Module 1E: Storage and Potato Prices

# Learning Outcomes

After completing this module you should be able to:

* Define the concept of the intertemporal law-of-one-price (LOP).
* Describe the relationship between cold storage and seasonal prices in the potato markets of Indian and Bangladesh.
* Estimate the monthly cost of cold storage as a percentage of the average selling price in the Indian potato market.
* Estimate the seasonal low to high price gap as a percentage of the average selling price in the Indian potato market.
* Use a dummy variable econometric model to estimate the seasonal low to high price gap in the U.S. market.

# Introduction to Intertemporal Law-of-One-Price (LOP)

Modules 1A through 1D examined commodity prices over space, using the lens of the spatial LOP. We now shift gears and begin to look at commodity prices over time, using the lens of the intertemporal LOP. The basic idea is quite simple. If a commodity is being stored through time then the expected future price of the commodity must be equal to the current price of the commodity plus the cost of storing the commodity. If the price difference was larger then arbitrage profits could be earned, and large scale arbitrage would bid up the current price and bid down the expected future price through additional storage. If the commodity is storable but is currently not being stored then it must be the case that the difference in the expected future price and the current price is less than the cost of storage. We refer to this situation as a “stock out”. These two conditions define the intertemporal LOP. More formally,

The intertemporal LOP tell us that:

* If a commodity is being stored then the difference between the current price and the expected future prices must equal the cost of storing the product between those two periods.
* If the commodity is not being stored (i.e., a stock out) then the price difference must be less than the cost of storage.

All commodities are storable to a certain degree (e.g., two weeks for the case of lettuce) and so it is important to be clear what we mean by storable. In this course we focus on storable crops with a single annual harvest. We do this because we are particularly interested in on how the price adjusts while the commodity is in storage and is gradually released for consumption (i.e., over the course of the marketing year). In this module we will focus on potatoes, which can be stored no more than a year (typically 8 months maximum). In future modules we will focus on corn, which can be carried across multiple years.

# Seasonality and Potato Prices in India and Bangladesh

It is useful to begin our discussion of storage in the context of food security. In Sub-Saharan Africa we saw that the problem was a lack of market integration, which meant that food did not quickly and efficiently move through space from regions with a surplus to regions with a shortfall. We now examine intertemporal aspect of food security. Specifically, the concern is that food may be in surplus during one time of the year, and in shortfall during another time of the year. Food prices will be low when food is in surplus and high when food is scarce. Storage is one solution to this problem. In a developed country food is stored at relatively low cost as a result of high capital investment in storage facilities. In developing countries storage facilities are often not available and are relatively costly when they are available. Hopefully you can see from the intertemporal LOP that a higher cost of storage implies a larger gap in the price of food when it is in surplus versus when it is scarce. The seasonal nature of commodity prices (i.e., the size of the price gaps) can be viewed as an intertemporal measure of market integration.

Consider the case of an important crop (potatoes) in two large developing countries: India and Bangladesh. India has a population of approximately 1.4 billion and Bangladesh has a population of about 165 million. Both countries have a high population density but Bangladesh particularly so given that the land mass of India is about 22 times larger than that of Bangladesh. These two South Asian countries are neighbors with Bangladesh situated to the east of India. Almost the entire population of Bangladesh consist of Bangla-speaking people (i.e., Bengalis), which is the third-largest ethnic group in the world (the first being the Han Chinese and the second being Arabs). India is much more diverse with over two thousand ethnic groups, a representation of all of the world’s major religions and four major languages ((Indo-European, Dravidian, Austroasiatic and Sino-Tibetan).

Per capita consumption of potatoes has grown steadily in both countries. For example, according to [PotatoPro](https://www.potatopro.com/bangladesh/potato-statistics), potato consumption in Bangladesh has grown steadily, rising from 7 kg per capita per year in 1990 to 24 kg per capita per year in 2005. Potatoes are a winter crop (October to March), mainly grown in the Punjab and other northern regions of India, and toward the south near the capital city of Dhaka in Bangladesh. Potatoes are typically grown for cash sale (versus subsistence consumption) and for this reason are important source of income for small farmers in both countries. In Bangladesh potatoes are second only to rice in terms of value of production.

Potato production and consumption has also grown in other developing countries. According to the [FAO](http://www.fao.org/3/i0200e/i0200e07.htm), potatoes are now the fourth most important global commodity (following maize, rice and wheat), and over half of global potato consumption takes place in developing countries. Potatoes are popular because of their high yield per unit of land and also because potatoes fit well into multi-cropping systems.

## Economics of Cold Storage for Potatoes in India

Potatoes which are intended to be stored for more than a couple of months must be placed into cold storage. In India potatoes which are not placed into cold storage are typically kept on the farm in “rustic” storage or other low-tech storage facilities. Rustic storage generally involves covering the potatoes with straw and plastic to keep them relatively cool, dry and dark. Farmers have reported losses of up to 40 percent when using rustic storage systems. Storage losses in well-constructed cold storage facilities are much smaller (rates of spoilage varies from year to year and depend on potato variety).

Farmers store potatoes to both reduce spoilage and to sell the potatoes at a higher price. If and respectively denote the quantity and the selling price with a rustic system (e.g., two months after harvest), and and denote the quantity and the selling price with a cold storage system (e.g., six months after harvest), then it follows that the profits for a farmer who rents cold storage facilities can be expressed as where is the rental cost of the cold storage service. When comparing a rustic system to a cold storage system, the percent difference between and is quite large (e.g., 25 - 30). As will be shown below, the percent difference between and is also quite large. The strong growth in the use of cold storage implies that must typically be positive.

The cost of placing potatoes in cold storage varies considerably from year to year. Given that cold storage has a fixed supply and is used to store an assortment of vegetables and fruits, it is likely that cold storage rental costs rise (fall) with a larger (smaller) than average potato crop. This relationship makes the evaluation of the spatial LOP difficult because the expected potato price differences across two periods will depend on both the average cost of storage and current production levels relative to average production levels.

A 2017 [report](https://timesofindia.indiatimes.com/city/agra/agra-cold-storages-to-throw-out-18-75-lakh-tonnes-of-potatoes-by-nov-end/articleshow/59933564.cms) indicates that “the charge for keeping potatoes in cold storage is Rs 220 per quintal, while the entire cost of growing potatoes is around Rs 850 per quintal, including cold storage costs.” (a quintal is 100 kg). To verify the accuracy of this point estimate let’s use a second source of data. In a February 2019 [report](https://www.downtoearth.org.in/news/agriculture/punjab-s-cold-storage-owners-struggle-to-stay-in-business-63240) it was noted that the owner of a cold storage facility charges Rs 80 - Rs 110 for storing a 50 kg bag of potatoes for a season (a “Rs” is an Indian Rupee). One quintal is 100 kg and so this storage cost is equivalent to RS 160 - Rs 220 per quintal. This compares well to the Rs 220 quintal storage cost quote from the previous report. These costs will be compared to the average selling price of potatoes in the next section.

