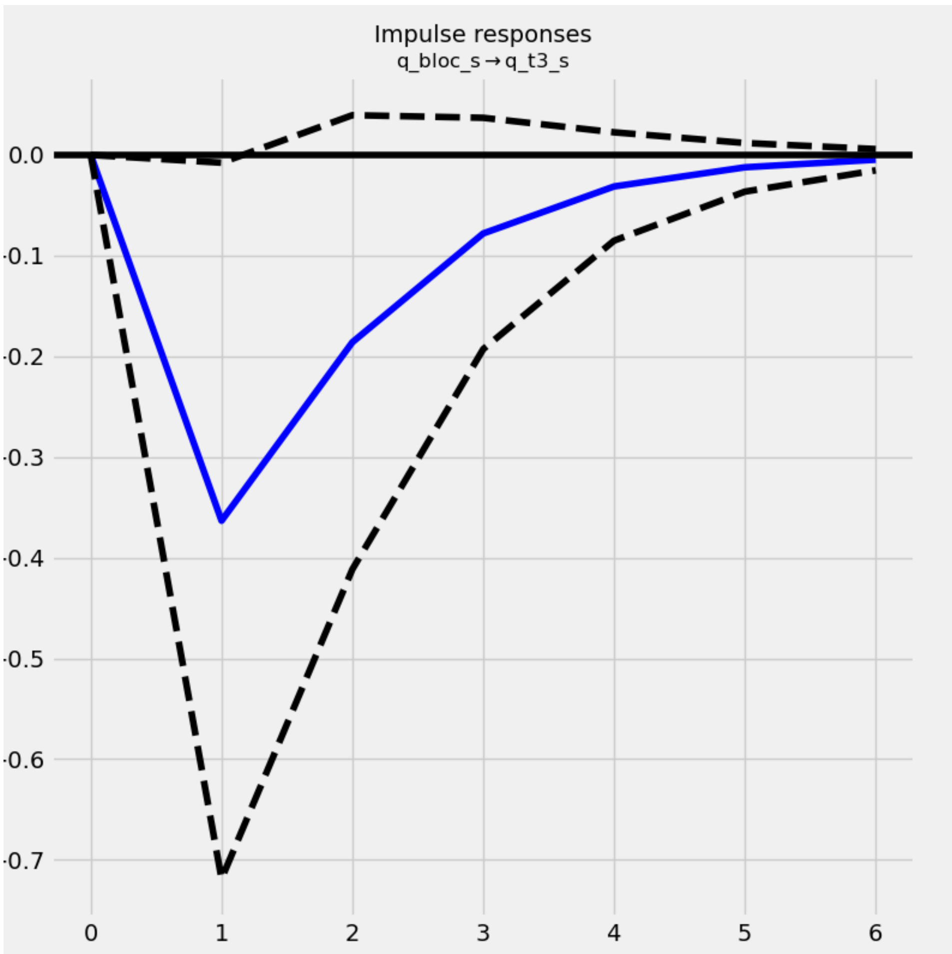

Figure 6.6: Response of quantity imported from the top three exporters to the U.S. (Chile, Australia, and Italy) from a one-unit impulse on the quantity imported of wine from U.K., France, Spain, and Germany

CHAPTER 7:
DISCUSSION

I started off this research project thinking that such a substantial increase in the tariff rate on wine from France, Italy, Spain, and the U.K. into the U.S. would most certainly have a negative impact on consumers. I saw this and thought that it may act as a natural experiment. And the effect the pandemic had on U.S. wine consumption would be held somewhat constant since all countries' exports would likely be impacted similarly.

My first thoughts were about how I can potentially measure the change. With the limitations on readily available domestic sales data in general and more specifically detailed data for sales of origin, I started looking into what other sorts of models may be useful in exploring the question of how U.S. wine prices were impacted by the additional tariffs and found that a vector autoregression model may be a good tool to investigate the relationship. Since other researchers used impact response functions from VAR models to gauge how both exogenous and endogenous variables change due to shocks in other inputs, I suspected a VAR could prove useful in illustrating what economic theory suggests.

