# The Three Epochs of Oil[1]

Eyal Dvir (Boston College) and Kenneth S. Rogoff (Harvard)

August 16, 2010

[1]We thank Jim Anderson, Deepa Dhume, Gita Gopinath, Lutz Kilian, Brent Neiman, and participants at the BC macro seminar, the Oil and the Macroeconomy conference at the Boston Fed, and the NBER for helpful comments on earlier versions, as well as Susanto Basu and Fabio Ghironi for helpful conversations. We also thank Vania Stavrakeva for help with computer code.
Dvir: eyal.dvir@bc.edu, Dept. of Economics, Boston College, 140 Commonwealth Ave., Chestnut Hill, MA 02467. Rogoff: krogoff@harvard.edu, Dept. of Economics, Harvard University, Littauer Center, 1805 Cambridge St., Cambridge, MA 02138

**Abstract**

We find strong evidence for changes in real oil price persistence and volatility across three well defined periods since 1861. We argue that historically, the real price of oil has tended to be highly persistent and volatile whenever rapid industrialization has coincided with uncertainty regarding access to supply. We extend the commodity storage model to incorporate both transitory and permanent demand shocks. When demand is subject to persistent growth shocks and supply is restricted, the role of storage is shown to be speculative, instead of its classic mitigating role. This is consistent with price persistence and volatility co-moving, as observed.

Keywords: Oil Price, Oil Shocks, Storage, Structural Change.

JEL classification: E0, Q4, L7, N5.

# 1 Introduction

Much has been written on the oil shocks of the 1970's as watershed events that have transformed the energy market; that, together with the lack of high frequency data from earlier periods, have led to an almost complete concentration on the post oil shock period among economists[1]. However, much can be learned about oil price behavior from the less recent past. The crude oil price time series illustrated in Figure 1 goes back to 1861[2]; even a cursory look reveals stark differences in the behavior of the series at different periods. First, from 1861 until about 1878, there was a period of extremely high volatility and generally high prices. Then came a much less volatile period, approximately between 1878-1973, in which prices were also generally lower. Finally, from 1973 onwards, we see a second period of high volatility accompanied again by higher prices.

Our first task in this paper is to document these differences and formally test for changes in behavior. We run two such tests, for changes in persistence and for changes in volatility. We find striking empirical similarities between the periods 1861-1878 and 1973-2009, in that oil prices were both significantly more persistent and significantly more volatile in these periods, both relative to the long period that separates them, i.e. 1878-1973. We also estimate that a further break in oil price volatility, but not in persistence, occurred around 1933, so that oil price in the period 1878-1933, while much less volatile than in 1861-1878, was still significantly more volatile than in the period 1933-1973[3].

What can explain the concurrence of price persistence and price volatility? We offer an informal, historical narrative, as well as a formal model. Our approach in this paper is to look for a unifying framework which is flexible enough to allow for the very different price behavior across periods that we observe. We find striking historical similarities between the two end-periods mentioned, 1861-1878 and 1973-2009, in terms of supply and demand factors affecting the market for oil. On the demand side, as we explain in greater detail in Section 5, both periods were years of intense industrialization in what was then becoming a major engine of the global economy: the U.S. in 1861-1878, and East Asia in 1973-2009. We see these as periods in which the demand side was characterized by persistent growth shocks. On the supply side, meanwhile, both periods featured uncertainty regarding the

[1] Pindyck (1999) is a notable exception.

[2] The series used in this paper is taken from British Petroleum's "Statistical Review of World Energy", revised annually and available at www.bp.com/statisticalreview. Prices are in 2009 $US per barrel. The series is comprised of three consecutive price series: US average price in 1861-1944, Arabian Light in 1945-1983, and Brent dated in 1984-2009.

[3] All of our tests reject the null of no break with very high levels of confidence. However, the confidence intervals around the exact break dates are large enough to suggest caution regarding the interpretation of individual historical events. Our emphasis will therefore be on the broad characteristics of the periods in question, rather than on the exact date of change from one period to the next.

continued access of consumer markets to oil. This was due to the monopoly of railroads on transportation in the former period, and to OPEC's ability to restrict access to easily exploitable reserves in the latter period (see Section 5 for details). The change-points we find in the data correspond to major changes in the oil industry's structure: 1878-9 saw the opening of the first long-distance pipeline, ending the railroads' monopoly over transport; 1933-4 saw the vast discoveries of East Texas oil bring government control to the industry; and 1972-3 saw the peak of U.S. oil production and the resulting rise to prominence of OPEC.

We further argue in this paper that this confluence of supply and demand factors can explain why we observe large and concurrent changes in both oil price persistence and oil price volatility over the years. We present a model, an extension of the canonical commodity storage model à la Deaton and Laroque (1992, 1996), in which we introduce growth dynamics to their well-known framework. This results in a model that can accommodate both $I(0)$ and $I(1)$ stochastic processes, so that periods of stable and stochastic trends can both be considered. The model can explain our main empirical findings: it predicts that in the presence of uncertainty regarding the trend, i.e. persistent growth shocks, rational storage behavior may act to *enhance volatility*. In the standard commodity storage framework, where uncertainty is in regard to deviations from trend only, storage always acts to *reduce volatility*. This feature of the storage model is in itself novel, as well as potentially useful in accounting, qualitatively, for observed patterns in the data. We present simulations in which this behavior can increase price volatility following persistent growth shocks when supply is restricted.

The large literature on commodity price behavior falls broadly into two major strands, depending on whether the commodity in question is perceived to be renewable. On the one hand, models of storage have been used mostly to study renewable commodities such as corn and wheat; see Wright (2001) and references therein for a survey of theory and evidence. Notable exceptions are papers which in the context of oil extend the storage model in various ways: Alquist and Kilian (2010), Ng and Ruge-Murcia (2000), and Routlege et al. (2000). These papers however do not seek to explain the different behavior of oil prices across historical periods as we do here, nor do they incorporate growth shocks explicitly into the framework. Apart from these exceptions, the study of non-renewable commodities, as oil is traditionally classified, has followed an altogether different path, strongly influenced by the seminal contribution of Hotelling (1931). Krautkraemer (1998) surveys the theory and evidence, which shows quite clearly that finite availability of oil - a separate issue from that of free access to currently available supplies - is not of first order significance in explaining oil price behavior. As a particularly striking example, note that proven world oil reserves have been *increasing* in recent decades, in spite of ever increasing production[4]. As a result, it may well be that technological advances in oil exploration and utilization will be enough to satisfy demand in the foreseeable future. That is the assumption that we make in this paper.

Our work is also related to the ongoing debate on oil and the macroeconomy (see the two recent surveys by Hamilton, 2008 and Kilian, 2008a, and references therein. Hamilton (2009a) provides an account of oil's contribution to the 2007-8 recession). This literature focuses on correctly identifying the sources of shocks to oil prices , usually utilizing post-1973, higher-requency data. We argue in this paper that a long-term view of the oil market should serve to complement and enrich this debate, since shocks to the oil market may have remarkably different effects on the real price of oil across historical periods, not only due to their origin on the supply or the demand side, but also because of the ability (or lack thereof) of key players in the market to restrict access to supplies. In particular, in periods when the ability to restrict access to supplies was lacking, the oil market showed remarkable flexibility and relative price stability, even in the face of massive disturbances in both supply and demand.

The paper proceeds as follows: section 2 presents our empirical findings on oil price behavior over time. Section 3 introduces the model, and Section 4 examines the model behavior under both transitory and permanent shocks. Section 5 puts our empirical findings in the context of the history of supply and demand for oil. Section 6 concludes.

## 2 Behavior of the Real Oil Price: Then and Now

Table 1 presents some sample statistics regarding each of the three periods delineated in the introduction, as well as the entire series. The differences between the sub-samples are quite clear: the mean price between the years 1861-1877, at \$50.9 (measured in 2007 U.S. dollars), was almost triple the mean price between 1878-1972, at \$17.2 . In contrast, there is only a 13% difference (again in 2007 U.S. dollars) between the mean price between the first period, 1861-1877, and the third, 1973-2009, at \$44.7 . A similar pattern holds for differences in the unconditional standard deviation of annual prices across these periods: at \$25.3, the standard deviation of price in the period 1861-1877 is five times higher than that of the period 1878-1972 (\$5.1), but only 12% higher than the standard deviation of price in the years 1973-2009 (\$22.2). The next lines of the table compare the rates of change across periods. Note in particular that the unconditional standard deviation of absolute price

[4]See the BP Statistical Review, published annually at www.bp.com/statisticalreview, for proven reserves and production data from 1980.

change (a common measure of volatility) behaves in an analogous manner: it is quite similar in the first and third periods (27.5% between 1861-1877 compared with 22.8% between 1973-2009), but remarkably lower in the intervening period. In sum, Table 1 shows that there seems to be much in common in terms of the behavior of real oil prices between the periods 1861-1877 and 1973-2009, while both periods look markedly different in most respects from the intervening period of 1878-1972.

## 2.1 Testing for Change in Persistence

Studies of the time series properties of real oil prices have taken one of the following approaches in the face of these clear non-linearities in the series: either analyzing the series as a whole, or, much more commonly, treating the series as composed of separate series "pasted together", and proceeding to analyze them in isolation. In one important category, that of determining whether or not oil prices exhibit a unit root, these different approaches have led to opposite conclusions. Pindyck (1999), an example of the former approach, ignores the aforementioned differences, and judges the entire series to be mean-reverting to a moving quadratic trend. A recent example of the latter approach, Hamilton (2009b) focuses his analysis on the third period only (with quarterly data). His conclusion that the real price of oil follows a random walk with no drift applies therefore by construction only to the period since the early 1970s[5].

In what follows we will treat both the assumption of a pure $I(0)$ process and the assumption of a pure $I(1)$ process as our null hypotheses, and test whether the series exhibits a shift from $I(0)$ to $I(1)$ (or vice versa) against *both* of these assumptions. In other words, instead of trying to decide whether the series as a whole or in part exhibits a unit root, we aim to determine whether it shows clear transitions from a stochastic trend to a deterministic one, and vice versa. In order to do that, we employ a series of tests proposed by Harvey, Leybourne, and Taylor (2006, HLT henceforth). This series of tests is a modified version of tests for change in persistence proposed earlier by Kim (2000), Kim et al. (2002), and Busetti and Taylor (2004), all of whom build on the unit root testing method of Kwiatkowski *et al.* (1992). What makes the HLT test of change in persistence appealing in our context is that it maintains its properties of consistency and appropriate size both under the $I(0)$ null and under the $I(1)$ null. This allows us to test for structural change without taking an a-priori stand regarding the null hypothesis. A disadvantage of the HLT modified test is that it does not take account of possible heteroskedasticity of the stochastic process. This is important, since oil price volatility does in fact exhibit structural breaks, as we shall see later in this

[5] Studies of higher frequency data from the 1990s (daily, weekly) do find a mean-reverting factor as well as a permanent factor. See for example Routledge et al. (2000), Schwartz and Smith (2000).

section. However, applying a bootstrap method to the tests can correct this issue: following Cavaliere and Taylor (2008), we use a wild bootstrap to arrive at heteroskedasticity - robust bootstrap p-values for our test statistics[6].