## Potato Selling Prices in India

An Indian government [report](https://agricoop.gov.in/sites/default/files/Monthly_Report_on_Potato_for_July__2019.PDF) provides monthly average potato prices in Delhi for the years 2015 to 2018. Let’s take a closer look at these prices:

pacman::p\_load(here, dplyr, ggplot2, lubridate)  
data\_delhi <- read.csv(here("Data", "potato\_delhi.csv"), header=TRUE, sep=",", stringsAsFactors = FALSE)  
data\_delhi

## month yr.2015 yr.2016 yr.2017 yr.2018 average stdev  
## 1 Jan 674.1 485.8 500.3 479.9 535.0 93.1  
## 2 Feb 603.9 527.5 470.2 464.7 516.6 64.8  
## 3 Mar 589.1 728.2 425.6 677.3 605.1 132.7  
## 4 Apr 446.1 927.3 428.5 829.0 657.7 257.8  
## 5 May 470.9 1212.7 410.6 1036.4 782.7 402.1  
## 6 Jun 743.2 1425.1 548.5 1116.1 958.2 390.3  
## 7 Jul 989.5 1465.7 671.9 1113.6 1060.2 328.2  
## 8 Aug 852.0 1445.9 585.3 1292.9 1044.0 396.1  
## 9 Sep 907.5 1322.8 636.4 1354.2 1055.2 345.6  
## 10 Oct 1204.6 1169.4 651.9 1423.2 1112.3 326.8  
## 11 Nov 942.6 837.1 703.6 1243.8 931.8 229.9  
## 12 Dec 613.9 503.5 626.0 592.9 584.1 55.4

We need to eliminate the month column in the data frame for the purpose of calculating summary statistics. After doing this we can calculate the mean of the four years worth of monthly prices.

# removes first column and creates new df called data\_delhi2  
data\_delhi2 <- data\_delhi[,-1]  
  
# creates new vector called mean\_annual\_price which takes the column means of data\_delhi  
 # sapply() function takes a df or vector, performns a function, and preserves the length of the original df   
 # the parentheses around the code just means we want R to print the output and store the result to mean\_annual\_price. If we didn't have the parentheses, then we'd have to run "mean\_annual\_price" to see the resulting vector  
(mean\_annual\_price <- sapply(data\_delhi2, mean))

## yr.2015 yr.2016 yr.2017 yr.2018 average stdev   
## 753.1167 1004.2500 554.9000 968.6667 820.2417 251.9000

# another option: dplyr function to calculate column mean  
data\_delhi %>% summarise\_if(is.numeric, mean)

## yr.2015 yr.2016 yr.2017 yr.2018 average stdev  
## 1 753.1167 1004.25 554.9 968.6667 820.2417 251.9

From the previous section we see that Rs 200 per quintal is a reasonable value to use for the cost of renting cold storage for the season. Let’s divide this storage cost value by the mean annual price of potatoes (calculated above) in order to determine the fraction of the average price which is allocated to storage costs.

(storage\_share <-200/mean\_annual\_price)

## yr.2015 yr.2016 yr.2017 yr.2018 average stdev   
## 0.2655631 0.1991536 0.3604253 0.2064694 0.2438306 0.7939659

Before discussing these results we can calculate the minimum and the maximum of the monthly prices for each of the four years:

(min\_annual\_price <- sapply(data\_delhi2, min))

## yr.2015 yr.2016 yr.2017 yr.2018 average stdev   
## 446.1 485.8 410.6 464.7 516.6 55.4

(max\_annual\_price <- sapply(data\_delhi2, max))

## yr.2015 yr.2016 yr.2017 yr.2018 average stdev   
## 1204.6 1465.7 703.6 1423.2 1112.3 402.1

Finally, we can calculate the difference between the minimum and the maximum monthly price (i.e., the gap) and express this difference as a fraction of the average monthly price.

(gap\_fraction <- (max\_annual\_price - min\_annual\_price)/mean\_annual\_price)

## yr.2015 yr.2016 yr.2017 yr.2018 average stdev   
## 1.0071481 0.9757530 0.5280231 0.9895045 0.7262494 1.3763398

What can we learn from this price data and the summary statistics?

1. As expected, prices tend to be lowest at the time of harvest (March - April) and highest at the time of planting, which is toward the end of the maximum eight month storage cycle (September - October).
2. Storage costs are typically 20 - 25 percent of the mean wholesale price of potatoes. This fraction is much higher if cost is compared to the harvest price and is much lower if compared to the planting price.
3. The gap between the minimum and maximum annual price is typically in the range of Rs 600 to Rs 700 per quintal. When expressed as a fraction of the average price, the gap is typically around 100 percent (2017 is clearly an outlier).
4. The price premium which can be achieved through the use of cold storage is much higher than the estimated Rs 200 per quintal cost of storage. Thus, cold storage is typically a profitable strategy for farmers.
5. There is considerable price variation across years. The cross-year coefficient of variation can be calculated by dividing the standard deviation in the last column by the average in the second last column. Without doing these calculations it can be seen that the coefficient of variation tends to be 30 percent or higher. This high level of price variation is problematic for income insecure farmers and for food insecure households.
6. 2017 saw a record crop of potatoes harvested and a policy-induced shortfall in demand. This double “whammy” caused monthly prices to drop by roughly 50 percent. If you re-read the 2017 [report](https://timesofindia.indiatimes.com/city/agra/agra-cold-storages-to-throw-out-18-75-lakh-tonnes-of-potatoes-by-nov-end/articleshow/59933564.cms) on potato prices in India you will learn that many farmers forfeited their potatoes (i.e., did not claim them from the storage facilities), which resulted in a large volume of potatoes being thrown away by the owners of the cold storage facilities. The media was very critical of this waste given the high levels of food insecurity in India.