With the long history of research in economics showing that the positive effect of tariffs on domestic prices, it was surprising that this model's impulse response functions were not able to weed out similar results with statistical significance at the

95% level. However, for the response of price from a one standard-deviation impulse in the duty rate was statistically significant at the 90% confidence level.

Most of the impulse response results were not statistically significant. The primary effect that was statistically significant was somewhat surprising—that a change in the quantity imported from the bloc of countries results in the opposite directional change in the quantity imported from the top three exporters not included in the bloc impacted by the country. This was surprising to me because research by Greear & Muhammad (2021) suggests that wine from various regions across the world are not substitutes for one another.

With data differentiating between white and red wine imports and production, I suspect we could look into which countries' wines are substitutes for one-another (ie, are countries which specialize in producing dry red wines only really substitutes for other countries who primarily produce dry reds?). However, this still would not really help the issues of statistical significance, which are likely due to the relationship between the number of observations included in the model (260) and the variables used (20*3—two lags) to model the VAR.

One route that could potentially mitigate the unexpected results from the model is using a structural VAR model instead. In a structural VAR, we define the relationships within the model and constrain the model with a set of conditions for how the variables ought to behave. I opted to use an unrestricted VAR because I believed the expected behaviors would easily be observed from the data since economic theory provides a strong case suggesting the imposition of tariffs reduces imports and increases domestic prices. Had I instead used a structural VAR for the analysis, we may see results with stronger statistical significance by virtue of the conditions set in the model.

Despite difficulties with interpreting causality in the impulse response functions, economic theory has provided us with justifications for the directional changes shown in the functions. And overall, this model may work well for forecasting, which is the most common use-case.

Though this specific application is not helpful in assessing causality in the model, it does provide us with another piece of evidence in support of the claim that an increase in tariff decreases the quantity imported from the impacted country. And if there were more observations available, I suspect we would be able to see statistically significant results in the impulse response functions. That being said, the model still would not be useful in estimating welfare changes in U.S. consumers.

CHAPTER 8:

CONCLUSION

This research explored the impact of the 2018 tariffs on European wine on domestic wine prices. The only statistically significant conclusion with 95% confidence that can be drawn from this model is a one-unit increase in the per-liter tariff rate on French, Germany, Spanish, and British wine results in a drop of over 150 million liters of wine imports from that group of countries in the two months that followed. The impact of the increase in the tariff rate on domestic wine prices was ambiguous until the confidence interval was decreased to 90%.

Though this model was able to provide a decent in-sample prediction for wine prices for the last three months of 2021, it wasn't able to provide insight into a causal relationship between the additional 25% tariff on the bloc of countries and domestic wine prices. The closest claim to causality in this research was that the additional tariff had a Granger-causal effect on the average wine price in the U.S. However, that doesn't show beyond a doubt that the additional tariff caused the change in price, rather the effect on price appears unlikely to be due simply to chance.

REFERENCES

Amiti, Mary, Redding, Stephen, & Weinstein, David. 2019 (Mar.). *The Impact of the 2018 Trade War on U.S. Prices and Welfare*. Tech. rept. w25672. National Bureau of Economic Research, Cambridge, MA.

Angrist, Joshua, & Pischke, Jörn-Steffen. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Cudia, Cynthia P. 2012. The Effect of Global Financial Crisis on the Philippines' Export Sector: A Vector Auto-Regression Analysis. **11**(2), 18.

Dal Bianco, Andrea, Boatto, Vasco Ladislao, Caracciolo, Francesco, & Santeramo, Fabio Gaetano. 2016. Tariffs and non-tariff frictions in the world wine trade. *European Review of Agricultural Economics*, **43**(1), 31–57.

Feenstra, R.C. 2015. *Advanced International Trade: Theory and Evidence - Second Edition*. Princeton University Press.

Flaaen, Aaron, Hortaçsu, Ali, & Tintelnot, Felix. 2020. The Production Relocation and Price Effects of US Trade Policy: The Case of Washing Machines. *American Economic Review*, **110**(7), 2103–2127.

Graham, A. 2018. *Kronecker Products and Matrix Calculus with Applications*. Dover Books on Mathematics. Dover Publications.