Table 2 presents the results of the HLT change of persistence test, using the real oil price series. Since the test is designed to find a single change-point, whereas the series exhibits two obvious break candidates, we conducted the test separately for periods 1 and 2, and for periods 2 and 3. The exact end points shown in the table (1881 and 1965) were chosen arbitrarily; the qualitative results are robust to small changes in these end points[7]. The table reports three sets of test statistics, all functions of Kim's (2000) ratio-based test statistic $K$. The null hypothesis in each test is that there is no change in persistence present; the alternative is that a single change in persistence occurs. The testing procedure first computes the value of $K$ for each potential break-point. As the true break-point is unknown, the procedure then computes the mean value ($MS$), the mean-exponential value ($ME$), and the maximum value ($MX$) taken over all these values of the test statistic. To test the alternative hypothesis of a change in persistence from $I(1)$ to $I(0)$, the same procedure it performed, with the statistics $MS^R$, $ME^R$, and $MX^R$ computed in the same way, taken over the reciprocal values of $K$. A third set of statistics - $MS^{MAX}$, $ME^{MAX}$, and $MX^{MAX}$ - test the alternative hypothesis that a change in persistence has occurred, regardless of direction, against either null. As is common in this literature, we test for a change in persistence from a trend - stationary process to a difference - stationary process (and vice - versa). It should be noted that by construction, the tests can be performed regardless of the type or order of time-dependence in the stationary process, as long as some regularity conditions hold (see Kwiatkowsky *et al.*, 1992 for original proofs). Estimates for the break-points follow Kim (2000). In parentheses are the bootstrap p-values[8].

Testing for change in persistence in the years 1861-1965, then, we find very strong evidence for a significant change from a high-persistence, local to $I(1)$ process, to a low-persistence $I(0)$ process, where the point of change is estimated at 1877. This is shown by

[6] The wild bootstrap was introduced by Hansen (2000). Rather than re-sampling, it generates a bootstrap sample by multiplying the residuals from regressing the series on the exogenous variables, with random numbers from a normal distribution. In this way the pattern of volatility present in the original shocks can be replicated in the bootstrap sample. Appendix A provides a detailed walk-through of the exact testing procedure we use.

[7] As a ratio-based test, the HLT change in persistence test is designed to reject the null if a point can be found after which the behavior of the series is statistically different than its behavior before that point. When two such points exist, and moreover the behavior of the series before the first point is similar to behavior after the second point, the test will lose power.

[8] Note that for our test to reject either null, price persistence on one side of a tentative break should be statistically different from price persistence on the other side of the break, but need not meet any particular critical value. Because of this, Caner and Kilian's (2001) criticism of KPSS tests, namely, that they may suffer from size distortions and therefore produce spurious rejections, does not apply to our paper.

the very high values of the relevant test statistics, $MS^R$, $ME^R$, and $MX^R$, which are all significant at the 1% level. Note the the statistics testing for any change, regardless of direction ($MS^{MAX}$, $ME^{MAX}$, and $MX^{MAX}$) are just slightly smaller, and still highly significant, implying that our rejection of both null hypotheses is quite clear. The statistics testing for a change from a low-persistence $I(0)$ to a local to $I(1)$ high-persistence process ($MS$, $ME$, and $MX$) are all very small and insignificant. In the period 1881-2009, we find strong evidence for a change in persistence from a low-persistence $I(0)$ process to a high-persistence, local to $I(1)$ process, with the point of change estimated at 1973. All three of the relevant test statistics, $MS$, $ME$, and $MX$, point to the same conclusion, and all are significant at the 5% level. Again we see that the statistics testing for any change are very similar in magnitude and significance, implying a clear rejection of both nulls. The statistics testing for change in the opposite direction are small and insignificant.

We use simulation results from Kim (2000) to calculate the 95% confidence interval around our change-points. As can be seen from Table 2, these are estimated somewhat imprecisely, with confidence intervals of 8 and 10 years for the first and second change-points respectively[9]. However the rejection of both the null of a pure $I(0)$ process and the null of a pure $I(1)$ process is quite clear. The series of real oil prices should not be seen as high- or low-persistence, but rather as having switched from one regime to another.

## 2.2 Testing for Change in Volatility

Apart from the rate of persistence, another time series aspect of real oil prices that has attracted much attention in the literature is their volatility. As already mentioned, volatility (as measured by the standard deviation of absolute rates of growth) was high before 1878, low from around that time until the early 1970s, then high again until the end of our sample in 2009. We therefore conducted a test for multiple breaks in oil price volatility, using the methods of Bai and Perron (1998, 2003). We define volatility as the mean absolute residual from a regression of oil price growth on its lagged value. The null hypothesis is that volatility is constant throughout. The test results are shown in Table 3, and illustrated in Figure 2. We identify three potential breakpoints: 1878, 1933, and 1972. All three test statistics against the null of no break are highly significant, implying that the series contains at least one breakpoint. However, in deciding how many breakpoints there are, the various criteria explored by Bai and Perron do not agree: their sequential procedure selects only

[9] Kim (2000) finds in simulations that for change-points that occur at the 25th or the 75th percentile of a given series, his procedure for estimating the change-point location has a maximum standard deviation of 1.6893 for T=100. We accordingly use this value, scaled to our sample length, to calculate our 95% CIs.

one breakpoint, in 1878, whereas the Bayesian Information Criterion selects all three[10]. A look at the coefficients denoting mean volatility in the different periods can explain this discrepancy. We see that the coefficients of Periods II (1879-1933) and IV (1974-2009) are very similar, and both are quite different from the coefficient for Period III (1934-1973). As Bai and Perron recognize (2003, pp. 15-16), in these cases the sequential procedure can be improved upon: the number of breaks should be chosen according to the last significant test statistic, instead of the usual practice of choosing according to the first insignificant test statistic. In the current case, as seen in Table 3, this improved sequential procedure puts the number of breaks at three, similarly to the BIC. Oil price volatility then has gone down by about half sometime in the last quarter of the nineteenth century (with 1878 as our best estimate), then gone down again by about two thirds around 1933. When it increased again, according to our estimate in 1972, it regained its level of the early twentieth century, but did not reach the heights set by oil prices before 1878. Note that 95% confidence intervals for the change-points are quite large; as in the change in persistence test, our confidence in the occurrence of breaks in the series' behavior is far stronger than our confidence in the exact dates of these breaks. Nevertheless, these years will be useful as anchors in our historical narrative in Section 5.

We can sum up our empirical findings as follows: real oil price from 1861-1877 (or 1878) was highly persistent and volatile, from 1878-1933 was not as persistent and less volatile, from 1934-1972 (or 1973) it was still not very persistent and displayed even lower volatility. Finally, from 1973 on the real price of oil returned to being highly persistent and volatile, though not as volatile as in the pre-1878 period. In Section 3 we present a model of the oil market that ties together these patterns of price behavior in a single framework.

# 3 An Extended Commodity Storage Model

Our model is an extension of the classic commodity storage framework. Chambers and Bailey (1996) and Deaton and Laroque (1996) extend the model to allow for autoregressive shocks. We extend it further to explicitly incorporate demand, and to allow for growth shocks.

## 3.1 Availability and Storage

Time is discrete, indexed by $t$. The market for oil consists of consumers, producers, and risk neutral arbitrageurs. The latter have at their disposal a costly storage technology which

[10]The LWZ information criterion also chooses 1878 as the only break; however, this criterion is known to perform badly when breaks are present (i.e. the alternative is true).

may be used to transfer any positive amount of oil from period $t-1$ to period $t$. Storage technology is limited by a non-negativity constraint, i.e. the amount stored at any period cannot drop below zero. This implies that intertemporal arbitrage, although potentially profitable, cannot always be achieved. In these cases the market is "stocked out". Let $A_t$ denote *oil availability*, as the amount of oil that can potentially be consumed at time $t$, since it has already been extracted from the ground, either in period $t$ or at some point in the past, and has not been consumed before period $t$. It is given by

$$A_t = X_{t-1} + Z_t, \tag{1}$$

where $X_{t-1}$ denotes the stock of oil transferred from period $t-1$ to $t$, and $Z_t$ denotes the amount of oil that is produced at time $t$. For simplicity, we assume that no oil is lost due to storage[11]. Decisions concerning both variables - how much to store, how much to produce - are assumed to have been made before period $t$ began. In period $t$ agents decide how to divide $A_t$ between current consumption $Q_t$ and future consumption, so that demand - the sum of current consumption and the amount stored for the future - must always equal current availability:

$$A_t = Q_t + X_t. \tag{2}$$

## 3.2 Demand for Oil

Let $Y_t$ denote an income parameter, which can be thought of as some function of world GDP. We can then write an inverse demand function for oil as follows:

$$P_t = P(Q_t, Y_t), \tag{3}$$

which is decreasing in its first argument, and increasing in its second. This inverse demand function constitutes a departure from the canonical model, where demand for the commodity is a function of its price alone, effectively assuming no income effects. This departure is a natural one to make, however, in the context of oil, as oil consumption and income are very highly correlated (see references in Section 5). We posit an inverse demand function in which only the *ratio* of consumption to income matters, i.e. invere deamnd is homogeneous of degree zero:

$$P_t = P(Q_t, Y_t) = P(\frac{Q_t}{Y_t}, 1) = p(q_t), \tag{4}$$

[11] Alternatively, we could have specified storage costs by a given loss percentage, as in Deaton and Laroque (1996).

where lowercase letters denote variables normalized by $Y_t$. We think of the normalized variables as "effective" amounts, in the sense that a growing income leads to higher energy needs, spreading any given amount of oil more thinly[12].

We will use a CES inverse demand function:

$$P_t = q_t^{-\gamma} = (a_t - x_t)^{-\gamma}, \tag{5}$$

where $\gamma > 1$ is the inverse elasticity of demand, and $a_t$, $x_t$ denote effective availability and storage in period $t$, respectively. It is natural to assume that the effective demand for oil is inelastic with respect to price. As equation (5) makes clear, for a given supply of oil, price is a function of the competing demands of current and future consumption. If the desire to consume more in the future grows (driven by expectations of future conditions), more oil is stored rather than consumed today, resulting in a price rise today even though supply has not changed.

Let $\overline{Y}_t$ denote trend income, i.e. the level of income that would prevail at time $t$ in a world without income shocks. $\overline{Y}_t$, which we think of as a measure of current production technology, is assumed to increase over time at a constant rate $\overline{\mu} > 0$. We now consider two alternative stochastic processes for $Y_t$: one where income moves around a deterministic trend, and another where the trend itself is stochastic. The former is a simple AR(1) process, analogous to the stochastic process that Deaton and Laroque (1996) consider for supply. Under this assumption we have:

$$\frac{Y_{t+1}}{\overline{Y}_{t+1}} = \left(\frac{Y_t}{\overline{Y}_t}\right)^{\rho} e^{\varepsilon_{t+1}}, \tag{6}$$

where $\rho \in (0,1)$ and $\varepsilon_{t+1} \sim N(0, \sigma_{\varepsilon}^2)$ is an iid shock. We think of this case as more closely relevant to income shocks in developed economies, where the economy exhibits business cycles around a stable trend. In the latter case, we assume instead:

$$Y_{t+1} = e^{\mu_{t+1}} Y_t, \tag{7}$$

such that

$$\mu_{t+1} = (1-\phi)\overline{\mu} + \phi\mu_t + \upsilon_{t+1}, \tag{8}$$

where $\phi \in (0,1)$ and $\upsilon_{t+1} \sim N(0, \sigma_{\upsilon}^2)$ is an iid shock. Dividing both sides of (7) by $\overline{Y}_{t+1}$ we

[12] A disadvantage of using normalized quantities is the difficulty in directly calibrating the model to actual observable quantities. This would be an important issue if we had the ability to perform such calibration. Unfortunately, quantity data for the oil industry - production, consumption, stocks - are not readily available for the full period (1861-2008). Our model is highly stylized as a result.

get:

$$\frac{Y_{t+1}}{\overline{Y}_{t+1}} = e^{\mu_{t+1}-\overline{\mu}} \frac{Y_t}{\overline{Y}_t}. \tag{9}$$

We think of this case as more relevant to income shocks in some developing countries, in particular quickly industrializing economies where very high growth rates can be persistent.

In principle the world demand for oil should be affected by developments in both types of economies, developed and developing, depending on the relative intensity of oil use. Specifically, a positive income shock in an advanced economy such as the twentieth century U.S., where the trend is stable, is expected to disappear over time. In contrast, a positive income shock in an emerging economy such as the late twentieth century China is perceived to have a more lasting effect on global demand, especially given its relatively high oil intensity.