## Seasonality in Bangladesh Potato Prices

A 2021 [study](http://bjae.bau.edu.bd/home/article/view/100) titled “Can Cold Storage Reduce Seasonal Variation in Prices of Agricultural Commodities/ A Case of Potato in Bangladesh” focuses on the seasonality of potato prices. The authors of the paper are particularly interested in knowing the effectiveness of cold storage as a mechanism for reducing the size of potato seasonal pricing gaps. The authors show that prior to 1988 there was very little cold storage capacity for potatoes in Bangladesh. Cold storage capacity has grown steadily since 1988 but its growth has not kept pace with the growth in potato production. Indeed, as of 2016 only about 60 percent of Bangladesh potato production was placed in cold storage.

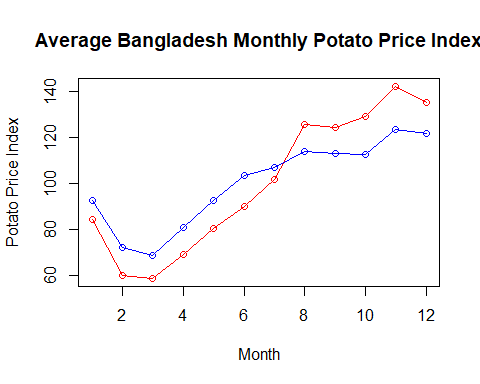
The 2021 [study](http://bjae.bau.edu.bd/home/article/view/100) on seasonality in potato prices in Bangladesh calculated a long term average monthly price index for potatoes in order to identify how potato prices typically change over the course of a season. As noted, the authors are interested in comparing the seasonal pattern of prices in the 1987 - 2018 period, which corresponds to heavy private-market investment in potato cold storage, with the seasonal pattern of prices in the earlier 1972 - 1987 period, which corresponds to a time of very little cold storage for potatoes.

The data in Table 1 of the 2021 Bangladesh study is not available in digital format and so it must be manually entered into a data frame. This is done as follows:

index\_before <- c(84.39, 59.86, 58.66, 69.15, 80.37, 90.06, 101.73, 125.46, 124.19, 129.06, 142.00, 135.02)  
index\_after <- c(92.44, 71.86, 68.46, 80.65, 92.26, 103.54, 106.67, 113.65, 112.98, 112.57, 123.21, 121.65)

This pair of index values can now be graphed:

plot(index\_before,type = "o",col = "red", xlab = "Month", ylab = "Potato Price Index",   
 main = "Average Bangladesh Monthly Potato Price Index")  
  
lines(index\_after, type = "o", col = "blue")



The graph shows the strong seasonality in potato prices, similar to what was seen for potato prices in India. The red schedule shows the average monthly price index before the large-scale investment in cold storage facilities, and the blue schedule shows the same index after the investment. It is clear that the investment significantly raised the low prices during the spring harvesting period, and lowered the high prices during the fall planting period. Similar to India, considerable seasonality in the pricing remains.

Later in this module we will learn that prices are expected to rise over the course of the marketing year, even if all potatoes are in cold storage. Indeed, the price must rise month by month at a rate which reflects the monthly cost of storage. As was seen in the previous section, the cost of storage is a sizable fraction of the commodity’s price. For India, the cost of storage was in the range of 20 - 25 percent of the average potato price. With eight months of storage this works out to a monthly storage cost of about three percent of the average potato price. It is beyond to scope of this analysis to determine if the actual average price rise is about three percent per month. Nevertheless, it is important to keep in mind the strong connection between the monthly price increase and the monthly cost of storage.

## Related Issues

If you reread the 2019 [report](https://www.downtoearth.org.in/news/agriculture/punjab-s-cold-storage-owners-struggle-to-stay-in-business-63240) on potato cold storage in India you will learn that despite the seemingly high return on using cold storage that is achieved by Indian potato farmers, the prices charged by cold storage suppliers are in some years barely enough to cover their operating and investment costs. The particular problem being discussed (i.e., cold storage owners not being paid by farmers) stems from farmers being unable to sell their potatoes due to India’s demonetisation. The more general problem appears to be problems with regional supply chains such as excess cold storage capacity in some regions and inadequate cold storage capacity in other regions. The authors emphasize that rotting potatoes when the technology exists to solve the problem appears highly problematic for a country with high rates of food insecurity. The Indian government currently subsidizes cold storage investment but it seems that a higher level of subsidy is required to solve the joint problem of rotting potatoes and cash flow problems for owners of cold storage facilities.

It is important to keep in mind that the previously-quoted storage costs do not include capital costs for farmers, either in the form of foregone interest if the farmer is a net saver, or interest payments if the farmer is carrying debt. In a developing country the cost of capital tends to be high, and this is an important component of the cost of storage. The cost of capital in India is in the 12 - 15 percent range and potatoes are typically in storage for six months. This means that an additional 6 - 7.5 percent of the potato price must be considered an additional implicit cost of storage.

It is also important to note that potatoes do not trade in international markets. This means that an over (under) production of potatoes can drive prices down (up) to relatively low (high) levels because there are no adjustments in imports or exports to stabilize the price. The current wholesale [price](https://www.tridge.com/intelligences/potato/IN/price) of fresh potatoes in India is about USD$0.14/kg. This compares to a potato price of about USD$0.85/kg in the U.S. This large cross country differential is quite different than the situation with tradable commodities such as palm oil. For internationally traded commodities the prices in most countries equal the international price with adjustments for the currency and transportation costs.

# Seasonality and Potato Prices in the U.S.

## Background

To familiarize yourself with cold storage for vegetables in the U.S. read the following [report](https://suncoastsvn.com/us-cold-storage-demand/). It is interesting to see the connection between the rapid growth in e-commerce and the corresponding surging investment in cold storage facilities. It is of interest to compare the reported 2018 cold storage capacity of 3.6 billion cubic feet in the U.S. to the 150 million cubic meters of cold storage capacity in India (see [here](https://www.downtoearth.org.in/news/agriculture/punjab-s-cold-storage-owners-struggle-to-stay-in-business-63240)). There are about 35 cubic feet in a cubic meter, which means that the U.S. capacity is roughly 100 million cubic meters – about two thirds that of India. Of course the population of India is a little more than four times that of the U.S.. and so on a per capita basis the U.S. has more cold storage capacity than India, which is to be expected.

How expensive is it to store potatoes in the U.S. in comparison to India? We learn from this [report](https://potatoassociation.org/wp-content/uploads/2014/04/A_ProductionHandbook_Final_000.pdf) that in 2007 the operating cost of a 100,000 ton storage facility in the U.S. state of Idaho worked out to $14.75 per ton. The ownership costs of the storage facility are not considered in the analysis because they are viewed as sunk from a decision making perspective. The base price of potatoes in 2007 was $110 per ton, which means that the storage operating costs represent about 14.75/110 (13.5 percent) of the price of the potato. This is well below the 0.20 - 0.25 ratio of operating cost to wholesale price of potatoes in India. One possible reason for this difference is that the price of potatoes is much lower in India than in the U.S.

