CrossMark

Macroeconomic Dynamics, **22**, 2018, 683–701. Printed in the United States of America.
doi:10.1017/S1365100516000511

DISCERNING TRENDS IN COMMODITY PRICES

Dimitri Dimitropoulos and Adonis Yatchew
University of Toronto

This paper emphasizes the use of data-driven techniques to discern trends in commodity prices. Whereas many previous papers rely on parametric assumptions, we take as our departure point the view that such trends are *inherently* nonparametric, driven by complex forces of innovation and obsolescence, competition and strategic behavior, resource depletion and discovery, and supply and demand. Our reference specification is the partial linear model $y_t = f(t) + z_t\beta + \varepsilon_t$ where macroeconomic variables z enter parametrically, and f is nonparametric, to be discovered using cross-validation.

We analyze data on 11 commodities—3 hydrocarbons and 8 metals. For the majority of these commodities, our data-driven estimates of f bear close similarity to band pass estimates which include long term trends and super cycles. The OPEC effect is estimated to have increased oil prices by over 50% on average and coal prices by about 25%. U.S. coal mining legislation is estimated to have increased coal prices by 9% to 15%.

Keywords: Commodity Prices, Oil Prices, Semiparametric Modeling, Trends, Supercycles

1. INTRODUCTION

From a statistical viewpoint, the Hotelling prediction—that the price of a depletable resource should rise at the rate of interest—has been decidedly rejected. Resource prices have risen and fallen. They are driven by a complexity of forces—innovation and obsolescence, competition and strategic behavior, depletion and discovery, and of course, supply and demand.

The tectonic shift in oil prices during the latter part of 2014 provides an especially vivid example. This is a story which contains each of the above ingredients. Technological breakthroughs—namely horizontal drilling and "fracking"—have dramatically increased recoverable resources. This, in turn, has fundamentally altered strategic behavior: whereas in the past OPEC might have coordinated a supply reduction to sustain prices, it is far more limited now in its ability to do so because shale producers (and others) can fill the gap. The peculiar features of the shale revolution have altered the nature of supply in critical ways. Not only are entirely new sources now technologically viable, they can be brought online in

We are especially grateful to Mathieu Marcoux for assistance on multiple aspects of this project, and to Lutz Killian and Angelo Melino for helpful comments. The support of the Social Sciences and Humanities Research Council is gratefully acknowledged. Address correspondence to: Adonis Yatchew, Economics Department, University of Toronto, Toronto M6S 3G7, Canada; e-mail: yatchew@chass.utoronto.ca.

© 2017 Cambridge University Press 1365-1005/17

https://doi.org/10.1017/S1365100516000511 Published online by Cambridge University Press

tiny increments—the cost of a productive shale well is *three orders of magnitude* smaller than Arctic or deep-sea projects. Thus, while the rhetoric suggests that well-endowed conventional suppliers are trying to drive out U.S. shale, the latter should be able to return quickly if prices increase. In short, shale has provided for a *scalable* response by many producers as market conditions change. This feature further limits unilateral or cartelized market power. Scalability of shale also reduces risks—wells do not last long, but capital requirements are low and lead-times short.

The richness of this story contrasts sharply with, for example, Hotelling's Rule. Thus, from the modeler's perspective, at one extreme one can opt for simple Hotelling-type trend models, and at the other, one can imagine supra-complex structural models which somehow incorporate supply and demand variables, technological change, strategic behavior, reserve-to-production ratios, and so forth.

This paper takes as its departure point the view that trends in commodity prices are inherently nonparametric, driven by complex forces such as those described above. For the purposes of this paper, we define trend to be a deterministic smooth function of time of unknown shape which we identify using data-driven methods. Our reference specification is the partial linear model, where trend is nonparametric and various macroeconomic and other variables enter parametrically. Throughout this paper, when we speak of detrending, we mean removal of a smooth trend effect (hence the need for flexible nonparametric modeling). How does one discern trends from short-term random price fluctuations? We suggest that a sensible approach is to use cross-validation to separate noise (even serially correlated noise) from systematic time effects.

We analyze data on 11 commodities—3 hydrocarbons and 8 metals—from the beginning of the 20th century into the 21st century. Some commodity groupings appear to exhibit shape similarity in their trends. In this connection, the recent plunge in world oil prices eerily follows on the heels of the drastic reductions in North American natural gas prices that occurred in recent years, both having been driven primarily by the fracking revolution but each occurring in starkly different market environments. The introduction of macroeconomic variables permits one to incorporate short run business dynamics as well as the effects of longer term economic growth.