Greear, Emily K., & Muhammad, Andrew. 2021. Tariff Elimination and the Competitiveness of Wine-Exporting Countries in Japan. *Agricultural and Resource Economics Review*, **50**(1), 76–98.

Heien, Dale, & Sims, Eric N. 2000. The Impact of the Canada-United States Free Trade Agreement on U.S. Wine Exports. *American Journal of Agricultural Economics*, **82**(1), 173–182.

Irwin, Doug A. 2010. Trade Restrictiveness and Deadweight Losses from US Tariffs. *American Economic Journal: Economic Policy*, **2**(3), 111–133.

Johnson, Henry. 2022 (3). *The Impact of 2018 Wine Tariffs*. `https://github.com/henrymjohnson/the-effect-of-tariffs-on-wine-prices`.

Kee, Hiau Looi, Nicita, Alessandro, & Olarreaga, Marcelo. 2021. Import Demand Elasticities and Trade Distortions. 18.

Krugman, P. 2020. Trump is abusing his tariff power, too. Jan.

Lütkepohl, Helmut. 2005. *New introduction to multiple time series analysis*. Berlin: New York : Springer. OCLC: ocm61028971.

Nordhaus, W. 2018. The Trump Doctrine on International Trade: Part 1. Oct.

Phillips, P. C. B. 1987. Time Series Regression with a Unit Root. *Econometrica*, **55**(2), 277.

Pinilla, V., & Ayuda, M.-I. 2002. The political economy of the wine trade: Spanish exports and the international market, 1890-1935. *European Review of Economic History*, **6**(1), 51–85.

Ricardo, D. 1821. *On the Principles of Political Economy and Taxation.* 3 edn. McMaster University Archive for the History of Economic Thought.

Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations.* McMaster University Archive for the History of Economic Thought.

Stock, James H, & Watson, Mark W. 2001. Vector Autoregressions. 44.

U.S. Bureau of Economic Analysis. 2022. *U.S. Bureau of Economic Analysis, Real Disposable Personal Income [DSPIC96].* Tech. rept.

U.S. Bureau of Labor Statistics. 2022. *U.S. Bureau of Labor Statistics, Producer Price Index by Industry: Wineries: Wines, Brandy, and Brandy Spirits [PCU3121303121300].* Tech. rept.

U.S. v. France, Spain, Germany, U.K. 2019 (Oct.). *European Communities and Certain member States — Measures Affecting Trade in Large Civil Aircraft.* Tech. rept. World Trade Organization.

Zhang, Lisha, Onel, Gulcan, & Seale, James L. 2021. Aircraft Dispute, Wine, and the US–EU Trade War: Implications for US Consumers and Wine Suppliers. *Journal of Agricultural and Applied Economics*, July, 1–21.

APPENDIX A:

DATA

A.1 Raw Data

The raw data files can be found at https://github.com/henrymjohnson/the-effect-of-tariffs-on-wine-prices/data.

A.2 Data Exploration Code

Much of the data was in a wide format and needed to be converted to being a long format.

```
def wide_to_long(df, col_name):
    df = pd.melt(df, id_vars=['Year'], var_name='month', value_name=
                                            col_name)

    df['month'] = df['month'].map(month_map)
    df['month'] = df['Year'].astype(str) + '-' + df['month']

    df.drop(columns='Year', inplace=True)
    return df
```

The following method was used to format dates to guarantee consistency across data sources.

```
def last_day_of_month(month):
    if pd.isnull(month):
        return

    d = month.split('-')
    date = dt.date(pd.to_numeric(d[0]), pd.to_numeric(d[1]), 1)
    return date.replace(day = calendar.monthrange(date.year, date.
                                            month)[1])
```