Given that monthly storage costs for potatoes are lower in the U.S. as compared to India, we should expect weaker seasonality in U.S. potato prices as compared to India (i.e., smaller annual price gaps). As previously noted, according to the intertemporal LOP, the price of potatoes should rise at a rate equal to the monthly cost of storage. Given that potatoes are in storage for say six months on average, it follows that the monthly potato price should increase by 0.135/6 = 0.0225 (2.25 percent) in the U.S. and 0.225/6 = 0.0375 (3.75 percent) in India. This means that the trough to peak in U.S. potato prices should be about 40 percent lower than the trough to peak in Indian potato prices.

## Examining the U.S. Consumer Price Index (CPI) for Potatoes

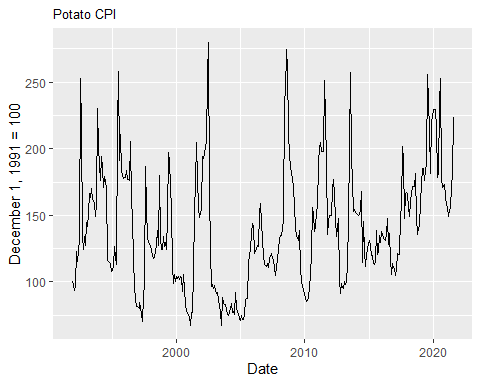
Data which shows the monthly average potato price in the U.S. does not appear to be available. This means that the U.S. potato price gaps must be estimated using a time series of potato prices. The St. Louis Federal Reserve Economic Data (FRED) has a monthly consumer price index (CPI) for potatoes with December 1991 serving as the base year (i.e., an index value equal to 100). Let’s read in this data to begin the formal analysis.

data <- read.csv(here("Data", "potato\_cpi\_data.csv"), header=TRUE, sep=",", stringsAsFactors = FALSE)  
data$period <- as.Date(data$observation, format = "%m/%d/%Y")  
head(data)

## observation potato\_cpi period  
## 1 12/1/1991 100.0 1991-12-01  
## 2 1/1/1992 98.3 1992-01-01  
## 3 2/1/1992 93.9 1992-02-01  
## 4 3/1/1992 97.0 1992-03-01  
## 5 4/1/1992 122.9 1992-04-01  
## 6 5/1/1992 115.7 1992-05-01

Let’s plot this data to get a general idea of the pricing patterns.

plot\_cpi <- ggplot(data, aes(x = period, y = potato\_cpi)) +   
 geom\_line() +   
 labs(title = "Potato CPI", y= "December 1, 1991 = 100", x = "Date") +   
 theme(plot.title = element\_text(size=10))  
  
plot\_cpi



Because the data is in index form with 100 as the first data point, it is acceptable to interpret the Y axis of the previous chart as percentage change relative to the 1991 base year price of potatoes. Notice that there has not been much of an upward trend in U.S. potato prices over the past 30 years. The high price volatility reflects the highly inelastic demand for potatoes combined with varying supply (mainly due to weather shocks and concentrated regional production) and a lack of international trade. Of course highly volatile potato prices are much less problematic for U.S. farmers who are protected by government crop loss subsidies and for relatively food secure U.S. consumers, as compared to the case of Indian farmers and households.

It might be tempting to program R to group the data by month and then average each group to construct monthly averages. This approach is acceptable if the data is stationary but it is not acceptable when working with non-stationary data. There is likely to be a semester 2 course in the MFRE program on time series analysis, which will emphasize the importance of not using non-stationary data when calculating summary statistics such as means and standard deviations. One could use the well-known Dickey-Fuller test to show that the potato CPI is non-stationary. We won’t do this because the test will almost certainly fail to reject the null hypothesis that the data is non-stationary (i.e., has a unit root) due to both seasonality and trends (most commodity prices are non-stationary and we should not expect potato prices to be any different).

The standard method of making non-stationary data stationary is to take the first difference. We can take the month by month first difference (e.g., March 2021 - February 2021) or the annual first difference (e.g., March 2021 - March 2020, February 2021 - February 2020). The latter approach is often used when there is strong seasonality in the data. Ideally we should estimate pricing seasonality using both methods. However, to keep things simple only the first method is used in the analysis below.

It is useful to examine a plot of the price index in first difference format. Begin by generating a new data series which is the first difference of the imported data series. The goal is to create a new data frame called *data\_diff*. We will then bind the first difference of the price index to *data\_diff*. To prepare for this transformation let’s eliminate the first column of *data* because it contains the imported dates, which are not being used in this analysis. Let’s also eliminate the first row so that the new data set will have the same number of rows as the set of first differenced index values.

data\_diff <- data[-1,-1]

Now difference the price index, bind it to *data\_diff* and eliminate the first column because there is no need to have two date columns.

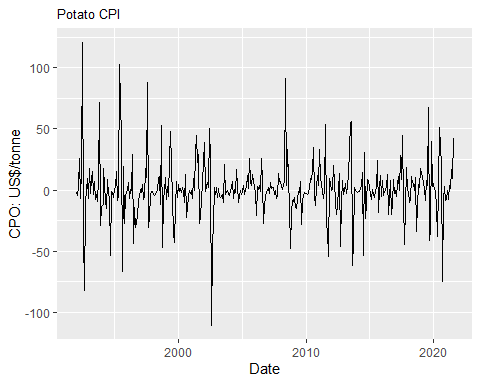
data\_diff$cpi\_diff <- diff(data$potato\_cpi)  
data\_diff <- data\_diff[,-1]  
head(data\_diff)

## period cpi\_diff  
## 2 1992-01-01 -1.7  
## 3 1992-02-01 -4.4  
## 4 1992-03-01 3.1  
## 5 1992-04-01 25.9  
## 6 1992-05-01 -7.2  
## 7 1992-06-01 16.5

# Alternative way  
# data\_diff <- data %>%  
# mutate(cpi\_diff = potato\_cpi - lag(potato\_cpi)) %>% #calc first diff  
# filter(year(period) > 1991) %>% # filter only if year > 1991  
# select(period, potato\_cpi, cpi\_diff) # select relevant variables

Now we can plot the price index in first difference format.

plot\_cpi\_diff <- ggplot(data\_diff, aes(x = period, y = cpi\_diff)) +   
 geom\_line() +   
 labs(title = "Potato CPI", y= "CPO: US$/tonne", x = "Date") +   
 theme(plot.title = element\_text(size=10))  
  
plot\_cpi\_diff



## Dummy Variables to Estimate Seasonality

Let be a dummy variable that takes on a value of 1 if and 0 otherwise. In this formulation, both and are a month index, with corresponding to January, corresponding to February, etc. This means there are 12 monthly dummy variables corresponding to the 12 months. If the current month in the data set is January then the January dummy takes on a value of 1 and all of the other dummies take on a value of 0. If the current month is February then the February dummy takes on a value of 1 and all of the other dummies take on a value of 0. And so forth.