Before proceeding, we provide a brief overview of the literature on deterministic trend models as applied to commodities.[1] Hotelling's (1931) seminal paper on optimal resource extraction has, for many decades, provided an important reference model for the pricing of exhaustible resources. The basic prediction, that the "resource rent", or net price (market price minus marginal cost), of an exhaustible resource will increase in the long run at the rate of interest, has formed the basis of important theoretical work on the economics of nonrenewable resources. However, Hotelling's model has typically been rejected by empirical researchers who instead have found that prices have also fallen, at least over certain periods of time.

An early empirical study of long run trends in nonrenewable resource prices was conducted by Barnett and Morse (1963) who sought to examine the scarcity of

https://doi.org/10.1017/S1365100516000511 Published online by Cambridge University Press

natural resources. They found that the relative prices of mineral commodities had actually been falling, and thus concluded that natural resource scarcity was not a problem. Smith (1978, 1979) reexamined Barnett and Morse's (1963) conclusions, by regressing the relative price of a resource, as measured by its price divided by a wholesale price index, on a constant and a linear time trend. Smith found no statistically stable time trends.

Heal and Barrow (1980) developed an arbitrage model where investors choose between investing in resources and investing in capital markets. Their model results in a differential equation for the evolution of the growth rate of a resource price in terms of the interest rate and the growth rate of industrial production, which the authors then estimate for various metal prices. Heal and Barrow found that it is the change in the interest rate, and not its level, which explains the growth rate of resource prices. Smith (1981) extended Heal and Barrow (1980) with a particular focus on long term movements of fuel and mineral price series. Smith compared the forecasting performance of a set of models for the growth rate of a resource price in terms of arbitrage behavior and found that the Heal-Barrow specification was consistently among the "best" model in terms of ex post forecasting performance.

Hotelling's original result was derived under idealized conditions. Since then, other researchers have extended his basic model to introduce additional factors including: aggregate economic growth [Stiglitz (1974), Solow and Wan (1976)]; exploration for new reserves [Pindyck (1978)]; uncertainty about the size of the resource base [Gilbert (1979)], or about demand and costs [Pindyck (1980)]; technological change in extraction methods [Slade (1982)]; imperfect competition [Salant (1982), Gilbert (1978)]; and tax distortions [Sweeny (1977), Dasgupta et al. (1980)]. Gaudet (2007) surveyed the theoretical literature on extensions of the Hotelling model, while Livernosis (2009) surveyed the empirical literature on Hotelling's model of resource extraction. Slade (2009) provided a survey of both the theoretical and empirical literatures on Hotelling-type models of resource extraction.[2]

An important implication of some of these extensions is the possibility of periods of decreasing and increasing prices. At least two types of arguments have been put forth to rationalize the existence of U-shape price paths. In a widely cited study Slade (1982) argues that if, initially, technological change is sufficiently rapid, then prices can fall over time. However, as reserves are depleted and the rate of technological change slows, prices eventually rise, leading to a U-shaped price path. Several others, for example, Pindyck (1978), Livernois and Uhler (1987), Swierzbinski and Mendelshohn (1989), have shown that exploration can give rise to U-shaped price paths if reserves are initially small and extraction costs depend on reserves; prices begin at relatively higher levels, fall as exploration expands the resource base, but then rise slowly as reserves decline.

Slade (1982) was the first to conduct empirical tests of U-shape price paths; she analyzed data on 11 natural resource commodities, thus setting off a series of such studies. Using a quadratic specification for time trend, Slade concluded that real

https://doi.org/10.1017/S1365100516000511 Published online by Cambridge University Press

prices of most of these commodities had reached their respective minima and were increasing. Since then, although employing different methodologies and datasets, Moazzami and Anderson (1994), Berck and Roberts (1996), and Pindyck (1999), all found at least some evidence of convex curvature in the price paths of many depletable resources.

Moazzami and Anderson (1994), expressed concerns over the potential bias that may arise if the time trend is unstable, and if there are short-run deviations from the long run trend. They proposed an "error correction" model that allowed for short-run price dynamics around a quadratic trend and found that, for the majority of the commodities they considered, a U-shaped curve captured the long run trend.