## 3.3 Supply of Oil

In the canonical commodity storage model, supply $Z_t$ varies according to some stochastic process $\psi_t$ around a predetermined mean $\widetilde{Z}_t$, and it is this variability in supply that creates an incentive for inter-temporal smoothing by the large pool of risk neutral arbitrageurs. As the literature has long recognized, demand and supply shocks in the canonical model are isomorphic: one can think of a negative realization of $\psi_t$ as representing an especially cold winter (demand) or a breakdown in a major pipe (supply). For this reason, since we model demand shocks explicitly, it would be redundant to model supply shocks separately. Our choice has to do, of course, with the argument of Section 5, but it is important to recognize that theoretically speaking we could just as well have modeled supply shocks.

We do model supply *choices*, however. In particular, we assume that either of the following two regimes holds: a regime where oil supply does not react at all to income shocks due to capacity constraints (such as railroad infrastructure or number of operational wells), and a regime in which oil supply fully accommodates any shock to demand (for example, when potential production is much higher than current production). We think of the former regime as describing supply behavior when access to excess supply sources is restricted, so that suppliers are constrained to produce at their installed capacity[13]. Under the latter regime, suppliers seek to stabilize prices by varying quantities as needed. We think of this regime as representing either perfectly competitive supply, where producers will offer any amount at a given price, or else the effect of purposeful government intervention, seeking to control market prices by adjusting supply.

[13]Naturally, capacity constraints can be relaxed in the medium run. However, as long as capacity does not fully accommodate all demand shocks, dynamic behavior will be qualitatively similar to the case where it does not react at all. A similar point has been made by Williams and Wright (1991).

Formally, in the former regime we assume that supply grows at the trend income rate $\overline{\mu}$, so that

$$Z_{t+1} = \widetilde{Z}\overline{Y}_t, \tag{10}$$

where $\widetilde{Z}$ is a supply parameter. Next period's oil supply depends then on current technology, since overall technological progress, which drives global GDP growth, applies to the oil extraction and exploration sectors as well, and therefore determines overall capacity.

This assumption deserves some comment. Of course, the total amount of oil existing in the earth's crust is finite. However technological progress is key to exploiting an increasing fraction of it over time. The global ratio of oil production to known oil reserves is slightly less than 2.5% , and has been quite steady at that level since 1985 (BP Statistical Review), even though global production has increased by about 39% from 1985 to 2009. The world economy is no closer to running out of oil now than it was in 1985 due to the rate at which new reserves are discovered and known reserves become exploitable due to better technology. This is the context which drives our modeling choice.

Note that in this regime oil supply depends on the technology driving income growth, but not on income growth itself. Therefore shocks to demand will drive a wedge between supply and demand, causing a shift in equilibrium price. In contrast, under the alternative supply regime oil suppliers will accommodate all income shocks, i.e. oil supply will be perfectly elastic. Next period's supply then will also depend on current income level (and growth rate if appropriate). Supply is then given by:

$$Z_{t+1} = \widetilde{Z}\overline{Y}_t \left(\frac{Y_t}{\overline{Y}_t}\right)^{\rho}, \tag{11}$$

for the AR(1) case or by:

$$Z_{t+1} = \widetilde{Z}e^{(1-\phi)\overline{\mu}+\phi\mu_t}Y_t, \tag{12}$$

for the stochastic trend case.

## 3.4 Storage of Oil

The defining characteristic of the canonical model is the availability of storage technology. Here we follow the literature closely. We assume free entry into the storage sector as well as risk neutrality, implying that the actions of arbitrageurs will raise or lower the current price until it is at a level which renders the strategy unprofitable in expectation, unless that would require holding negative stocks, at which case inter-temporal arbitrage will be incomplete. In all other cases, i.e. when equilibrium at time $t$ is fully optimal, the price of oil must obey the following arbitrage condition:

$$P_t = \beta E_t[P_{t+1}] - C, \tag{13}$$

where $\beta = 1/(1+r)$ is the discount factor, and $r > 0$ is the exogenously given interest rate. The parameter $C > 0$ denotes the per barrel cost of storage[14]. Equilibrium price $P_t$ must be such that there is no incentive to increase or decrease $X_t$, the amount stored.

The inter-temporal price condition (13) does not hold in the case of a stockout, i.e. the case where $X_t = 0$ because the storage non-negativity constraint is binding. In this case arbitrageurs expect the future price of oil to be sufficiently lower than the current price that they would sell any amount of oil they had, except that they have nothing left to sell; every barrel of extracted oil is being used for consumption. As a result, current price is above its unconstrained level:

$$P_t > \beta E_t[P_{t+1}] - C. \tag{14}$$

Note that storage involves an intertemporal choice, whereas in our model the production decision does not. This is worth mentioning since models of the oil market which emphasize non-renewability imply that producers must decide whether to extract a barrel of oil today or tomorrow, which leads to a conflation of supply and storage decisions. In our model these decisions are completely separate.

### 3.5 The Rational Expectations Equilibrium

The canonical commodity storage model is a rational expectations model with one state variable - availability of oil $A_t$ - and one choice variables - storage of oil $X_t$. A solution of the model - the rational expectations equilibrium - consists of a storage rule, which specifies the level of storage for every possible value of the state variable. Determination of price and consumption follows immediately from this rule. In our extended version of the model the rule retains its salient characteristics, well known from the literature (see below). But in the extended version, similarly to the AR(1) case considered by Chambers and Bailey (1996), storage is also the function of one (or two) exogenous variables, depending on assumptions regarding the income process. Relative income $Y_t/\overline{Y}_t$ - how far above or below its mean is the current level of income - serves as the second state variable of the model when we assume that income follows a stable trend. For the case where income is subject to growth shocks, we need a third state variable: the current growth rate of income, denoted by $\mu_t$.

[14] The cost of storing a barrel of oil have most likely decreased over time. We ignore this for simplicity, since accounting for a downward trend in storage cost cannot explain the observed changes in price persistence or volatility.

In order to solve the model we express all quantity variables in their normalized forms. The model can be then be summarized by two (or three) transition functions (for the state variables) and one response equation (for the decision variable). We therefore arrive at a $2 \times 2$ framework: two alternatives for the demand process and two for the supply regime. Agents in the model observe all the state variables every period, and decide on storage accordingly, taking into consideration expectations regarding the next period's price, and implicitly producers' behavior.

The transition functions for the stable trend case are:

$$a_{t+1} = \frac{x_t + z_{t+1}}{(Y_t/\overline{Y}_t)^{\rho-1} e^{\overline{\mu}+\varepsilon_{t+1}}}, \tag{15}$$

$$\frac{Y_{t+1}}{\overline{Y}_{t+1}} = \left(\frac{Y_t}{\overline{Y}_t}\right)^{\rho} e^{\varepsilon_{t+1}}, \tag{16}$$

where equation (15) is derived by normalizing equation (1) by $Y_{t+1}$ and using (6). Effective supply $z_{t+1}$ is arrived at by dividing either equation (10) or (11) through by $Y_t$, depending on the supply regime in effect.

For the stochastic trend case, there are three transition functions:

$$a_{t+1} = (x_t + z_{t+1})/e^{\mu_{t+1}}, \tag{17}$$

$$\frac{Y_{t+1}}{\overline{Y}_{t+1}} = e^{\mu_{t+1}-\overline{\mu}} \frac{Y_t}{\overline{Y}_t}, \tag{18}$$

$$\mu_{t+1} = (1-\varphi)\overline{\mu} + \varphi\mu_t + \upsilon_t, \tag{19}$$

where the transition function (17) is derived by normalizing equation (1) by $Y_{t+1}$ and using (7) instead. Here as well, the supply regime in effect determines how we arrive at $z_{t+1}$: dividing either equation (10) or (12), as appropriate, by $Y_t$.

The response equation for both cases is:

$$(a_t - x_t)^{-\gamma} = \beta E_t[P_{t+1}] - C. \tag{20}$$

Note importantly that equation (20), which determines optimal storage, holds only when the state variables are such that the optimal storage is non-negative. If the state variables dictate negative storage, this response condition breaks down and we have simply $P_t = a_t^{-\gamma}$.

The existence and uniqueness (under certain general conditions) of the rational expectations equilibrium, as well as its important properties, have been proven in the literature. In particular, Chambers and Bailey (1996) prove these properties for the case of auto-correlated supply shocks. However, commodity storage models generally cannot be solved analytically even in their most simple form, due to the non-negativity constraint. We therefore follow the literature since Gustafson's (1958) original contribution and proceed to solve the model numerically. This can be done using a variety of methods[15]. For computational reasons, we choose to use the spline collocation method (see Judd, 1998, and Miranda and Fackler, 2002 for a discussion, and appendix B for more details).

The storage rules in our extended model are identical in form to the ones that result from the canonical model. The difference is that in the extended model these rules hold for the *normalized* variables instead of the original quantities. In other words, effective storage has a relationship with effective availability in the extended model, under both sets of assumptions regarding demand, and both supply regimes, that is qualitatively similar to the relationship between actual storage and actual availability in the canonical model.

This last point is clearly shown in Figure 3, which exhibits the optimal storage rule as well as the corresponding equilibrium price, both as functions of effective oil availability $a_t$ (on the horizontal axis), with the other state variable(s) held constant[16]. The figure is qualitatively similar regardless of our assumption on income's stochastic process or the supply regime.

**Figure 3 Here.**

In the figure, points on the curves that correspond to a particular level of effective availability represent the rational expectations equilibrium - effective storage and the resulting equilibrium price - that would prevail if effective oil availability were indeed at that level. As illustrated in the figure, effective storage is zero when the effective amount of oil available is low, then after a kink at $\overline{a}$ it rises monotonically[17]. The marginal propensity to store is always less than one; that is because a rise in storage must lower the expected future price, as it raises future availability of oil. The kink in the storage rule occurs when an additional barrel of stored oil will generate an expected profit of zero:

$$\overline{a}^{-\gamma} = \beta E[z^{-\gamma}] - C, \tag{21}$$

[15] Williams and Wright (1991), Chap. 3, survey the numeric methods applied to commodity storage models in the literature.

[16] Certain assumptions need to be made regarding the model's parameters in order to solve the model numerically. Demand elasticity $-1/\gamma$ is set at -0.2. The cost of storage $C$ is 0.02 per barrel. The discount factor $\beta$ is set at 0.97. The trend income growth rate $\overline{\mu}$ is set at 0.02, the income persistence parameter $\rho$ is set at 0.6, and the growth persistence parameter $\phi$ is set at 0.45. Effective supply capacity $\widetilde{Z}$ is set at $e^{\overline{\mu}}$. Lastly, the income shock's standard deviation $\sigma$ is set at 0.1, and the growth shock's standard deviation $\upsilon$ is set at 0.02.

[17] See Deaton and Laroque (1992), Theorem 1. When availability is relatively low (oil is temporarily scarce), agents will sell off all existing inventories of oil while the equilibrium price is high, expecting it to fall in the following period. Storage will be therefore zero, and indeed would have been negative had it been possible - agents would want to hold a short position.

where $z$, recall, is normalized production (which is a function of current and trend income; see below). The kink at $\overline{a}$ is also seen in the equilibrium price function $p$: when oil is relatively abundant, i.e. $a > \overline{a}$, the price function is more elastic. That is because once storage kicks in, a rise in oil availability causes a less than proportionate rise in the amount available for consumption, since there is now competing demand from arbitrageurs who are keen to increase their stocks.