Let be the price index for U.S. potatoes in month of year . The regression model which has as the dependent variable and eleven monthly dummies as the explanatory variables can be written as follows.

Notice that the December dummy has been left out. This is required because the regression includes an intercept.

We can interpret as the difference between the average potato index for month and the average potato index for December. In the U.S. potatoes are harvested in October and so we expect a relatively low value for the December price index. This means that through should take on positive values, because the price index for potatoes is expected to be higher in January through September than the price index for December. This also means that and should take on negative values because the price index for potatoes is expected to be lower in October and November than the price index for December.

As noted above, the previous equation should not be estimated as is because the price index is quite likely to be non stationary. Instead we will estimate the first difference of the equation:

Fortunately the coefficients are preserved and so the estimates of these coefficients can be interpreted as discussed above. For example, the estimate of can be interpreted as the difference between the average January price index minus the average December price index.

## Estimating the Differenced Dummy Variable Model

To estimate the dummy variable model we must create a set of monthly dummy variables, beginning with December of 2011. There are several different ways to do this. A method which highlights the power of R uses the following code. See *potatoes\_code\_only.R* for more details about this method.

df\_month <- factor(month(data$period))  
dummies <- model.matrix(~df\_month+0)  
head(dummies)

## df\_month1 df\_month2 df\_month3 df\_month4 df\_month5 df\_month6 df\_month7  
## 1 0 0 0 0 0 0 0  
## 2 1 0 0 0 0 0 0  
## 3 0 1 0 0 0 0 0  
## 4 0 0 1 0 0 0 0  
## 5 0 0 0 1 0 0 0  
## 6 0 0 0 0 1 0 0  
## df\_month8 df\_month9 df\_month10 df\_month11 df\_month12  
## 1 0 0 0 0 1  
## 2 0 0 0 0 0  
## 3 0 0 0 0 0  
## 4 0 0 0 0 0  
## 5 0 0 0 0 0  
## 6 0 0 0 0 0

dimnames(dummies)[[2]] <- month.abb

This set of dummies contains a variable for December. Recall that we will estimate the model without the December dummy. To eliminate the December dummy use:

dummies <- dummies[,-12]

We can now take the first difference of the dummy variables and store in *dummies\_diff*:

dummies\_diff <- diff(dummies)  
head(dummies\_diff)

## Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov  
## 2 1 0 0 0 0 0 0 0 0 0 0  
## 3 -1 1 0 0 0 0 0 0 0 0 0  
## 4 0 -1 1 0 0 0 0 0 0 0 0  
## 5 0 0 -1 1 0 0 0 0 0 0 0  
## 6 0 0 0 -1 1 0 0 0 0 0 0  
## 7 0 0 0 0 -1 1 0 0 0 0 0

Let’s create the final data set by binding together *data\_diff* and *dummies\_diff*.

data\_full <- cbind(data\_diff,dummies\_diff)   
head(data\_full)

## period cpi\_diff Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov  
## 2 1992-01-01 -1.7 1 0 0 0 0 0 0 0 0 0 0  
## 3 1992-02-01 -4.4 -1 1 0 0 0 0 0 0 0 0 0  
## 4 1992-03-01 3.1 0 -1 1 0 0 0 0 0 0 0 0  
## 5 1992-04-01 25.9 0 0 -1 1 0 0 0 0 0 0 0  
## 6 1992-05-01 -7.2 0 0 0 -1 1 0 0 0 0 0 0  
## 7 1992-06-01 16.5 0 0 0 0 -1 1 0 0 0 0 0

The *data\_full* data frame has the required variables to run the desired regression, which is given by

The code for running this regression is as follows (the “0” added at the end of the dummy variables tells R to suppress the intercept):

reg\_cpi <- lm(cpi\_diff ~ Jan + Feb + Mar + Apr + May + Jun + Jul + Aug + Sep + Oct + Nov + 0, data = data\_full)  
summary(reg\_cpi)

##   
## Call:  
## lm(formula = cpi\_diff ~ Jan + Feb + Mar + Apr + May + Jun + Jul +   
## Aug + Sep + Oct + Nov + 0, data = data\_full)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -119.562 -6.881 -0.992 6.733 111.095   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## Jan 2.0487 3.5959 0.570 0.5692   
## Feb 0.6908 4.8517 0.142 0.8869   
## Mar 2.4729 5.6414 0.438 0.6614   
## Apr 3.6283 6.1474 0.590 0.5554   
## May 2.8204 6.4369 0.438 0.6615   
## Jun 20.6791 6.5387 3.163 0.0017 \*\*   
## Jul 29.9845 6.4617 4.640 4.95e-06 \*\*\*  
## Aug 39.0466 6.1993 6.299 9.15e-10 \*\*\*  
## Sep 13.8798 5.7053 2.433 0.0155 \*   
## Oct -3.5526 4.9178 -0.722 0.4705   
## Nov 2.0254 3.6517 0.555 0.5795   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 20.56 on 345 degrees of freedom  
## Multiple R-squared: 0.2284, Adjusted R-squared: 0.2038   
## F-statistic: 9.283 on 11 and 345 DF, p-value: 1.28e-14

When looking at the estimated coefficients for the 11 dummy variables we can see that they all have the expected sign expect for November. We were expecting a negative sign for November because the November price is expected to be below the December price since November is relatively closer to the October harvest date. Notice that it is only the June through September dummies which are statistically significant.

How can these results be interpreted? The lack of statistical significance from October to May corresponds to the period that the potatoes are in storage. This means that it is not possible to detect a price rise in the price of potatoes while they are in storage. It is only when storage is no longer feasible (June through September) that the price rises in a significant way. This likely suggests that the monthly cost of storage is very low because according to the intertemporal LOP prices must rise at a rate equal to the monthly cost of storage. It is not clear where the supply of potatoes come from during the June to September period. Perhaps they are imported and/or possibly there is a high cost supplier of potatoes in the warmer climates of California or Florida. In any case, the price rises sharply during this non-storage period.