Similar to Slade (1982), Berck and Roberts (1996) first estimated a quadratic trend-stationary model for the price series of a variety of metals, and found U-shapes. They performed unit-root tests and found that almost all of the price series they considered are difference-stationary, thus casting doubt on their initial Slade-like findings.

Citing Perron's (1989) work on the bias that structural breaks can cause in unit-root tests, several authors, e.g. Ahrens and Sharma (1997), Lee, List, and Strazicich (2006), re-examined the stationarity issue using the 11 original commodities studied by Slade (1982). Ahrens and Sharma applied an Augmented Dicky-Fuller test and found evidence of trend-stationarity in 6 out of the 11 commodity price series. Lee, List, and Strazicich (2006) employed a Lagrangian multiplier unit root test that allowed for endogenously determined structural breaks in the presence of a deterministic quadratic trend, and found that all 11 of the price series were trend-stationary.

With an eye towards forecasting, Pindyck (1999) studied the long run evolution of energy prices, in particular oil, natural gas, and coal. He first fitted a quadratic trend model, and performed variance ratio tests to see whether prices follow unit roots. After finding U-shaped price paths and evidence of mean-reversion (albeit slow) towards a fluctuating trend, Pindyck proposed a model of mean-reversion to a stochastic trend, à la Ornstein-Uhlenbeck, and argued that such a model is more consistent with the theory of depletable resource prices.

There is also a growing literature on the use of band pass filter methods [Christiano and Fitzgerald (2003)] to estimate time effects across various ranges of frequencies. These include Cuddington and Jerrett (2008), Jacks (2013), and Erten and Ocampo (2013). Jacks find evidence of decades-long deviations in commodities prices from their long run trends, giving credence to the hypothesis of super cycles. Using a similar approach, Erten and Ocampo decompose the time series of nonoil commodity prices, metal prices, agricultural prices, and crude oil prices into cyclical components of varying frequencies. After filtering out long run trend, they find evidence of super cycles that last in the range of 30 to 40 years. Erten and Ocampo then compute the correlation between these price super cycles and global macroeconomic output indicators, and find that commodity price super cycle components are more correlated with lags of the world GDP, suggesting that world GDP cycles lead super cycles in commodity prices.

https://doi.org/10.1017/S1365100516000511 Published online by Cambridge University Press

These band pass filter methods are inherently nonparametric, and so in this sense, are similar to the ideas we develop here. However, our approach differs: first, our identification of the "time effect" is accomplished using data-driven methods while band pass filter methods divide the "time effect" somewhat arbitrarily into frequency ranges; second, our approach easily permits the introduction of additional covariates—such as business cycle and macro variables—directly into the model.

2. MODEL AND ESTIMATION

Consider the partial linear model

$$y_t = f(t) + z_t\beta + \varepsilon_t, \tag{1}$$

where f is an unknown smooth function of time t ; z_t is a vector of observable variables; β is a conformable vector of unknown parameters; and ε_t is an error term which may possess heteroskedastic and time series properties. We choose this specification for several reasons. First, the approach allows highly flexible modeling of trend, which is our primary focus. Second, macroeconomic effects included in z_t are arguably more amenable to parametric modeling; for example, one might expect that increases in unemployment and lower economic growth rates would have monotone adverse effects on base metals prices such as copper. Third, the vector z_t may also include dummy variables which influence commodity prices, such as the profound shift in strategic behavior by OPEC in 1973, or U.S. Federal legislation affecting coal mining in 1969 and 1970. Fourth, while one could consider models where all right hand side variables enter nonparametrically, such an approach quickly leads to highly imprecise estimation as a result of the "curse of dimensionality". In short, the partial linear specification provides a reasonable compromise between flexibility and precision.

Of course, the parametric variables (for example, economic growth rates in OECD or developing countries) may also be trending. Define $E(z_t|t) \equiv g(t)$ where the vector function g models smooth trend effects in the parametric variables. Write $E(y_t|t) \equiv h(t) = f(t) + g(t)\beta$ and

$$y_t - h(t) = [z_t - g(t)]\beta + \varepsilon_t. \tag{2}$$

The left-hand-side represents "detrended" price, and the right-hand-side the linear combination of detrended parametric variables which purportedly influence price. The model in equation (2) has a "double residual" structure which permits $\sqrt{n}-$ consistent estimation of the parameters β [Robinson (1988)]. In particular, one performs nonparametric regressions of the dependent variable and each of the parametric variables on time trend to obtain, say $\hat{h}$ and $\hat{g}$. Substituting these into equation (2) one then obtains $\hat{\beta}$ using OLS. Since the residuals are likely to be serially correlated and heteroskedastic, we rely upon Newey-West (1987) standard errors adapted to the current semiparametric setting [see e.g. Yatchew (2003),

https://doi.org/10.1017/S1365100516000511 Published online by Cambridge University Press

pp. 51–52]. To estimate f one removes the estimated parametric effects and performs nonparametric regression on the model $y_t - z_t\hat{\beta} \cong f(t) + \varepsilon_t$.