Here's an example of code used to aggregate data points from the USITC into a single dataframe from being sheets in an excel spreadsheet. It utilizes the functions above to reorient the datasets and clean up the dates.

```
imports_ukfrspde_dict = pd.read_excel(ukfrspde_imports_data,
                                  sheet_name=None)

imports_ukfrspde_df = pd.DataFrame()
imports_ukfrspde_df['month'] = months
imports_uk_df = imports_ukfrspde_df.copy()
imports_fr_df = imports_ukfrspde_df.copy()
imports_de_df = imports_ukfrspde_df.copy()
imports_sp_df = imports_ukfrspde_df.copy()

for name, sheet in imports_ukfrspde_dict.items():
    if name != 'Query Parameters' and name != 'Query Results':
        sheet_df = pd.DataFrame(sheet)
        sheet_df.columns = sheet_df.iloc[0]
        sheet_df = sheet_df.iloc[1: , :]
        sheet_df = sheet_df[['Country','Year','JAN', 'FEB', 'MAR', '
                                          APR', 'MAY', 'JUN', 'JUL',
                                           'AUG', 'SEP', 'OCT', 'NOV
                                          ', 'DEC']]

        temp_fr_df = sheet_df.loc[(sheet_df['Country'] == 'France')]
        temp_de_df = sheet_df.loc[(sheet_df['Country'] == 'Germany')
                                          ]
        temp_sp_df = sheet_df.loc[(sheet_df['Country'] == 'Spain')]
        temp_uk_df = sheet_df.loc[(sheet_df['Country'] == 'United
                                          Kingdom')]
```

```
temp_all_df = sheet_df.loc[(sheet_df['Country'].str.contains
                                    ('Total'))]

temp_fr_df2 = wide_to_long(temp_fr_df, col_name=name)
temp_de_df2 = wide_to_long(temp_de_df, col_name=name)
temp_sp_df2 = wide_to_long(temp_sp_df, col_name=name)
temp_uk_df2 = wide_to_long(temp_uk_df, col_name=name)
temp_all_df2 = wide_to_long(temp_all_df, col_name=name)

temp_fr_df2['month'] = temp_fr_df2['month'].map(lambda x:
                                    last_day_of_month(x))
temp_fr_df2['month'] = temp_fr_df2['month'].astype('
                                    datetime64[ns]')
imports_fr_df = imports_fr_df.merge(temp_fr_df2, on='month')

temp_de_df2['month'] = temp_de_df2['month'].map(lambda x:
                                    last_day_of_month(x))
temp_de_df2['month'] = temp_de_df2['month'].astype('
                                    datetime64[ns]')
imports_de_df = imports_de_df.merge(temp_de_df2, on='month')

temp_sp_df2['month'] = temp_sp_df2['month'].map(lambda x:
                                    last_day_of_month(x))
temp_sp_df2['month'] = temp_sp_df2['month'].astype('
                                    datetime64[ns]')
imports_sp_df = imports_sp_df.merge(temp_sp_df2, on='month')

temp_uk_df2['month'] = temp_uk_df2['month'].map(lambda x:
                                    last_day_of_month(x))
```

```
        temp_uk_df2['month'] = temp_uk_df2['month'].astype('
                                          datetime64[ns]')
        imports_uk_df = imports_uk_df.merge(temp_uk_df2, on='month')

        temp_all_df2['month'] = temp_all_df2['month'].map(lambda x:
                                          last_day_of_month(x))
        temp_all_df2['month'] = temp_all_df2['month'].astype('
                                          datetime64[ns]')
        imports_ukfrspde_df = imports_ukfrspde_df.merge(temp_all_df2
                                          , on='month')

imports_ukfrspde_df.sort_values('month', ascending=True, inplace=
                                  True)
imports_ukfrspde_df.head()
```

The following was used to calculate various duty rate values.

```
# Duty Rate per liter of quantity
df['duty_r_ukfrspde'] = df['calculated_duties_adj_ukfrspde_imports']
                                   / df['quantity_ukfrspde_imports']
df['duty_r_italy'] = df['calculated_duties_adj_italy_imports'] / df[
                                  'quantity_italy_imports']
df['duty_r_australia'] = df['calculated_duties_adj_australia_imports
                                  '] / df['
                                  quantity_australia_imports']
df['duty_r_chile'] = df['calculated_duties_adj_chile_imports'] / df[
                                  'quantity_chile_imports']
df['duty_r_top3'] = df['calc_duties_top3_adj'] / df['quantity_top3']
df['duty_r_row'] = df['calculated_duties_adj_row_imports'] / df['
                                  quantity_row_imports']
```

```
# Charges, Insurance, and Freight rates per liter of quantity
df['cif_r_ukfrspde'] = df['
                                charges_insurance_freight_adj_ukfrspde_imports
                                '] / df['quantity_ukfrspde_imports
                                ']
df['cif_r_italy'] = df['charges_insurance_freight_adj_italy_imports'
                                ] / df['quantity_italy_imports']
df['cif_r_australia'] = df['
                                charges_insurance_freight_adj_australia_import
                                '] / df['
                                quantity_australia_imports']
df['cif_r_chile'] = df['charges_insurance_freight_adj_chile_imports'
                                ] / df['quantity_chile_imports']
df['cif_r_top3'] = df['cif_top3_adj'] / df['quantity_top3']
df['cif_r_row'] = df['charges_insurance_freight_adj_row_imports'] /
                                df['quantity_row_imports']
```