# 4 Dynamic Behavior of Storage and Price in the Extended Model

We can now examine the dynamic behavior of the model following shocks to the income process. We first examine the dynamic behavior of our extended model under the assumption that shocks to income are AR(1), and show that our extended model behaves very similarly to the model analyzed and estimated by Deaton and Laroque (1996). In particular, storage in these conditions serves its classic purpose, to mitigate shocks: arbitrageurs transfer stocks from times of plenty to times of want. This is true under both supply regimes. Therefore storage cannot explain price volatility when shocks follow an AR(1) process in levels, a well known result. We then proceed to show that the model's dynamic behavior is quite different when we assume that demand is subject to growth shocks and supply is inelastic. Storage will then act to magnify shocks in the model, thereby increasing price volatility[18].

## 4.1 AR(1) Shocks to Income

Consider first a positive and persistent shock to income, when the stochastic process is AR(1). The shock's effects on effective availability, effective storage and equilibrium price are depicted in Figure 4. The figure exhibits the results of the following simulation: we let the system run for 15 periods, where each period a new value of $\varepsilon_t$ is drawn from the appropriate distribution. We perform 10,000 repetitions of this simulation, with the figure showing mean values for each period. This produces the baseline case. We then repeat this exercise with one change, namely that in period 2 there occurs a three standard deviation positive shock to $\varepsilon_t$. The mean results of 10,000 repetitions of this simulation are shown in dashed lines. Lastly, we repeat the simulation while restricting storage to zero at all times. This allows us to highlight its role in dynamic behavior. All of the above simulations were

[18] Kilian (2008c) argues that oil price movements that cannot be explained by either supply or industrial demand shocks should be thought of as shocks to precautionary demand. The endogenous storage response in our model is also separate from the direct effects of the shock, however it responds to the expected mean price, rather than to its expected volatility as is the case with precautionary demand.

done under the restrictive supply regime. Results of the same simulations done under the flexible supply regime are shown in Figure 5.

The shock to income results in *effective* availability dropping sharply, leading therefore to an immediate rise in current equilibrium price, as a fixed amount of oil must satisfy a larger thirst for it. These twin effects on effective availability and price subside gradually over time, as the income shock dissipates, and the system returns to its steady state.

The shock's effects on storage are more complex. Arbitrageurs are caught between two contradictory forces following this type of shock: on the one hand, the rise in current prices and drop in current effective availability dictate a drop in optimal effective storage according to the storage rule (see Figure 3). On the other hand, due to the shock's persistence future income is also expected to be higher than average, implying higher expected future prices and therefore an increased incentive to store at any level of effective availability. However, the former effect must dominate the latter in the case of an AR(1) income process, as future relative income is expected to be *less* than current relative income, implying that future effective availability is expected to be higher than its current value, and accordingly that the future price is expected to be lower than its current, post-shock level. As a result, the effect of a positive shock to income would be to reduce storage until the system reverts back to its steady state.

**Figure 4 Here.**

We see in the system's response to a temporary and persistent demand shock the underlying reason for the disappointment expressed by Deaton and Laroque (1996) regarding the storage model's inability to account for the auto-correlation seen in commodity prices. When shocks are transitory, i.e. when the system is stationary, storage acts as a countervailing force: in Figure 4, when the equilibrium price is above its steady state level, storage is smaller than its own steady state level, in a partial compensation for the shock. This in itself does contribute to a higher equilibrium price in the next period, as observed in the data. However, persistence of the shock only serves to reduce the magnitude of this response, since the connection between current and future conditions formed by the shock's persistence substitutes in part for the inter-temporal connection that is due to arbitrage. Therefore an AR(1) shock does not deliver the added persistence that Deaton and Laroque (1996) are looking for. We can see this clearly when we compare the system's response with and without storage. Where storage is not possible, availability must drop by more, and equilibrium price rises by more, relative to the case where storage is possible.

Figure 5, where supply is flexible, shows a very similar pattern. Since suppliers react to the shock, storage is less important, and the difference between the cases with and without storage are smaller. It is still the case however that storage serves as a countervailing force, i.e. it mitigates the effects of the demand shock.

**Figure 5 Here.**

## 4.2 Growth Shocks to Income

Our extended model allows us, as we have seen, to incorporate growth shocks into the storage framework. In this case, storage does not act as a countervailing force anymore; indeed, immediately following the shock storage tends to *magnify* the shock's effect on equilibrium price. Figure 6 demonstrates the effects of a positive and persistent shock to income growth, in this case a three standard deviation positive shock to $\nu_t$. Here again supply is restricted. As in the AR(1) case, a positive demand shock lowers effective availability and raises the equilibrium price. However, in this case the shock brings about a transition to a new steady state in which effective availability is expected to be at a lower level permanently, accompanied by a permanently higher price level. Importantly, due to positive auto-correlation in the stochastic process, this transition is spread over several periods, providing a role for storage.

As in the AR(1) case, arbitrageurs are subject to contradictory forces: the current rise in price induces a corresponding drop in storage, while the prospect of higher prices in the future induces a storage increase. However, the crucial difference between the two cases is that here, due to the shock's persistence, equilibrium price in the future is expected to *increase* relative to the current, post-shock price. As a result, in the stochastic trend case the storage-increasing effect of future prices is stronger than the storage-decreasing effect of the current price. Storage in the transition period is therefore *higher* in expectation relative to the expected path it would follow had the shock not occurred. In the stochastic trend case, the shock's persistence magnifies the storage response instead of diluting it as in the stable trend case[19].

**Figure 6 Here.**

It is revealing to compare our simulated response to the case where storage is not allowed. Due to the shock to income growth, storage rises sharply upwards, leading to a slightly *higher* equilibrium price relative to the no-storage case. Effective availability in the periods of transition to the new steady state is high relative to the no-storage case, since storage

[19]The same logic applies to the opposite case, where income growth suffers a negative shock, but with a caveat. Storage response, in this case a decrease, is stronger the more persistent the growth shock. However, since storage cannot be non-negative, this effect is bounded in the negative growth shock case.

remains positive throughout the transition (in expectation of higher prices in the future). Therefore we see a slower expected convergence to steady state price relative to the no-storage case.

In Figure 7 we see the same shock affecting a system in which supply is flexible. The magnifying effect of storage is completely absent: since suppliers will react to the shock by increasing supply, price is expected to return to its baseline level, and the role of storage in dynamic behavior is quite small.

**Figure 7 Here.**

We see then that in the presence of growth shocks and restricted supply, a rise in current price will be associated with an increase in optimal storage, rather than a decrease as would always be the case in the canonical model. Storage in this case does not "lean against the wind", as is its customary role; it actually magnifies the shock somewhat, by increasing demand exactly when it is already high, in preparation for even higher demand in the future. This behavior then provides the link between price persistence and price volatility. We saw that when prices are not persistent, either because demand is stationary or because supply is flexible, storage can only serve to reduce the effects of shocks. Only when prices are persistent, due to the presence of growth shocks as well as restrictions on access to supplies, could storage act to increase the effect of shocks, and therefore increase price volatility immediately after the shock, above and beyond what it would otherwise be without storage.

This result accords well with our empirical findings: periods in which persistent growth shocks are dominant should be periods in which price exhibits extra volatility, relative to periods in which AR(1) shocks are more prevalent and / or supply is flexible. At this point we must be cautious not to read too much into this result, as we offer no calibration to the data. Our contribution is to show that the oil market can only behave in a speculative way under certain conditions. In the next section, we provide more detail on the historical contexts in which the oil market operated, in particular the interaction between industrialization and the structure of the oil industry.

# 5 Industrialization and Market Structure: Transition Points in Context

In Section 2 we identify three points of transition. In 1877-8 and again in 1972-3, oil price persistence and volatility both changed, while in 1934 we find a change in volatility, but no change in persistence. Of these three points, only 1934 can be linked to a major oil discovery, that of the East Texas Oil Field a few years earlier. All three points of transition, we will argue, had to do with changes in market structure. In 1878 began the construction of Tidewater, the first long-distance pipeline, which eventually ended railroad monopoly over the transportation of oil. In 1934 the Federal government gained virtual control over U.S. oil production. In 1970 the East Texas Oil Field peaked, ending U.S. control over excess exploitable reserves, and signalling the rise to prominence of OPEC[20].

These three points also show striking similarities and differences from a demand point of view as well. Of these three points, 1934 is special also in that it occurs in the midst of a worldwide recession. Both 1877-8 and 1972-3 were years in which the global economy, and with it demand for oil, were booming, driven by the large-scale industrialization of the United States and of East Asia, respectively. Rapid industrialization is by definition a transitional stage, and as such it features growth rates that are on the one hand unsustainably high and on the other hand quite persistent, since the process of industrialization often stretches over decades.

We will argue that the two observed periods in which oil prices were both highly persistent and highly volatile occurred because two conditions were simultaneously met in each of these periods: access to supplies was restricted and demand was unsustainably high. It is important to note that U.S. industrialization was far from over when the first such period ended in 1878, and that post-war industrialization in East Asia was well underway by 1973, our estimate of the beginning of the second period of high persistence and high volatility. These years were not turning points in oil demand, rather they signified major structural changes in the petroleum industry, in which key players with the ability to restrict access to supplies either emerged or declined in importance. When only one of the conditions was met, as for example happened during both World Wars (when demand was unsustainably high but supply was unrestricted), the market was significantly less persistent and less volatile. This necessary confluence of demand and supply factors has been relatively rare looking back all the way to 1860, but of course has been the reality in the oil market in recent decades[21].

It is worth emphasizing that we do not focus on the source of shocks to the oil market, nor do we attempt to identify these shocks. Our main objective is to account for the radically

[20] The oil industry was, of course, much bigger in the late 20th century compared to its size a hundred years earlier; oil is also no longer used mainly for illumination as it did at first. It is notable however that prominent features of the industry remain relatively unchanged: oil was internationally traded from the very beginning, demand for it being global. Moreover, its efficiency as a source of energy made it indispensable to consumers from the earliest stages, a feature of the industry that remains crucial to this day.

[21] We stress supply restrictions rather than reserve depletion since there is no evidence that "running out of oil" was ever a real danger. Oil security, the danger of not having access to existing oil, was on the other hand very real. In this regard, "capacity constraints" must be viewed as mechanisms to restrict supply, since by their nature these constraints - derricks, strorage tanks - can be loosened in the medium run, whereas the amount of extractable oil in any given field cannot be increased beyond a certain point.

different behavior of oil prices across different regimes, rather than determining the genesis of shocks to oil prices within a given regime. Studies such as Hamilton (2009b) and Kilian (2008b, 2009) exclude this type of analysis due to their intentional focus on the experience of recent decades.

## 5.1 From Rail to Pipe: Explaining the 1877-8 Change-Point

By 1865, the Oil Regions in northwestern Pennsylvania, where all commercial oil production was located, were well served by three different railroads[22]. These firms enjoyed an oligopolistic position, as both production and refining were highly competitive. For illustration purposes, in 1877, the year before Standard Oil consolidated its control of refining, the "open fare" for rail transport of crude oil from the Oil Regions to New York was \$1.40 per barrel (Bentley, 1979, page 28); this amounted to 58% of the average price of a barrel of crude oil in that year according to our data. Williamson and Daum (1959, Chap. 17) estimate the per barrel cost of carriage by rail at no more than \$0.40 around that time, giving us an idea of the margins involved. Rockefeller's vision was a large refining concern that could bargain effectively with the railroads; Standard's business advantage was well understood at the time to consist of the special "rebates" that it was in a position to demand from the railroads[23].

In 1878, oil producers who were trying to break the joint monopoly of transportation and refining started construction on the world's first long-distance pipeline, the Tidewater. It was completed in May of 1879, and was intended to bypass both the railroads and Standard Oil's refineries. In the face of this technological breakthrough, Rockefeller proceeded to construct Standard's own long-distance pipelines, choosing to destroy the railroads' monopoly on transportation in order to strengthen Standard's monopoly on refining. This spelled the end of attempts to increase profitability by restricting access to oil. Having invested in his own infrastructure, and given the very low transportation costs it afforded him[24], Rockefeller's strategy now was to sell as much oil as possible.