Estimation of nonparametric regressions can be accomplished using a range of smoothers, such as kernel, loess, and spline type estimators.[3] In some cases, for example in modeling oil prices, certain macroeconomic variables (e.g. unemployment) are likely to be endogenously determined. In such cases, we use lagged values of macro variables as instruments.

Implementation of cross-validation in a time-series setting requires some modification. Consider, for example $y_t = h(t) + \varepsilon_t$. If the residuals are independent, then in order to prevent over-fitting, the smoothing parameter, say λ, is commonly selected by omitting each observation, one at a time, and predicting the function value at that point based on the remaining observations. However, if the residuals are strongly positively correlated, then inclusion of nearby observations can also lead to over-fitting. Therefore, in implementing cross-validation, for each t, we omit observations in a neighborhood of t [see e.g. Aneiros-Pérez and Quintela-del-Río (2001)]. That is, we minimize

$$CV(\lambda) = \frac{1}{T}\sum_{t=1}^{T}\left[y_t - \hat{h}_{-N(t)}(t,\lambda)\right]^2, \tag{3}$$

where $\hat{h}_{-N(t)}$ indicates that the function is estimated at t by omitting the $t - th$ observation as well as observations in a neighborhood $N(t)$ of t. Given that we are working with annual data, as a practical matter we omit five consecutive observations including that at time t.[4]

To formally assess specifications, we rely on "residual regression tests". Suppose, one wants to test $H_o : f \in \overline{\Im}$ against $H_1 : f \in \Im$ where $\overline{\Im}$ is a restricted family of functions, such as the set of quadratic functions, and $\Im$ is a broader class, such as a set of smooth nonparametric functions. Let $\hat{\varepsilon}_t$ be the estimated residuals under the null hypothesis. Define

$$U = \frac{1}{\lambda T^2}\sum_t\sum_{s\neq t}\hat{\varepsilon}_s\hat{\varepsilon}_t K_{st}, \tag{4}$$

where λ is the smoothing parameter, K_{st} is the kernel matrix with zeros on the diagonal, and off-diagonal entries defined as $K_{st} = \frac{1}{2}$ if $|x_s - x_t| \leq \lambda$.[5] Let $s_U^2 = \frac{2}{\lambda^2 T^4}\sum_t\sum_{s\neq t}\hat{\varepsilon}_s^2\hat{\varepsilon}_t^2 K_{st}^2$. Under a range of conditions, U/s_U converges in distribution to a standard normal where one rejects the null for large positive values of the statistic. Tests of this type have been applied in many nonparametric and semiparametric settings.[6] Gao and Hawthorne (2006) study this test within the partial linear model framework where time trend is nonparametric—the model considered here.

https://doi.org/10.1017/S1365100516000511 Published online by Cambridge University Press

3. DATA

We study the price series of 11 commodities: 3 fuels (oil, natural gas, coal) and 8 metals (copper, nickel, zinc, iron, tin, silver, lead, aluminum) for the period 1901 to 2014. The data extend those used by Slade (1982) but with some modifications which we outline here. Like Slade, our main source is Manthy (1978), from which we obtain data on oil, natural gas, coal, copper, zinc, iron, lead, and the wholesale price index prior to 1973. The main differences are the sources of data for nickel, tin, and aluminum; for these, we use prices (available to 1957) given in Potter and Christy (1962).[7] Silver prices are obtained from the U.S. Bureau of the Census (1975), [Slade uses Schurr (1960) and the American Metal Market's *Metal Statistics* (1979)].