A.3 Modeling Code

A.3.1 Granger Causality test function

```
def grangers_causation_matrix(data, variables, endog, lag, test=test
                                , verbose=True):
    granger_df = pd.DataFrame(np.zeros((len(endog), len(variables)))
                                    , columns=variables, index=
                                    endog)
    for c in granger_df.columns:
        for r in endog:
```

```
            test_result = grangercausalitytests(data[[r, c]], maxlag
                                                =lag, verbose=False)
            p_values = [round(test_result[i+1][0][test][1],4) for i
                                                in range(lag)]
            if verbose: print(f'Y = {r}, X = {c}, P Values = {
                                                p_values}')
            min_p_value = np.min(p_values)
            granger_df.loc[r, c] = min_p_value
    granger_df.columns = [var + '_x' for var in variables]
    granger_df.index = [var + '_y' for var in endog]

    return granger_df
```

A.3.2 Stationarity Testing

The following code is used to test for stationarity.

```
def stationarity_tests(col):
   print('Augmented Dickey-Fuller Test:')
   unit_root_test = adfuller(col, autolag='AIC')
   dfoutput = pd.Series(unit_root_test[0:4], index=['t-stat:','p-
                                     value:','lags:','observations:'
                                     ])
   for key, value in unit_root_test[4].items():
      dfoutput['critical value (%s):' % key] = value
   print(dfoutput)
   print('\n')
   pp = PhillipsPerron(col)
   print(pp.summary().as_text())
```

The following code makes stationarity transformations using seasonality and trend decomposition using LOESS.

```
def make_stationary(x):
    adf_test = adfuller(x, autolag='AIC')
    if adf_test[0] > adf_test[4].get('1%'):
        dc = STL(x, seasonal_deg=1, trend_deg=1, robust=True).fit()
        adj = dc.observed - dc.trend - dc.seasonal - dc.weights

        dc_dict = {
            'adjusted': adj,
            'trend': dc.trend,
            'seasonal': dc.seasonal,
            'weights': dc.weights
        }

        return dc_dict
    else:
        return
```

A.3.3 Plotting Time Series'

The following is an example of the code used to plot the price time series.

```
price_line_2018_plot = sns.lineplot(data=train_df.loc['2018-01-01':'
                                    2021-12-31']['price'].rolling(3).
                                    mean())
price_line_2018_plot.set(title='Avg Wine Price in U.S. City 2018 to
                                    Present (3-Month Avg)', ylabel='
                                    Average Price', xlabel='Month')
# shade in the timespan of the additional tariff
price_line_2018_plot.axvspan(
    xmin=train_df['tariff'].where(train_df['tariff']).
```

```
                                    first_valid_index(),
    xmax=train_df['tariff'].where(train_df['tariff']).
                                    last_valid_index(),
    color='gray',
    alpha=0.2
)
price_line_2018_plot.yaxis.set_major_formatter(FuncFormatter(lambda
                                y, _: '\${:.2f}'.format(y)))
plt.gcf().set_size_inches(12, 8)
plt.show()
```