In terms of a price gap, the highest price occurs in August where the average price is about 40 percent higher than the December price. The harvest price and the December price appear to be similar and so our best estimate of the price gap as a percentage of the average price is 40 percent. This is well below the 100 percent price gap which was estimated for India.

# Simulated Potato Prices in Australia

## Goal of this Section

The previous sections were devoted to an empirical examination of the intertemporal LOP. In this section we build a model of the intertemporal LOP, similar to how a model was build for the spatial LOP in our study of the North American lumber market and the global tomato market. The model will be calibrated to the Australian potato market. Australia was chosen because it operates as a closed market and therefore makes the modeling assumptions more realistic.

Australian potato production is assumed to take place in March of each year. Monthly demand and the monthly cost of storage are both assumed to remain fixed over time (i.e., no shifts in the demand schedule and no connection between the cost of storage and the amount to be stored). The potato stock pile must be allocated to monthly consumption in a way that ensures a price increase over time which is consistent with the intertemporal LOP. The analysis initially assumes that all potatoes must be consumed within on year. This assumption is later relaxed by assuming that potatoes can be stored from one marketing year to the next marketing year. This type of storage necessarily involves processed potatoes (e.g., frozen French fries of dehydrated potatoes) versus fresh potatoes.

## Background

Potatoes are an important horticultural crop in Australia. Due to concerns over disease transmission, imports of fresh potatoes are largely banned. Exports of fresh potatoes from Australia have averaged less than three percent of production over the past three years. The majority of potatoes grown in Australia are processed (e.g., French fries). Over the past five years, Australian exports and imports of processed potatoes have averaged to 6.6 percent and 4.8 percent respectively. These trade data come from the ComTrade database.

The above data suggests that the Australian potato market can well approximated as a closed market. In the analysis below we will assume that all of the potatoes produced in Australia are consumed domestically, either fresh or processed. We will not distinguish between fresh and processed potatoes, and we will further assumed that the monthly demand schedule is constant throughout the year.

Most Australian potatoes are grown in New Southwales, South Australia and Tazmania (see <https://silo.tips/download/the-potato-industry-in-new-south-wales>). Potatoes are harvested in many different months depending on the type of potato being produced (e.g., early versus mid versus late season) and the region. We will simplify by using the March harvest date for the midseason potatoes in the Riverina District as the representative harvesting time period.

## Data

According to FAOstat, the average level of Australian potato production for 2015 - 2019 was (in metric tonnes)

Q\_year <- 1160760

The average annual value over this same time period was (in Australian dollars)

Value <- 717000000

Divide annual value by annual production to obtain an estimate of the average price per tonne:

(P <- Value/Q\_year)

## [1] 617.6987

Divide annual production by 12 to obtain an estimate of monthly consumption (in tonnes):

(x <- Q\_year/12)

## [1] 96730

## Demand Curve

The inverse demand schedule can be expressed as . Invert this to obtain . The elasticity of demand evaluated at the calculated level of monthly consumption and average price is given by . Substitute in and then solve for to obtain . Finally, solve for and substitute in to obtain

Assume that

E <- -0.5

This gives

(b <- -1/E\*P/x)

## [1] 0.01277161

and

(a <- (1-1/E)\*P)

## [1] 1853.096

Let’s double check by substituting x into the inverse demand curve: we should see monthly consumption.

a - b\*x

## [1] 617.6987

## Intertemporal Law-of-One-Price (LOP)

To keep things simple, assume that the Australian potato producers can rent potato storage facilities at a cost of dollars per tonne per month. If the potatoes are stored instead of sold then then the opportunity cost of capital must also be considered. Assume that the foregone interest earnings are percent per month. This means that the total cost of storage per tonne for month is . Storage will only take place if the additional revenue from storing the potatoes, , is greater than or equal to the total cost of storage, .

As more and more potatoes are stored in the market, will be driven up and will be driven down. Arbitrage in the cash market for selling the commodity and in the market for storage will ensure that the in equilibrium which emerges, the marginal revenue from storing will equal to the marginal cost of storing. That is, .

If arbitrage in the cash and storage markets is costless then the intertemporal LOP is similar to the spatial LOP.

**Intertemporal LOP**

1. If Storage is positive then prices in successive periods must satisfy .
2. If there are zero stocks in storage then .

## Simulated Prices with Guess Value for P0

You may recognize the equilibrium pricing equation, , as a linear first order difference equation with constant coefficients. The solution to this equation (see <https://mjo.osborne.economics.utoronto.ca/index.php/tutorial/index/1/fod/t>) is given by:

If we knew the starting price, , then we would have the full pricing solution. For now we will assume

P0 <- 590

We can use the above equation along with to generate a price series for the 12 months “between” harvesting periods. To keep things simple we assume that all potatoes are harvested on March 1 and so the first month is March, the second month is April, etc. This means that is the March 1 price, is the April 1, price, etc.

We begin by assigning values to the remaining parameters of the model. Specifically, we must assign a value for the storage cost per tonne per month, , and the monthly cost of capital, . Assume that

m <- 1.5  
r <- 0.002

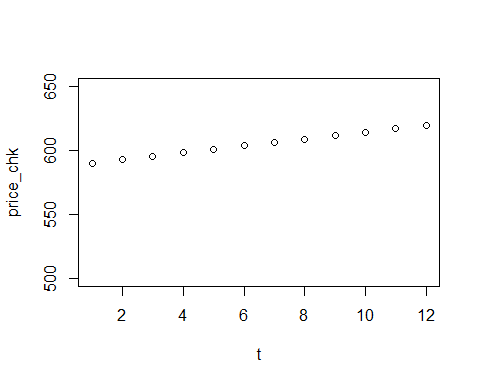
It will be tedious to write the above equations 12 times and so we will use an R loop. We create a price vector called, *price\_chk*, and then loop through the pricing formula, each time storing the calculated price as successive elements for the pricing vector:

price\_chk <- numeric(12)  
for(i in 1:12){  
 t <- i - 1  
 price\_chk[i]<- (1+r)^t/r\*(r\*P0+m)-m/r  
}  
price\_chk

## [1] 590.0000 592.6800 595.3654 598.0561 600.7522 603.4537 606.1606 608.8729  
## [9] 611.5907 614.3139 617.0425 619.7766

You should verify that the 12 simulated prices satisfy the LOP equation: . A graph of the 12 prices looks as follows:

t <- 1:12  
plot(t,price\_chk, ylim=c(500,650))



The previous plot looks linear but you can see from the equation that it is slightly non-linear.