Additional sources are needed to update the data to 2014. The price series for oil, natural gas and coal were augmented using prices published by the U.S. Energy Information Administration (EIA). We used the EIA (2014c) *Refiner Acquisition Cost of Crude Oil* to update our series for oil from 1973 to 2014. The EIA (2013) *U.S. Natural Gas Wellhead Price* series was used to update natural gas prices to 2012, at which point this series was discontinued; the *Henry Hub Natural Gas Spot Price* series, also available from the EIA (2014b), was used to complete the series to 2014. For coal, we used the price series for bituminous coal available from the EIA's *Annual Energy Review* (2011) to update our series for 1973 to 2011, and from the EIA's *Annual Energy Outlook* (2014a) for 2012 to 2014.

For metals prices, the U.S. Geological Survey (2013b) *Historical Statistics for Mineral and Material Commodities in the United States* was used to extend the time series of copper, nickel, zinc, tin, silver, lead, and aluminum to 2012. London Metal Exchange prices were used to complete the series through 2014. These are available from the World Bank Cross Country Database (2014a) and can be retrieved through Quandl (https://www.quandl.com). Additional sources include the U.S. Geological Survey (1999, 2013a) *Metals Prices in the United States Through 1998* and *Minerals Yearbook* and the *Steel Mill Producer Price Index*, available from the U.S. Bureau of Labor Statistics (BLS) (2014b). The wholesale price index was updated using the *Producer Price Index (Commodities)* as published by the BLS (2014a).[8]

Below, we introduce macroeconomic variables to control for short term business cycle dynamics, as well as longer term economic growth trends. We include variables on the growth rate of the Gross Domestic Product[9] (GDP) of 20 OECD countries (Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, the United Kingdom, and the United States) as well as six developing countries (Argentina, Brazil, Chile, China, India, and Indonesia). The source for these GDP data was the Maddison Project (2013) *Statistics on World Population, GDP and Per Capita GDP 1-2010,* updated to 2014 using the World Bank (2014b) *World Development Indicators* databank. The U.S. unemployment rate series was constructed from Lebergott (1957) for the period 1901–1942, and

https://doi.org/10.1017/S1365100516000511 Published online by Cambridge University Press

TABLE 1. Goodness of fit and specification tests

	Quadratic trend		Band pass long term Trend		Band pass long term Trend + Super cycles		Nonparametric trend	
	R^2	U-stat	R^2	U-stat	R^2	U-stat	R^2	U-stat
Oil	0.76	6.53	0.76	6.61	0.94	−2.76	0.96	−1.46
Gas	0.77	10.90	0.85	8.34	0.97	−1.81	0.97	−0.43
Coal	0.70	8.98	0.72	8.86	0.94	−3.74	0.97	−1.77
Copper	0.20	9.56	0.31	5.55	0.65	−3.33	0.77	−2.51
Nickel	0.47	9.27	0.54	6.12	0.77	−2.72	0.83	−1.80
Zinc	0.03	1.01	0.05	0.47	0.23	−1.96	0.23	−1.82
Iron	0.43	11.20	0.65	1.56	0.76	−3.63	0.92	−1.21
Tin	0.20	11.00	0.35	8.22	0.73	−3.59	0.80	−2.53
Silver	0.33	11.30	0.45	8.29	0.83	−3.58	0.90	−1.97
Lead	0.02	7.20	0.21	3.69	0.53	−2.26	0.42	−0.98
Aluminum	0.84	6.89	0.85	6.02	0.93	−1.21	0.90	−0.25

from the BLS (2014c) for the period 1943–2014. The U.K. unemployment rate series was constructed from Denman and McDonald (1996) for the period 1901 to 1970, and from the U.K. Office of National Statistics (2014) for the period 1971 to 2014. Consistent unemployment data for this time period across all OECD countries were not available.

4. RESULTS

4.1. Quadratic, Band Pass, and Nonparametric Model Estimates

Our initial objective is to estimate models which consist solely of deterministic functions of time, that is, models of the form $y_t = f(t) + \varepsilon_t$. We compare pure nonparametric specifications against models estimated using band pass filter methods as well as simple quadratics such as those studied by Slade (1982). Estimates using these various methods are depicted in Figure 1; goodness-of-fit measures and tests of specification are in Table 1.

We begin by estimating a simple quadratic trend. Then, we apply the cross-validation function in equation (3) to determine the appropriate level of smoothing of each price series.[10] These are used to obtain pure nonparametric estimates of the function $f(t)$.