A.3.4 Johansen Test Code

```
def johansen_test(var_df, critical_value=0.05):
    output = coint_johansen(var_df, -1, 3)
    d = {'0.90': 0, '0.95': 1, '0.99': 2}
    trace_stat = output.lr1
    cvt_values = output.cvt[:, d[str(1-critical_value)]]
    cvm_values = output.cvm[:, d[str(1-critical_value)]]

    print('Cointegration Johansen Test:')
    print('{:<30}'.format('Variable') + '{:<30}'.format('T-Stat >
                                    CVs (trace, eig)') + '{:<20}'.
                                    format('Significant'))
    print('--'*36)
    for col, trace, cvt, eig in zip(var_df.columns, trace_stat,
                                    cvt_values, cvm_values):
        print('{:<30}'.format(col) + '{:<30}'.format('{:<7}'.format(
                                        format(trace, '.2f')) + \
```

```
                ' > ' + '{:<9}'.format('(' + format(cvt, '.2f') + ', ')
                                                  + '{:<7}'.format(
                                                  format(eig, '.2f')  +
                                                  ')')) + '{:<20}'.
                                                  format(str(trace > cvt
                                                  )))
```

A.3.5 Lag Order Selection Code

```
endog = ['price', 'bott', 'exp_q', 'q_bloc', 'q_it', 'q_cl', 'q_au',
                                    'q_row', 'prop_it', 'prop_au', '
                                   prop_cl', 'prop_bloc']
exog = ['disp_inc', 'duty_r_bloc', 'tariff', 'duty_r_cl', 'duty_r_au
                                   ', 'duty_r_it', 'duty_r_row', '
                                   cif_world']
train_df['tariff'] = train_df['tariff'].astype(int)

model = VAR(endog=train_df[endog], exog=train_df[exog])
lag_orders = model.select_order(maxlags=16)
lag_orders.summary()
```

A.3.6 Modeling Code

Fitting the model

```
endog = ['price_s', 'duty_r_bloc_s', 'bott_s', 'exp_q_s', 'q_bloc_s'
                                   , 'q_it_s', 'q_cl_s', 'q_au_s', '
                                   q_row_s', 'prop_it_s', 'prop_au_s'
                                   , 'prop_cl_s', 'prop_bloc_s']
exog = ['tariff', 'disp_inc_s', 'duty_r_cl_s', 'duty_r_au_s', '
                                   duty_r_it_s', 'duty_r_row_s', '
```

```
                                    cif_cl_s', 'cif_r_au_s', '
                                    cif_r_it_s', 'cif_r_bloc_s']
train_df['tariff'] = train_df['tariff'].astype(int)

model = VAR(endog=train_df[endog], exog=train_df[exog])
tsmf = model.fit(1)
tsmf.summary()
```

Impulse Response Example

```
impulse_response = tsmf.irf(6)
ir_plt = impulse_response.plot(impulse='duty_r_bloc_s', response='
                                    price_s', signif=0.1, stderr_type=
                                    'mc')
plt.gcf().set_size_inches(8, 8)
plt.savefig('../figures/impulse-response_duty_r_bloc_on_price.png')
plt.show()
```

Forecasting Example

```
prices = tsmf.fittedvalues['price_s'] + train_df['price_seas'] +
                                    train_df['price_trend'] + train_df
                                    ['price_weights']
prices
price_data = pd.DataFrame()
price_data['pred_price'] = prices
price_data = price_data.merge(train_df['price'], left_index=True,
                                    right_index=True)

price_plot = price_data[-36:].plot(title='Actual vs Predicted Price'
                                    , ylabel='Price', xlabel='Month')
```

```
plt.show()
```

Durbing Watson Test

```
dw_output = durbin_watson(tsmf.resid)
mod_columns = ['price_s', 'duty_r_bloc_s', 'bott_s', 'exp_q_s', '
                                   q_bloc_s', 'q_t3_s', 'q_row_s', '
                                   prop_t3_s', 'prop_bloc_s', 'tariff
                                   ', 'disp_inc_s', 'duty_r_t3_s', '
                                   duty_r_row_s', 'cif_r_bloc_s']


for col, val in zip(mod_columns, dw_output):
    print('{:<62}'.format(str(col), ':'), round(val, 3))
```

APPENDIX B:

TEST RESULTS

B.1 Granger Causality Test

Table B.1: Granger-Causality Test Results for Variables

	price_x	bott_x	disp_inc_x	exp_q_x	tariff_x	q_bloc_x	q_t3_x	q_row_x	prop_bloc_x	duty_r_bloc_x	cif_r_bloc_x	q_it_x	prop_it_x
price_y	1.0000	0.0000	0.0000	0.0096	0.0486	0.0000	0.0000	0.0000	0.0000	0.1081	0.0127	0.0000	0.0000
duty_r_bloc_y	0.7280	0.2053	0.0003	0.5204	0.0000	0.0333	0.1287	0.4104	0.7311	1.0000	0.7877	0.1668	0.6392
bott_y	0.0000	1.0000	0.0000	0.0020	0.0411	0.0000	0.0000	0.0000	0.0002	0.0247	0.0000	0.0000	0.0000
exp_q_y	0.0440	0.0000	0.1369	1.0000	0.4222	0.0002	0.0000	0.0319	0.0000	0.3796	0.0000	0.0001	0.0000
q_bloc_y	0.0000	0.0000	0.0000	0.0000	0.0074	1.0000	0.0000	0.0000	0.0000	0.0024	0.0000	0.0000	0.0000
q_t3_y	0.0000	0.0000	0.0000	0.0002	0.0904	0.0002	1.0000	0.0000	0.0085	0.1541	0.0000	0.0000	0.0000
q_row_y	0.0008	0.0000	0.0000	0.0257	0.0566	0.0423	0.2956	1.0000	0.1775	0.0019	0.0487	0.0078	0.1531

	duty_r_it_x	cif_r_it_x	q_au_x	prop_au_x	duty_r_au_x	cif_r_au_x	q_cl_x	prop_cl_x	duty_r_cl_x	cif_r_cl_x	duty_r_t3_x	prop_row_x	cif_r_row_x	duty_r_row_x
price_y	0.005	0.0001	0.0038	0.0504	0.0000	0.0000	0.0000	0.0439	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
duty_r_bloc_y	0.045	0.6660	0.2551	0.5557	0.2415	0.5028	0.0737	0.2330	0.2669	0.4201	0.0520	0.0878	0.5378	0.0766
bott_y	0.000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0069	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
exp_q_y	0.000	0.0425	0.0000	0.0168	0.0708	0.0015	0.0383	0.2230	0.0841	0.2014	0.1674	0.0556	0.0747	0.0218
q_bloc_y	0.000	0.0068	0.0000	0.0008	0.0000	0.0000	0.0250	0.1318	0.0000	0.0002	0.0000	0.0000	0.0000	0.0000
q_t3_y	0.000	0.0600	0.0027	0.0005	0.0000	0.0000	0.0343	0.0266	0.0000	0.0001	0.0001	0.0000	0.0000	0.0049
q_row_y	0.000	0.0863	0.3191	0.2858	0.0000	0.0000	0.4073	0.3853	0.0000	0.0155	0.0030	0.0006	0.0002	0.0000

B.2 Johansen Cointegration Test

Table B.2: Johansen Cointegration Test Results with Significance at the 95% Confidence Level.

Variable	T-Stat	Critical Values (trace, eig)	Significant
price_diff1	1085.07	(179.52, 54.96)	True
calc_duty_diff1_diff12	825.38	(143.67, 48.88)	True
calc_duty_rate_diff1_diff12	653.64	(111.78, 42.77)	True
bottled_diff1_diff12	499.00	(83.94, 36.63)	True
exp_q_diff1_diff12	351.33	(60.06, 30.44)	True
prop_imp_diff1_diff12	233.72	(40.17, 24.16)	True
imp_q_w_diff1_diff12	152.35	(24.28, 17.80)	True
imp_q_diff1_diff12	87.48	(12.32, 11.22)	True
cif_w_diff1_diff12	27.21	(4.13, 4.13)	True

B.3 Durbin Watson

Table B.3: Durbin Watson statistic for serial correlation.

Variable	Statistic
price_diff1	1.639
bottled_diff1_diff12	2.356
exp_q_diff1_diff12	2.171
calc_duty_diff1_diff12	1.98
calc_duty_rate_diff1_diff12	2.009
cif_w_diff1_diff12	2.165
imp_q_w_diff1_diff12	2.254
imp_q_diff1_diff12	2.128
prop_imp_diff1_diff12	2.109