## Solution Value for P0

In the previous section we guessed at the starting value for the price series, . To solve for the unknown value of begin by substituting the optimal pricing equation, , into the inverse demand schedule, , and then solve for . This gives us an expression for the optimal monthly consumption as a function of our unknown value for :

Let denote the size of the beginning stockpile, which consists of stocks brought in, plus the March 1 harvest, . Let denote the level of stocks which are carried over to the next year. That is, is the same as if the problem was to be solved again, one year later. Assume initially that and . In this case initial stocks, , will equal the size of the harvest:

S\_in <- 0  
S\_out <- 0  
H <- Q\_year  
(S\_0 <- S\_in + H)

## [1] 1160760

Market clearing requires where for the current problem since there are 12 months. In words, the sum of monthly consumption must explain the change in stocks from the beginning of the crop year to the end of the crop year.

After substituting in the previous expression for this market clearing condition can be rewritten as

where . Solve this equation for to get

The variable is a standard finite sum of a geometric sequence (see <https://mathworld.wolfram.com/GeometricSeries.html>). It can be shown that

To derive the solution value for note that

N <- 12  
(Z <- (1+r)/r\*((1+r)^N-1))

## [1] 12.15715

Given our previous assumption that it follows that

(P0\_star <- (a\*r+m)\*N/(Z\*r) - m/r - b/Z\*(S\_0-S\_out))

## [1] 600.0192

## Graph of Price and Stocks over Time

Given the solution value for in the previous equation, we now know how prices will evolve over the 12 month period in accordance to the intertemporal LOP. It would also be useful to observe how the potato stock pile gradually depletes over the 12 months. After harvest is complete, stocks evolve according to . Recall that . Thus, . We can now use an R loop to observe price and stocks over the 12 month period.

t2 <- 1:12  
stocks <- numeric(12)  
price <- numeric(12)  
price[1] <- (1+r)^1/r\*(r\*P0\_star+m)-m/r  
stocks[1] <- S\_0 - a/b +1/b\*price[1]  
for(t in 2:12){  
price[t]<- (1+r)^t/r\*(r\*P0\_star+m)-m/r  
stocks[t]<- stocks[t-1]-a/b +1/b\*price[t]  
}  
price

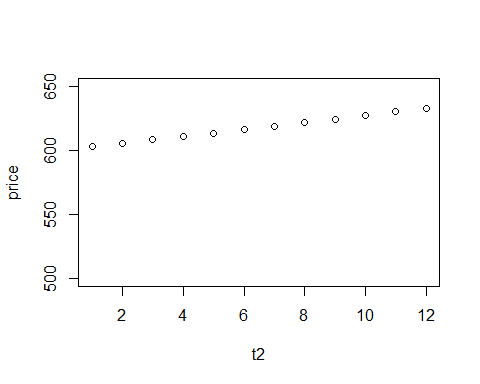
## [1] 602.7192 605.4246 608.1355 610.8517 613.5735 616.3006 619.0332 621.7713  
## [9] 624.5148 627.2638 630.0184 632.7784

stocks

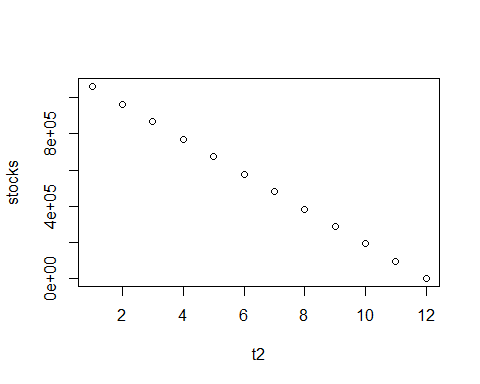
## [1] 1.062857e+06 9.651661e+05 8.676873e+05 7.704212e+05 6.733682e+05  
## [6] 5.765287e+05 4.799032e+05 3.834921e+05 2.872957e+05 1.913147e+05  
## [11] 9.554928e+04 -9.604264e-10

Graph the price and stocks data points.

plot(t2,price, ylim=c(500,650))



plot(t2, stocks)



## Two Identical Successive Years

Suppose harvest at the beginning of a second year was identical to the first year harvest. Moreover, there were no stocks carried into or out of the second year. This means that the set of prices in year 2 will be identical to the set of prices in year 1.

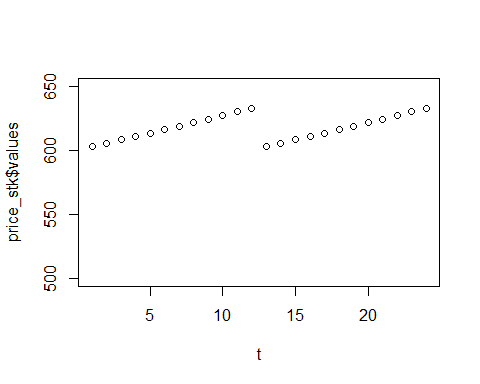
To program this in R we create a matrix with the first set of prices in the first column and the second set of prices in the second column. We then convert to a dataframe and use the Stack function to create a single set of values

price\_2 <- cbind(price, price)  
price\_2 <-as.data.frame(price\_2)  
price\_stk <- stack(price\_2)  
price\_stk

## values ind  
## 1 602.7192 price  
## 2 605.4246 price  
## 3 608.1355 price  
## 4 610.8517 price  
## 5 613.5735 price  
## 6 616.3006 price  
## 7 619.0332 price  
## 8 621.7713 price  
## 9 624.5148 price  
## 10 627.2638 price  
## 11 630.0184 price  
## 12 632.7784 price  
## 13 602.7192 price.1  
## 14 605.4246 price.1  
## 15 608.1355 price.1  
## 16 610.8517 price.1  
## 17 613.5735 price.1  
## 18 616.3006 price.1  
## 19 619.0332 price.1  
## 20 621.7713 price.1  
## 21 624.5148 price.1  
## 22 627.2638 price.1  
## 23 630.0184 price.1  
## 24 632.7784 price.1

Graph the new (double) set of prices

t <- 1:24  
plot(t,price\_stk$values, ylim=c(500,650))



## Stockout and the Intertemporal LOP

It should be obvious that potato merchants have no incentive to store from year 1 to year 2 because doing so will result in a lower selling price and positive storage costs. When zero storage from one year to the next is optimal we call this a market *stockout*.

It should be obvious that if beginning stocks for year 1 are less than the year 2 harvest then the market will stockout. In this case prices in year 1 will uniformly be above prices in year 2, in which case merchants have an even stronger negative incentive to carry stocks from year 1 to year 2.