Following Cuddington and Jerrett (2008), Erten and Ocampo (2013), and Jacks (2013), we then apply the Christiano-Fitzgerald (2003) filter to extract an estimate of long term trend using frequencies in excess of 70 years; these are displayed in Figure 1 as "Band Pass Long Term Trend" (BPLTT). Next we extract "super cycles" using frequencies between 20 and 70 years and combine these with the long term trend estimate to obtain "Band Pass Long Term Trend + Super Cycles" (BPLTTSC).

https://doi.org/10.1017/S1365100516000511 Published online by Cambridge University Press

https://doi.org/10.1017/S1365100516000511 Published online by Cambridge University Press

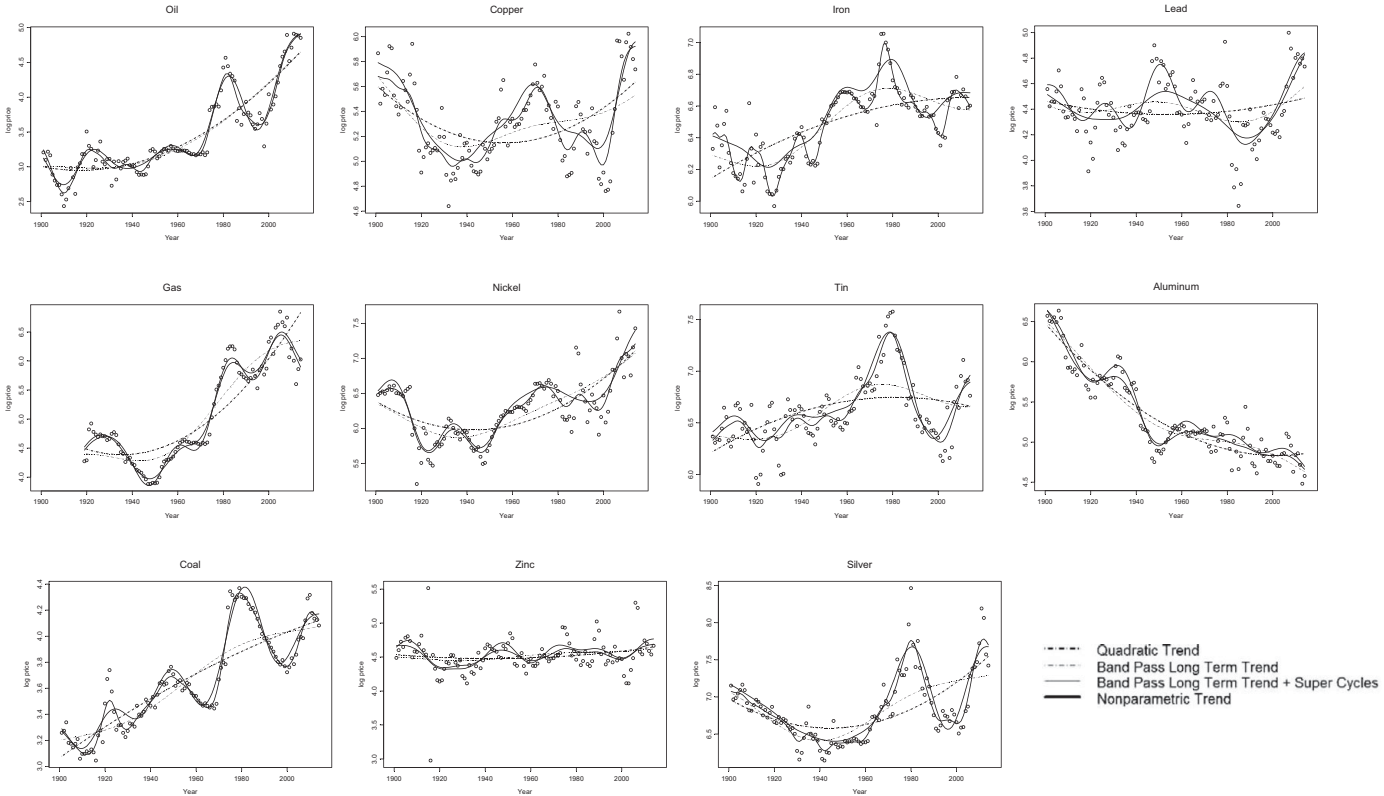


FIGURE 1. Quadratic, band pass, and nonparametric models.