These events occurred against a backdrop of ever rising demand, both domestic and foreign. The United States was going through rapid industrialization at the time, eventually overtaking Britain as the world's leading center of manufacturing. During the period, the share of world industrial output made in the U.S. rose spectacularly, from 7.2% in 1860, to 14.7% in 1880, to 23.6% in 1900. In absolute numbers, U.S. manufacturing production rose by a factor of three in the two decades between 1860 and 1880, and by a factor of eight

[22] Yergin (1991) is the source for all historical facts, unless noted otherwise.

[23] Standard's business advantage over independent refiners as a result of its strong bargaining position is estimated by Bentley (1979) to have been \$1.00 per barrel of refined oil.

[24] Williamson and Daum (1959, p. 458) estimate that per barrel transportation costs in Standard Oil's own pipelines were between \$0.12 - \$0.20, less than half the rail cost of carriage.

between 1860 and 1900[25]. U.S. population more than doubled from 1860-1900, rising from 31.8 million to 76.4 million, while GDP per capita rose almost as fast, from $2,445 in 1860 to $4,091 in 1900 (constant 1990 dollars)[26]. As a result domestic consumption of illuminating oil rose from 1.6 million barrels (mb) in 1873-75 to 12.7 mb in 1899, while that of machine lubricating oil rose even more, from 0.2 mb in 1873-75 to 2.4 mb in 1899 (Williamson and Daum [1959], pp. 489, 678). Even as urban communities in the United States and Europe shifted to gas or electric lighting, kerosene remained in high demand in other parts of the world. By the turn of the twentieth century, there was increasing demand for gasoline, from the burgeoning auto industry.

The transition point we identify in 1877-8 was therefore the starting point of sweeping changes in market structure, brought about in an environment of rapidly growing demand. Before 1878 the railroads were using their monopolistic position to limit the supply of crude to the markets in the interest of rent extraction. After 1878 that power was slipping away from them at a fast clip; by 1884 Rockefeller's network of long-distance pipelines was essentially complete, and the railroads were sidelined. Moreover, since Standard Oil owned the vast majority of long-distance pipelines, and with demand expected to continue unabated, there was no player in the market who had both the interest and the capability of limiting supplies[27].

## 5.2 Oil Glut and Government Control: Explaining the 1933-4 Change-Point

This state of affairs continued until the early 1930's, when the discovery of the East Texas Oil Field created an oil glut of proportions heretofore unknown. Against the backdrop of Depression, the result was a slump in price that threatened the entire industry, eventually leading to Federal regulation. The U.S. government came to effectively control supplies: since East Texas production was far below its potential, and given the authority to raise and lower production quotas as circumstances required, the U.S. government (both Federal and state, in particular the Texas Railroad Commission) had the power to increase or decrease oil supply almost at will. Over the decades since, while it still had that power, the U.S. government would use it to stabilize the market on numerous occasions. It increased production enormously during World War II, as well as during supply crises involving the Middle East,

[25] Bairoch (1982) is the source for the U.S. absolute and relative industrial output numbers.

[26] Figures for U.S. population and GDP per capita are from Maddison (2003).

[27] The producers in the Oil Regions were relatively small-scale, and had repeatedly failed in their attemps to control production. There was one exception: in 1887-8, there was a willingness to cooperate on the part of Standard Oil, and production reduction was achieved for a short while.

in 1953 (Iran), 1956 (Suez), and 1967 (Six-Day War). When the surge of oil was no longer needed, it had the power to reduce production once more. U.S. regulation thus acted as an automatic stabilizer (Yergin, 1991, page 259). This had the effect of reducing the standard deviation of supply and demand shocks, and accords well with the observed reduction in volatility that we date to 1934, around the time that this mechanism went into effect. Quite the opposite from the railroads' rent extraction strategy before 1878, U.S. government agencies aimed to stabilize price by adjusting quantity as needed. The supply of oil, far from limited, was in effect quite flexible.

## 5.3 U.S. Oil Peak and OPEC: Explaining the 1972-3 Change-Point

Our third transition point is 1973, where we find that oil price persistence and volatility both increased. In 1970 U.S. oil production reached its peak. In March 1971 the Texas Railroad Commission, for the first time since World War II, allowed production at 100% capacity; the ability of U.S. government agencies to increase production in times of need was gone (Yergin, 1991, pp. 567-8). Excess capacity existed now only in the Middle East, giving the rulers of these countries the same kind of market power enjoyed by the railroads almost a century earlier: the ability to extract large rents from consumers by limiting access to oil supplies. In the Gulf, 1972 saw an agreement of "participation" of oil producers, i.e. the transfer of some ownership rights of the oil resources located on their land from the international oil companies to the governments. These developments changed fundamentally the nature of the market: the oil producing countries were now owners (whole or part) of their reserves, the only easily-exploitable oil reserves left in the world. In November 1973, of course, OPEC member states acted to dramatically restrict the West's access to oil supplies[28].

As in the early years of the oil industry, these events were occurring in a period of increasing demand. The demand for oil is driven, first and foremost, by income. In recent decades, world GDP and global oil production have moved in lockstep; the International Energy Agency estimates long-run income elasticity of world oil demand at about 0.5, i.e. each percentage point increase in world GDP is accompanied by a 0.5% increase in the global demand for oil (IEA [2006,2007]). This may in fact be an underestimate: Gately and Huntington (2002) find that income elasticity of oil demand in OECD countries is 0.55, but for non-OECD countries the income elasticity may be as high as 1. The IEA estimates that in 2001-2005, China's higher-than-average propensity to consume oil may have raised the

[28] There is a debate in the literature on whether OPEC can be shown to have acted collusively to withhold supplies from the market. Clearly OPEC's degree of control over prices has been inconsistent over the years, however that in itself does not settle the issue: Smith (2005, 2008) and Almoguera and Herrera (2007) are recent references. In our context, what matters is only OPEC's ability to restrict access to the world's only easily exploitable reserves of oil. This ability is undisputed (Smith, 2008).

global income elasticity to 0.8 (IEA [2007]).

This time it was East Asia that was industrializing fast: first Japan, then Taiwan and South Korea, and finally China. Japan's GDP per capita, for example, more than tripled in two decades, rising from $3,986 in 1960 to $13,428 in 1980. Japanese GDP by itself already equaled 37% of U.S. GDP by 1980 (all comparisons in 1990 international dollars, Maddison [2003]). China industrialized slightly later, but the pace of its industrialization in the final decades of the twentieth century was as rapid as that of the U.S. a century earlier, if not more: between 1980 and 2000, Chinese industrial production rose by a factor of 9; between 1970 and 2000, it grew by a factor of 21. In relative terms, the Chinese share of world industrial output was only 0.7% 1970; it has increased to 6.3% in 2000[29]. The IMF's World Economic Outlook (2008) projects that Asia's share of global trade and manufacturing will continue to soar in the coming decades, despite the short term dislocations caused by the current financial crisis. Overall, the average growth rate in Asia (excluding Japan) between 1973-2001 was 5.4%, compared with 2.1% for Western Europe, 3.0% for the United States, and 2.7% for Japan during the same years (Maddison [2003]). It seems clear that Asia, outside of Japan, was still very much in the midst of an era of industrialization at the onset of the present century, with no end in sight. This is similar to the situation of the U.S. economy about a hundred years earlier.

We see therefore a repeat, on a much broader scale, of important features from the market environment that prevailed before 1878: a combination of supply limits and ever rising demand[30]. As in the earlier period, with supply limited in this fashion, shocks to demand would be fully incorporated into the price. Since these shocks are very persistent, in an era where the trend in demand is uncertain, we would expect the price of oil to be more persistent, relative to a period where these limits on supply are not binding. This persistence in the price of oil can be reasonably expected to continue as long as demand shocks are persistent, or until the ability of OPEC to effectively limit access to supplies no longer exists, either due to an independent source of oil, or to an alternative source of energy.

# 6 Conclusion

We argue in this paper that a long-term view is essential to understanding the dynamic behavior of oil prices. We show that shocks to the oil market can have remarkably different effects on the real price of oil across historical periods, not because of their origin on the

[29] Statistics for China are from the World Bank's World Development Indicators database.

[30] OECD growth was also high at various points during this time period, a fact that no doubt has been important in the timing of the first and second oil shocks (see Barsky and Kilian [2002, 2004]). However this fact cannot account for the very high price persistence that we observe in this period.

supply or the demand side, but rather due of the ability (or lack thereof) of key players in the market to restrict access to supplies. With effective restrictions on access to excess supplies, growth shocks can generate oil prices that are both highly persistent and, through an endogenous storage response, highly volatile. On the other hand, without these restrictions, the same growth shocks will be quickly accommodated, and will not lead to increased persistence or volatility. In this regard, it is immaterial whether the growth shocks originate on the demand or the supply side.

The literature, for obvious reasons, has focused on the extremely persistent and volatile post-1973 period. Given the geo-political realities, this period is unlikely to end in the foreseeable future. However, it is useful to recognize that throughout most of the history of oil, the ability to restrict access to supplies was actually sorely lacking, with the oil market showing remarkable flexibility and relative price stability as a result. This held true even in years when oil supply or demand were experiencing great upheavals, such as during World War II and the postwar re-building of Europe. Although rare, the history of oil shows that shifts in industry structure do occur, and the structural breaks in price behavior associated with these shifts are testimony to their importance.

# References

[1] Almoguera, Pedro A., and Ana María Herrera. 2007. "Testing for the Cartel in OPEC: Noncooperative Collusion or Just Noncooperative?" https://www.msu.edu/~herrer20/research.html.

[2] Alquist, Ron, and Lutz Kilian. 2010. "What Do We Learn from the Price of Crude Oil Futures?" *Journal of Applied Economics*, 25(4): 539-573.

[3] Bai, Jushan and Pierre Perron. 1998. "Estimating and Testing Linear Models with Multiple Structural Changes." *Econometrica* 66(1): 47-78.

[4] Bai, Jushan and Pierre Perron. 2003. "Computation and Analysis of Multiple Structural Change Models." *Journal of Applied Econometrics*, 18: pp. 1-22.

[5] Bairoch, Paul. 1982. "International Industrialization Levels from 1750 to 1980." *Journal of European Economic History*, 11: 269-310.

[6] Barsky, Robert B. and Lutz Kilian. 2002. "“Do We Really Know that Oil Caused the Great Stagflation? A Monetary Alternative.” *NBER Macroeconomics Annual* 2001: 137–83.

[7] Barsky, Robert B. and Lutz Kilian. 2004. "Oil and the Macroeconomy Since the 1970s." *Journal of Economic Perspectives*, 18(1): 115-134.

[8] Bentley, Jerome T. 1979. *The Effects of Standard Oil's Vertical Integration into Transportation on the Structure and Performance of the American Petroleum Industry, 1872-1884.* New York: Arno Press.

[9] British Petroleum. 2010. *BP Statistical Review.* http://www.bp.com/statisticalreview.

[10] Busetti, Fabio and A.M. Robert Taylor. 2004. "Tests of Stationarity Against a Change in Persistence." *Jornal of Econometrics*, 123(1): 33-66.

[11] Caner, Mehmet and Lutz Kilian. 2001. "Size Distortions of Tests of the Null Hypothesis of Stationarity: Evidence and Implications for the PPP Debate." *Journal of international Money and Finance,* 20(5): 639-657.

[12] Cavaliere, Giuseppe and A.M. Robert Taylor. 2008. "Testing for a Change in Persistence in the Presence of Non-stationary Volatility." *Jornal of Econometrics,* 147(1): 84-98.

[13] Deaton, Angus and Guy Laroque. 1992. "Behaviour of Commodity Prices." *Review of Economic Studies,* 59(1): 1-23.

[14] Deaton, Angus and Guy Laroque. 1996. "Competitive Storage and Commodity Price Dynamics." *Journal of Political Economy,* 104(5): 896-923.