When the market stocks out, prices jump down with the arrival of the new harvest, and then gradually rise back up. If there are repeated stockouts (e.g., 5 - 6 years) then the pricing pattern looks like a “saw tooth” (see <https://en.wikipedia.org/wiki/Sawtooth_wave>).

It is important to keep in mind that when the market stocks out the intertemporal LOP continues to hold when transitioning from year 1 to year 2. Recall the second component of the LOP, which states: If there are zero stocks in storage then . This equation allows for the outcome where the price one day after the new harvest is lower than the price one day before the new harvest.

## Positive Carry Over

The more interesting case is when year 1 beginning stocks are large relative to the year 2 harvest. In this case the month 12 price of year 1 might be below the month 1 price of year 2. If the price difference is large enough then merchants will have an incentive to carry over potatoes from year 1 to year 2.

For example, suppose beginning potato stocks for year 1 equal 1.2 million tonnes. This value is about 3.4 percent higher than the year 2 harvest, which is assumed to equal the long term average of 1,160,760 tonnes. If the pricing model specified above is re-run with and not carryout stocks (i.e., ) the following set of (approximate) equilibrium monthly prices emerge.

559 561 564 566 569 572 574 577 580 582 585 588

In year 2, the 12 monthly (approximate) prices are: 603 605 608 611 614 616 619 622 625 627 630 633

It should be clear that the year 1 month 12 price of $588/tonne is well below the year 2 month 1 price of $603/tonne. The intertemporal LOP no longer holds. Some stock from year 1 should be carried over to year 2.

As stock is shifted from year 1 to year 2, the year 1, month 12 price will increase and the year 2, month 1 price will decrease. Stocks will continue to shift until the intertemporal LOP holds when transitioning from year 1 to year 2.

## Equilibrium Level of Carry Over Stocks

What remains to be determined is the actual amount which is carried over from year 1 to year 2. Let denote the month 12 price in year 1 and let denote the month 1 price in year 2. If carry over from year 1 to year 2 is positive then according to the LOP equation, we must have = (1+r)P\_{12}^1 + m$.

We can substitute the previous equations into this formula and solve for , which is the amount of stock carried over from year 1 to year 2. The steps involved are somewhat complex, and so we will jump to the final solution. Assuming that are beginning stocks for year 1 and is the level of harvest for year 2, we have:

Suppose is equal to 1,160,760 tonnes, which is the 2015 - 2019 average level of potato production in Australia. Further suppose that is equal to 1.2 million tonnes. We can use the previous formula to identify that amount of potatoes that will be carried over from year 1 to year 2 (presumably in a processed form). The formula will be build using three separate components so that it is easier to follow.

H2 <- Q\_year  
S0 <- 1200000  
S\_out\_A <- (a\*r+m)\*N/(b\*r)\*(1-(1+r)^12)  
S\_out\_B <- (1+r)^12\*S0 - H2  
S\_out\_C <- (1+r)^12 + 1  
(S\_out <- (S\_out\_A+S\_out\_B)/S\_out\_C)

## [1] 4450.514

You might expect the above formula to return a value of zero if year 1 beginning stocks are low relative to the size of the year 2 harvest. What you will see instead is a negative value for . More sophisticated programming is required if want the restriction that stocks can’t be negative to be imposed on the model. You must keep an eye on the value of . If it is negative then zero carry over stocks are optimal and the market should be viewed as stocked out.

## Price Path with Positive Carry Over Stocks

Lets continue with the previous example by solving the model for 24 months of potato prices. We will first solve the 12 month model from the perspective of year 1 with beginning stocks equal to 1.2 million tonnes and tonnes of potatoes carried over from year 1 to year 2. We will then solve the model for year 2, assuming a normal harvest and tonnes of potatoes added to the year 2 stockpile.

For year 1 we have:

(P0\_star\_yr1 <- (a\*r+m)\*N/(Z\*r) - m/r - b/Z\*(S0-S\_out))

## [1] 563.4713

stocks\_yr1 <- numeric(12)  
price\_yr1 <- numeric(12)  
price\_yr1[1] <- (1+r)^1/r\*(r\*P0\_star\_yr1+m)-m/r  
stocks\_yr1[1] <- S0 - a/b +1/b\*price\_yr1[1]  
for(t in 2:12){  
price\_yr1[t]<- (1+r)^t/r\*(r\*P0\_star\_yr1+m)-m/r  
stocks\_yr1[t]<- stocks\_yr1[t-1]-a/b +1/b\*price\_yr1[t]  
}  
price\_yr1

## [1] 566.0982 568.7304 571.3679 574.0106 576.6587 579.3120 581.9706 584.6345  
## [9] 587.3038 589.9784 592.6584 595.3437

stocks\_yr1

## [1] 1099229.749 998665.595 898307.951 798157.230 698213.845 598478.212  
## [7] 498950.745 399631.861 300521.978 201621.513 102930.885 4450.514

And for year 2 we have:

(P0\_star\_yr2 <- (a\*r+m)\*N/(Z\*r) - m/r - b/Z\*(S\_0+S\_out))

## [1] 595.3437

stocks\_yr2 <- numeric(12)  
price\_yr2 <- numeric(12)  
price\_yr2[1] <- (1+r)^1/r\*(r\*P0\_star\_yr2+m)-m/r  
stocks\_yr2[1] <- S\_0 +S\_out - a/b +1/b\*price\_yr2[1]  
for(t in 2:12){  
price\_yr2[t]<- (1+r)^t/r\*(r\*P0\_star\_yr2+m)-m/r  
stocks\_yr2[t]<- stocks\_yr2[t-1]-a/b +1/b\*price\_yr2[t]  
}  
price\_yr2

## [1] 598.0344 600.7305 603.4319 606.1388 608.8511 611.5688 614.2919 617.0205  
## [9] 619.7545 622.4940 625.2390 627.9895

stocks\_yr2

## [1] 1.066941e+06 9.688822e+05 8.710352e+05 7.734000e+05 6.759773e+05  
## [6] 5.787673e+05 4.817705e+05 3.849874e+05 2.884184e+05 1.920639e+05  
## [11] 9.592425e+04 -7.057679e-10

The most important result from this last simulation is that the price continues to rise over all 24 months, despite the fact that new harvest arrives in month 13. It seems counterintuitive that the price rises when new supply becomes available. However, if you think about this for a while there can be no other outcome. A merchant is willing to carry inventory from year 1 to year 2 only if price increase is sufficient to satisfy her storage costs. The LOP tells us that as long as storage is positive, the price must rise according to .