[15] Hamilton, James D. 2008. "Oil and the Macroeconomy." in *The New Palgrave Dictionary of Economics, 2nd Edition*, eds. Steven N. Durlauf and Lawrence E. Blume. New York: Palgrave Macmillan.

[16] Hamilton, James D. 2009a. "Causes and Consequences of the Oil Shock of 2007-08." *Brookings Papers on Economic Activity*, Spring 2009: 215-59.

[17] Hamilton, James D. 2009b. "Understanding Crude Oil Prices." *Energy Journal*, 30(2): 179-206.

[18] Hansen, Bruce E. 2000. "Testing for Structural Change in Conditional Models." *Journal of Econometrics,* 97(1): 93-115.

[19] Harvey, David I., Stephen J. Leybourne, and A.M. Robert Taylor. 2006. "Modified Tests for a Change in Persistence." *Journal of Econometrics* 134(2): 441-469.

[20] Hotelling, Harold. 1931. "The Economics of Exhaustible Resources." *Journal of Political Economy* 39(2): 137-175.

[21] International Energy Agency. 2006. *Oil Information*. Paris: OECD.

[22] International Monetary Fund. 2008. *World economic Outlook.* Washington, DC.

[23] Gustafson, Robert L. 1958. "Carryover Levels for Grains: A Method for Determining Amounts That Are Optimal under Specified Conditions." USDA Technical Bulletin 1178.

[24] Judd, Kenneth L. 1998. *Numerical Methods in Economics.* Cambridge, MA: MIT Press.

[25] Kilian, Lutz. 2008a. "The Economic Effects of Energy Price Shocks." *Journal of Economic Literature,* 46(4): 871-909.

[26] Kilian, Lutz. 2008b. "Exogenous Oil Supply Shocks: How Big Are They and How Much Do They Matter for the U.S. Economy?" *Review of Economics and Statistics,* 90(2): 216-240.

[27] Kilian, Lutz. 2009. "Not All Oil Price Shocks Are Alike: Disentangling demand and Supply Schocks in the Crude Oil Market." *American Economic Review* 99(3): 1053–69.

[28] Kim, Jae-Young. 2000. "Detection of Change in Persistence of a Linear Time Series." *Journal of Econometrics,* 95(1): 97-116.

[29] Kim, Jae-Young, Jorge Belaire-Franch, and Rosa Badillo Amador. 2002. "Corrigendum", *Journal of Econometrics,* 109(2): 389-392.

[30] Krautkraemer, Jeffrey A. 1998. "Nonrenewable Resource Scarcity," *Journal of Economic Literature,* 36(4): 2065-2107.

[31] Kwiatkowski, Denis, Peter C.B. Phillips, Peter Schmidt, and Yongcheol Shin. 1992. "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root." *Journal of Econometrics,* 54: 159-178.

[32] Maddison, Angus. 1998. *Chinese Economic Performance in the Long Run.* Paris: OECD.

[33] Maddison, Angus. 2003. *The World Economy: Historical Statistics.* Paris: OECD.

[34] Miranda, Mario J. and Paul L. Fackler. 2002. *Applied Computational Economics and Finance.* Cambridge, MA: MIT Press.

[35] Ng, Serena, and Francisco J. Ruge-Murcia. 2000. "Explaining the Persistence of Commodity Prices." *Computational Economics,* 16(1-2): 149–171.

[36] Pindyck, Robert S. 1999. "The Long-Run Evolution of Energy Prices." *Energy Journal,* 20(2): 1-27.

[37] Routledge, Bryan R., Duane J. Seppi, and Chester S. Spatt. 2000. "Equilibrium Forward Curves for Commodities." *Journal of Finance,* 55(3): 1297-1338.

[38] Smith, James L. 2005. "Inscrutable OPEC? Behavioral Tests of the Cartel Hypothesis." *Energy Journal,* 26(1): 51-82.

[39] Smith, James L. 2008. "Organization of the Petroleum Exporting Countries (OPEC)," in *The New Palgrave Dictionary of Economics, 2nd Edition*, eds. Steven N. Durlauf and Lawrence E. Blume. New York: Palgrave Macmillan.

[40] Schwartz, Eduardo, and James E. Smith. 2000. "Short-Term Variations and Long-Term Dynamics in Commodity Prices." *Management Science,* 46(7): 893-911.

[41] Williams, Jeffrey C. and Brian D. Wright. 1991. *Storage and Commodity Markets.* Cambridge, England: Cambridge University Press.

[42] Williamson, Harold F. and Arnold R. Daum. 1959. *The American Petroleum Industry.* Vol. I*, The Age of Illumination 1859-1899.* Evanston, IL: Northwestern University Press.

[43] Wright, Brian D. 2001. "Storage and Price Stabilization." in *Handbook of Agricultural Economics.* Vol 1B*, Marketing, Distribution, and Consumers,* eds. Bruce L. Gardner and Gordon C. Rausser*,* 817-862. New York: Elsevier.

[44] Wright, Gavin. 1990. "The Origins of American Industrial Success, 1879-1940." *American Economic Review,* 80(4): 651-68.

[45] Yergin, Daniel. 1991. *The Prize.* New York: Simon & Schuster.

**Table 1: Sample Statistics of Oil Price Series**

| | Subsamples | | | Full Sample |
|---|---|---|---|---|
| | 1861-1877 | 1878-1972 | 1973-2009 | 1861-2009 |
| Price (2009 $US) | | | | |
| Mean | 50.9 | 17.2 | 44.7 | 27.8 |
| Std. Deviation | 25.3 | 5.1 | 22.1 | 20.2 |
| Annual price change (abs.) | | | | |
| Mean | 39.0% | 14.2% | 22.1% | 19.0% |
| Std. Deviation | 27.5% | 13.7% | 22.8% | 19.6% |

Data source: BP Statistical Review.

**Table 2: Testing for Change in Persistence of Real Oil Price**

| Sample | 1861-1955 | 1881-2009 |
|---|---|---|
| Change-point (direction) | 1877 ($I(1)$ to $I(0)$) | 1973 ($I(0)$ to $I(1)$) |
| 95% CI | (1873, 1881) | (1968, 1978) |

*Mean Score Statistics:*

| | | |
|---|---|---|
| $MS$ | 0.03 (0.954) | 31.98** (0.033) |
| $MS^R$ | 127.61*** (0.002) | 0.10 (0.918) |
| $MS^{MAX}$ | 125.99*** (0.002) | 31.86** (0.033) |

*Mean-Exponential Statistics:*

| | | |
|---|---|---|
| $ME$ | 0.02 (0.966) | 36.67** (0.035) |
| $ME^R$ | 271.22*** (<0.001) | 0.06 (0.888) |
| $ME^{MAX}$ | 264.24*** (<0.001) | 36.53** (0.036) |

*Maximum Statistics:*

| | | |
|---|---|---|
| $MX$ | 0.22 (0.830) | 83.78** (0.032) |
| $MX^R$ | 575.38*** (<0.001) | 1.86 (0.342) |
| $MX^{MAX}$ | 561.67*** (<0.001) | 83.40** (0.032) |

Note: Only statistically significant change-points are listed. Three asterisks (***) denote significance at the 1% level, two asterisks(**) denote significance at the 5% level. T-statistics are in parentheses. Data source: BP Statistical Review.

**Table 3: Testing for Change in Volatility of Oil Price, 1861-2009**

Estimated breakpoints, confidence intervals, and period mean volatilities:

| | Breakpoint | 95% CI | Mean Volatility (t-statistic) |
|---|---|---|---|
| Period I | 1878 | 1873-1896 | 0.381 (9.4) |
| Period II | 1933 | 1931-1946 | 0.195 (9.0) |
| Period III | 1972 | 1950-1973 | 0.061 (2.3) |
| Period IV | | | 0.228 (8.5) |

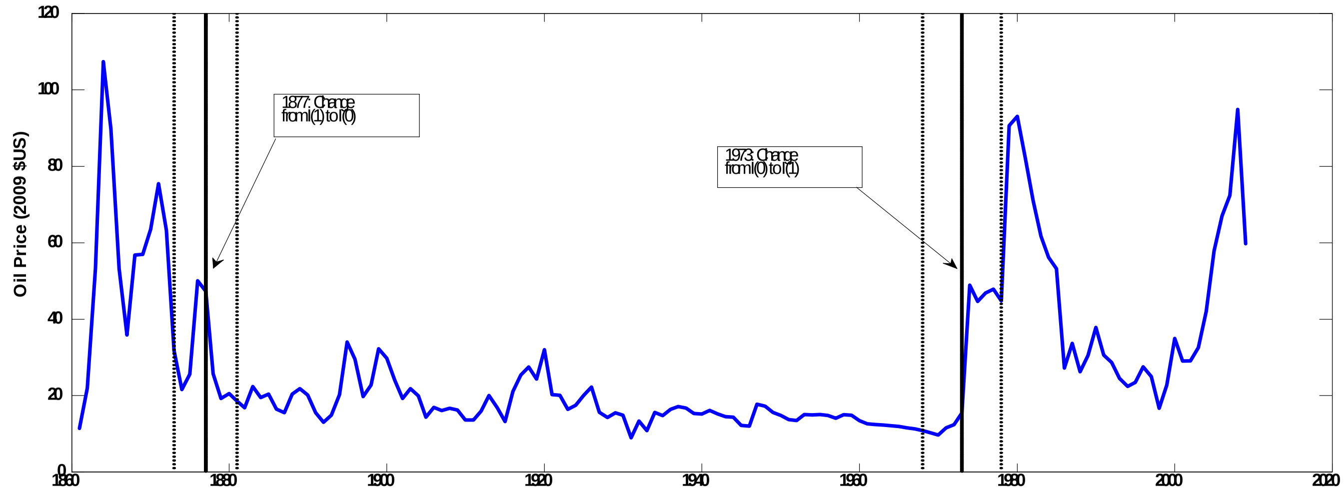
The Bai-Perron test statistics:

| Test Statistic | Value |
|---|---|
| $\sup F_T$(1 break vs. no breaks) | 21.89*** |
| $\sup F_T$(2 breaks vs. no breaks) | 14.85*** |
| $\sup F_T$(3 breaks vs. no breaks) | 16.00*** |
| Sequential Procedure: | |
| $\sup F_T$(2 breaks vs. 1 break) | 7.66 |
| $\sup F_T$(3 breaks vs. 2 breaks) | 32.07*** |
| $\sup F_T$(4 breaks vs. 3 breaks) | 3.15 |

Three asterisks (***) denote significance at the 1% level, two asterisks(**) denote significance at the 5% level.

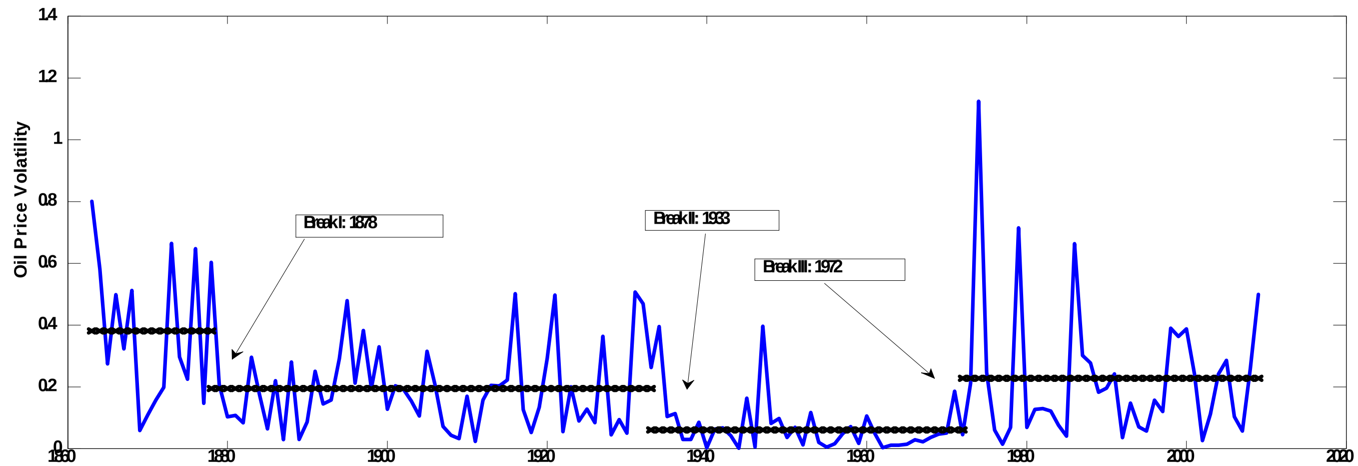
Data source: BP Statistical Review.

Figure 1: Real Oil Price, 1861-2009




Data source: BP Statistical Review. Change-point estimates and 95% confidence intervals taken from Table 2.

Figure 2: Real Oil Price Volatility, 1861-2009




Data source: BP Statistical Review. Change-point estimates taken from Table 3.

Figure 3: Rational Expectations Equilibrium

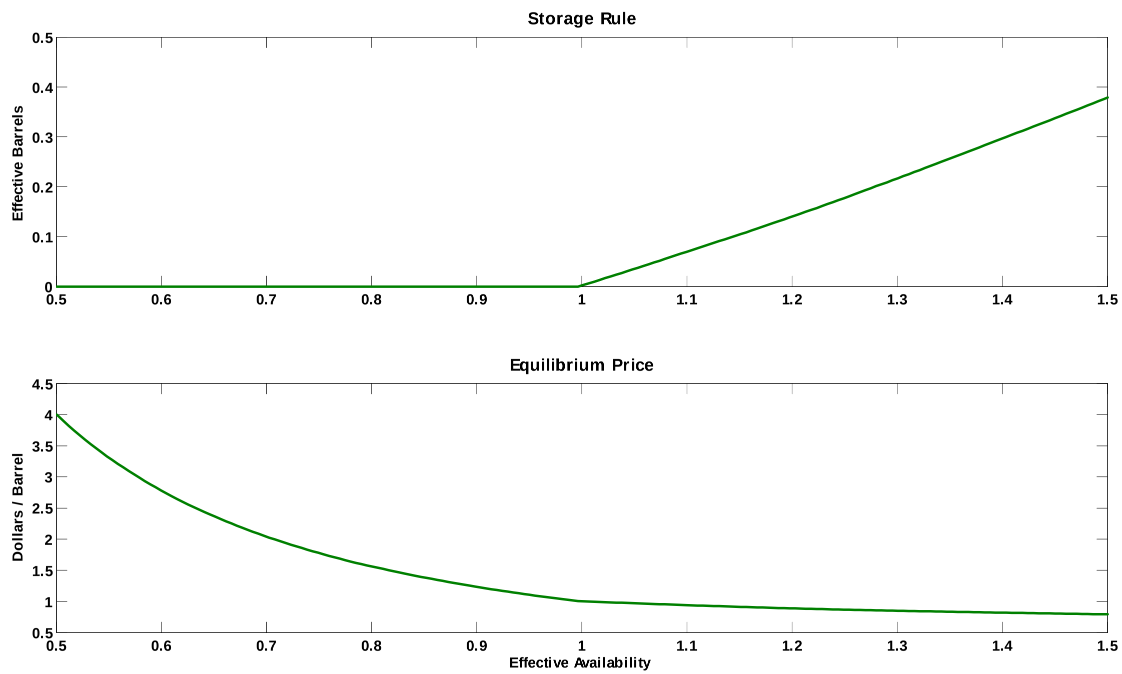


Figure 4: Dynamic Behavior: AR(1) Process with Restricted Supply

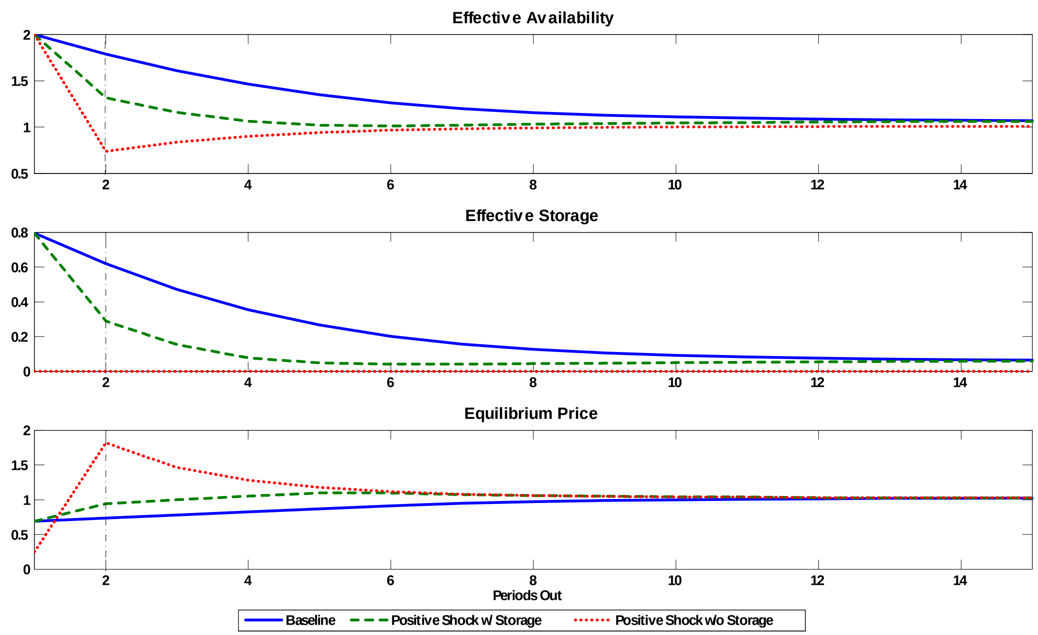


Figure 5: Dynamic Behavior: AR(1) Process with Flexible Supply

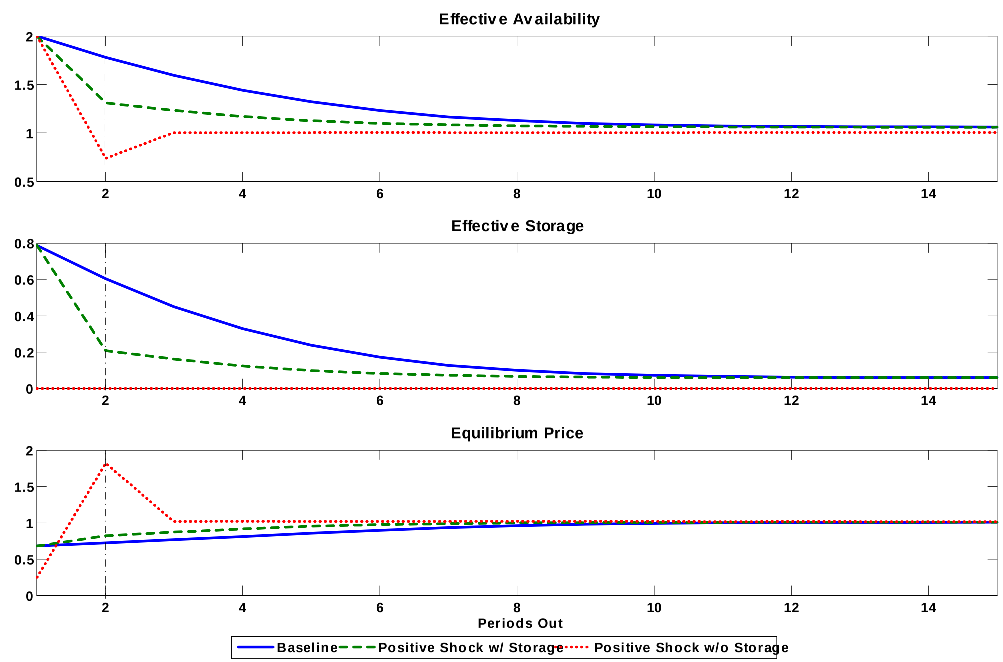


Figure 6: Dynamic Behavior: Growth Process with Restricted Supply

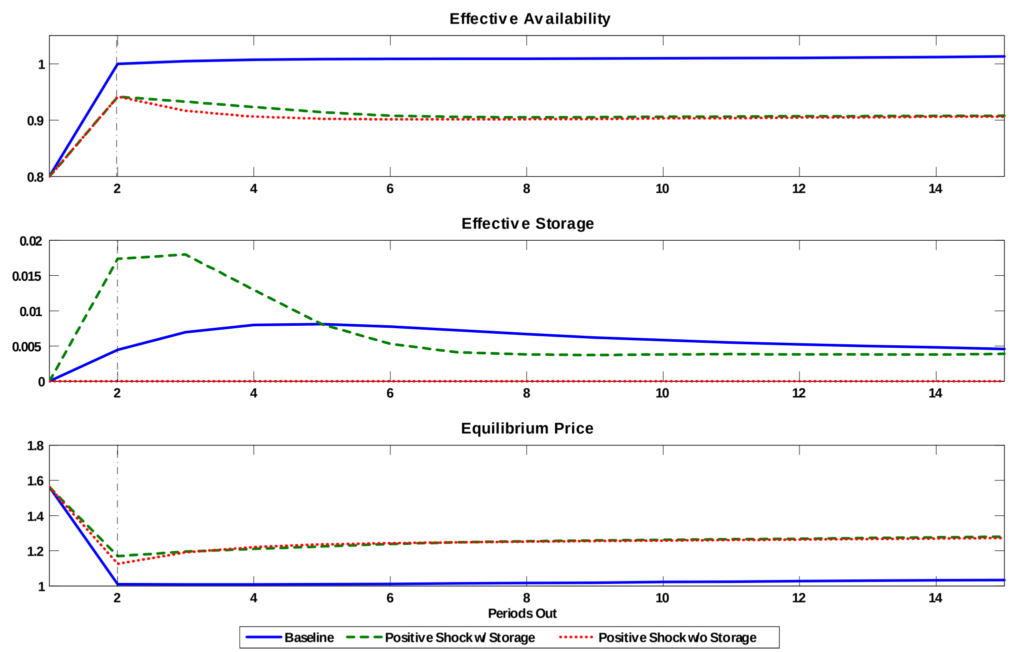
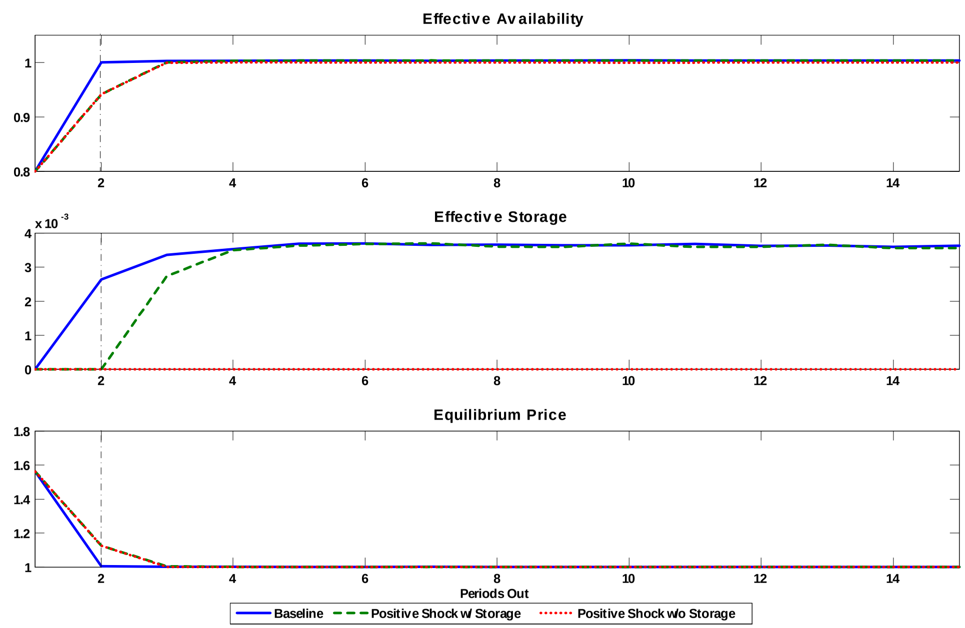


Figure 7: Dynamic Behavior: Growth Process with Flexible Supply



# A Methods Used to Test for Change in Persistence

Consider a time series $y_t$, where $t = 1, ..., T$. Assume the series can be decomposed into the sum of a deterministic trend, a random walk, and a stationary error:

$$y_t = \xi t + r_t + \varepsilon_t, \tag{22}$$

where $r_t$ is the random walk component:

$$r_t = r_{t-1} + u_t. \tag{23}$$

Let the errors $u_t$ be iid with mean zero and variance $\sigma_u^2$. Then one can test the null hypothesis of $I(0)$ by positing $H_0 : \sigma_u^2 = 0$ against the alternative $H_1 : \sigma_u^2 > 0$. The test is constructed as follows: let $e_t$ denote the residuals from a regression of $y_t$ on a constant and a trend. Then consider the following test statistic:

$$K = \frac{1}{\widehat{\sigma}_\varepsilon^2} \sum_{t=1}^{T} S_t^2, \tag{24}$$

where $\widehat{\sigma}_\varepsilon^2$ is the estimated error variance, and $S_t$ denotes the partial sum process:

$$S_t = \sum_{i=1}^{t} e_i, \quad t = 1, ..., T. \tag{25}$$

A value of this test statistic that is higher than an appropriate critical value would imply a rejection of the $I(0)$ null. Kim (2000), later modified and corrected by Kim *et al.* (2002), and Busetti and Taylor (2004), apply this method to the question of change in the rate of persistence of a series. With the same null hypothesis, consider the following two alternative hypotheses:

$$H_{01} : \begin{cases} \sigma_u^2 = 0, & t = 1, ..., \tau T \\ \sigma_u^2 > 0, & t = \tau T + 1, ..., T \end{cases}, \tag{26}$$

$$H_{10} : \begin{cases} \sigma_u^2 > 0, & t = 1, ..., \tau T \\ \sigma_u^2 = 0, & t = \tau T + 1, ..., T \end{cases}. \tag{27}$$

The point $\tau T$ at which the change from $I(0)$ to $I(1)$ (under $H_{01}$) or vice versa (under $H_{10}$) is assumed unknown, and is estimated during the testing procedure. The test is carried out as follows. At each possible change-point (i.e. at all points in the range $[\tau_l T, \tau_u T]$), compute two sets of residuals: let $\bar{e}_t$ denote the residuals from a regression of $y_t$ , $t = 1, ..., \tau T$ on a constant and a trend, and let $\widetilde{e}_t$ denote the residuals from a similar regression for the observations $t = \tau T + 1, ..., T$. Define the partial sum processes accordingly:

$$\widetilde{S}_t = \sum_{i=\tau T+1}^{t} \widetilde{e}_i, \quad (28)$$

$$\overline{S}_t = \sum_{i=1}^{t} \overline{e}_i. \quad (29)$$

For all $\tau \in [\tau_l, \tau_u]$ define the following statistic:

$$K_{[\tau T]} = \frac{[T(1-\tau)]^{-2} \sum_{t=\tau T+1}^{T} \widetilde{S}_t^2}{[\tau T]^{-2} \sum_{t=1}^{\tau T} \overline{S}_t^2}. \quad (30)$$

Had $\tau$ been known with certainty, this statistic (evaluated at $\tau$) could be used to test the null of a pure $I(0)$ process against the alternative $H_{01}$. A high value of the statistic would imply a rejection of the $I(0)$ null. Since in general the true $\tau$ is not known, Kim(2000) suggests using three functions of the sequence of $K_{[\tau T]}$ over the range $\tau \in [\tau_l, \tau_u]$. The limits $\tau_l, \tau_u$ are arbitrarily chosen, commonly 0.2 and 0.8, respectively[31]. The three functions are given by:

$$\begin{aligned} MS &= \frac{1}{\tau_u - \tau_l + 1} \sum_{t=\tau_l T}^{\tau_u T} K_t, \\ ME &= \log \left[ \frac{1}{\tau_u - \tau_l + 1} \sum_{t=\tau_l T}^{\tau_u T} \exp{(K_t)}^{0.5} \right], \\ MX &= \max_{t=\tau_l T, \ldots, \tau_u T} K_t. \end{aligned}$$

Busetti and Taylor (2004) show that it is possible to use the reciprocal of $K_t$ to test the $I(0)$ null against $H_{10}$. We define the functions $MS^R$, $ME^R$, $MX^R$ in a similar manner, substituting $K_t^{-1}$ for $K_t$ everywhere. A third set of test statistics can be used to test the null against *any* change in persistence, whether from $I(0)$ to $I(1)$ or vice versa. These are

[31] In our 1861-1965 sample, we choose instead $\tau_l = 0.1$ and $\tau_h = 0.9$, since the change-point occurs quite early in the sample. If we choose the more common range, the change-point becomes 1879, the earliest year allowed.

defined as follows:

$$\begin{aligned} MS^M &= \max\{MS, MS^R\} \\ ME^M &= \max\{ME, ME^R\} \\ MX^M &= \max\{MX, MX^R\} \end{aligned}$$

Further, HLT show that all nine statistics, in a modified form, can also be used to test an $I(1)$ null against $H_{01}$, $H_{10}$, or both. The modification they propose corrects the test statistics so that they have the correct size under *both* $I(0)$ and $I(1)$. HLT show that this can be achieved by multiplying the relevant test statistic by $\exp(-bJ)$, where $b$ is a constant[32], and $J$ denotes the Wald statistic for testing the joint hypothesis that in the following regression:

$$y_t = \gamma_0 + \gamma_1 t + \gamma_2 t^2 + ... + \gamma_9 t^9,$$

the coefficients of all higher order trends (i.e. $\gamma_2, ..., \gamma_9$, quadratic trend and above, in the standard case) are zero[33]. HLT also allow for the test to include local to unit root behavior as well as true unit root behavior, so that $H_{01}$can be thought of as a significant change in persistence from $I(0)$ to a rate of persistence that is very close to 1, but not necessarily exactly 1. The same holds for $H_{10}$.

The HLT procedure assumes homoskedasticity, which is inappropriate for our purposes given the results of section 2.2. We therefore employ the correction suggested by Cavaliere and Taylor (2008), i.e. we introduce a wild bootstrap. We generate bootstrap samples by multiplying the vector $e_t$, $t = 1, ..., T$ by random numbers taken independently from a $N(0,1)$ distribution, and repeating the process 10,000 times. The bootstrap samples by construction replicate the heteroskedasticity patterns in the price data. We then proceed to calculate bootstrap analogs for all test statistics using bootstrap analogs of $\overline{e}_t$ and $\widetilde{e}_t$ , i.e. using residuals obtained by regressing the bootstrap samples on a constant and a trend. This produces for every test statistic 10,000 bootstrap analogs, which we then use to calculate the bootstrap p-values shown in Table 2. A complication is that the HLT modified test-statistics are sized correctly only in the homoskedastic case, and are therefore not well suited for the wild bootstrap procedure. However these statistics are computed directly from, and are uniformly lower than, the non-modified test statistics. We bootstrap the non-modified statistics, and compute bootstrap p-values using the lower modified statistics.

[32] HLT provide different values for $b$ for different levels of significance. We choose the values appropriate for the 1% level, which are largest and therefore would reduce the likelihood of rejection.

[33] Due to computational reasons, we limit the polynomial in our tests to the $6^{th}$degree. The value of the Wald statistic is quite robust to changes in the degree of the polynomial.

These p-values should therefore be viewed as biased upwards.

A final step in the testing procedure, taken only if the tests indicate that a change in persistence has indeed taken place, is to estimate the change-point $\tau T$. Kim (2000) suggests the formula:

$$\Lambda(\tau) = \frac{[T(1-\tau)]^{-2}\sum_{t=\tau T+1}^{T}\widetilde{e}_t^2}{[\tau T]^{-2}\sum_{t=1}^{\tau T}\overline{e}_t^2},$$

where, in the case of rejecting the null in favor of $H_{01}$, the estimated change-point is:

$$\tau_{01} = \arg\max_{t=\tau_l,\ldots,\tau_u}\Lambda(\tau),$$

and in the case of rejecting the null in favor of $H_{10}$, the estimated change-point is:

$$\tau_{10} = \arg\min_{t=\tau_l,\ldots,\tau_u}\Lambda(\tau).$$

# B Solving the Model

We solve the model numerically using spline collocation (see Judd, 1998, and Miranda and Fackler, 2002)[34]. The following summarizes the logic of this method, using as an example our simplest case of two state variables (effective availability $a$ and realtive income $y$) and one choice variable (effective storage $x$), when supply is restricted. Other versions of the model are solved in exactly the same way, changing the variables and equations as appropriate.

The collocation method approximates an unknown function (in this case, the equilibrium price function) by a linear combination of known functions of the state variables. Approximating the price function leads to better results compared with approximating the storage rule directly, since the former function is relatively smoother. We have:

$$f(a,y) = [a - x(a,y)]^{-\gamma} \approx \sum_{m=1}^{N_a}\sum_{n=1}^{N_y} b_{m,n}\phi_{m,n}(a,y), \tag{31}$$

where $N_a$, $N_y$ denote chosen number of points in each state variable's support, called the collocation nodes, where the equilibrium condition (13) must hold exactly. This produces $N_a N_y$ nonlinear equations, one for each collocation node, and a corresponding number of coefficients $b_{m,n}$ to be determined. We use polynomial splines for the bivariate basis functions $\phi_{m,n}(a,y)$, since splines perform better in approximating functions with a derivative discontinuity as we have here.

[34] We use the CompEcon Toolbox for MATLAB, provided by Paul Fackler on his website: www4.ncsu.edu/~pfackler/compecon/toolbox

The expectation operator is dealt with by a discretization of the known distribution of the shock, i.e. by assigning probabilities $w_k$ to particular points on the distribution's support $\varepsilon_k$, where $k = 1, ..., K$, such that the continuous density is approximated by the two vectors $(\varepsilon_k, w_k)$. The method then proceeds to find the coefficients of the linear approximation by using a double iteration. First, given an initial guess for the coefficients $b_{m,n}$, it finds the values of the choice variable for which the equilibrium condition holds at the collocation nodes. The equilibrium condition at node $i$ must then hold exactly, allowing us to solve for $x_i$:

$$(a_i - x_i)^{-\gamma} = \beta \sum_{k=1}^{K} \sum_{n=1}^{N_a} \sum_{m=1}^{N_y} w_k b_{m,n} \phi_{m,n} [\frac{x_i + \widetilde{Z}/y_i}{y_i^{\rho-1} e^{\overline{\mu}+\varepsilon_k}}, y_i^{\rho} e^{\varepsilon_k}] - C. \tag{32}$$

It is now possible to use the vector $x$ generated in this way to update the coefficients $b_{m,n}$ by solving the following system of equations:

$$(a_i - x_i)^{-\gamma} = \sum_{m=1}^{N_a} \sum_{n=1}^{N_y} b_{m,n} \phi_{m,n}(a_i, y_i), \tag{33}$$

so that at every node $i$ the approximating function is exactly equal to the equilibrium price. These updated coefficients are now used again to produce updated values for the response variable at each of the collocation nodes, and so on until the coefficients converge.