A number of observations are in order. For oil, coal, zinc, and aluminium, the quadratic and BPLT estimates are almost indistinguishable, with very similar R^2 values. For natural gas and the other metals, the quadratic goodness of fit is substantially inferior to that of the BPLTT specification. In all cases, the U-statistic specification test strongly rejects the quadratic specification; the BPLTT specification is rejected in all cases except for zinc and iron. The inclusion of super cycles in the band pass framework leads to much better fit to the data for all commodities, as does modeling using our nonparametric approach.

Strikingly, for the three hydrocarbons the BPLTT and super cycle specifications track the nonparametric estimates closely (see first column of graphs in Figure 1) with correspondingly similar R^2 values (Table 1). Nonparametric models provide superior fits for copper, nickel, iron, tin, and silver relative to the BPLTTSC specifications. Keeping in mind that one rejects the null hypothesis for large *positive* values of the U-statistic, none of the BPLTTSC or nonparametric specifications are rejected.

Returning to the graphical results in Figure 1, casual inspection suggests a degree of shape similarity among the three hydrocarbons. Note the downturn in natural gas and coal prices post-2008 augurs the drop in oil prices which begins in mid-2014 and is not yet fully evident in the annual data. In the second column, copper and nickel exhibit some shape similarity, as do iron, tin, and silver in the third column. Aluminum is the only commodity which displays a strong secular downward trend.

4.2. Inclusion of Macro Variables

We now incorporate macroeconomic variables parametrically. These exhibit strong joint significance for most commodities considered here. Because of the intimate relationship between oil prices and macroeconomic performance, and the potential for substitution among hydrocarbons,[11] we use instrumental variable techniques to estimate macroeconomics effects for all three hydrocarbons. As instruments, we use lagged values of the macroeconomic variables. Results are reported in Tables 2A and 2B.

Our semiparametric oil model is estimated for 1901–2014 and for the post-war period 1946–2014. Neither OECD nor emerging economies growth rates have a significant impact on price. The average U.S./U.K. unemployment rate is marginally significant only for the full data period. The change in strategic behavior by OPEC beginning in late 1973 initiated an era during which oil prices were on average in excess of 50% higher over the subsequent four decades after taking account of trend and macro effects.

U.S. natural gas prices are available as early as 1919, however, this market goes through extended periods of regulated price constraints for inter-state trade; these are finally eliminated in the early 1990's. As natural gas markets are primarily continental, we use U.S. growth and unemployment rates as explanatory variables; we report results for 1919–2014 and the recent 1993–2014 period. For the full

https://doi.org/10.1017/S1365100516000511 Published online by Cambridge University Press

https://doi.org/10.1017/S1365100516000511 Published online by Cambridge University Press

TABLE 2A. Macro variable estimates—hydrocarbons

	Oil (1901–2014)		Oil (1946–2014)		Natural gas (1919–2014)		Natural gas (1993–2014)		Coal (1901–2014)		Coal (1946–2014)	
	Coef	SE	Coef	SE	Coef	SE	Coef	SE	Coef	SE	Coef	SE
OECD Growth	−0.0006	0.0093	−0.0158	0.0128	—		—		—		—	
EMERGE Growth	−0.0006	0.0049	0.0066	0.0067	—		—		—		—	
U.S. Growth	—		—		0.0019	0.0026	0.0571	0.0175**	−0.0004	0.0012	−0.0005	0.0013
U.S. Unemployment	—		—		0.0072	0.0056	−0.0346	0.0638	−0.0009	0.0018	−0.0082	0.0047*
U.S./U.K. Unemployment	−0.0305	0.0165*	−0.0731	0.0667	—		—		—		—	
OPEC Dummy	0.5354	0.0714**	0.5675	0.1063**	0.1634	0.2222	—		0.2544	0.0550**	0.2758	0.0416**
OPEC Market Share	—		—		—		0.1378	0.0830*	—		—	
Regulatory Dummy	—		—		—		—		0.0911	0.0374**	0.1457	0.0255**
R^2	0.96	(0.95)	0.98	(0.97)	0.97	(0.97)	0.95	(0.91)	0.98	(0.97)	0.98	(0.96)
Test of Joint	χ^2_4	p-val	χ^2_4	p-val	χ^2_3	p-val	χ^2_3	p-val	χ^2_4	p-val	χ^2_2	p-val
Significance	141.0	< 0.01	184.0	< 0.01	2.80	0.423	59.7	<0.01	30.4	< 0.01	76.6	< 0.01

Notes: Where necessary, macro variables are instrumented using their lagged values. Standard errors and tests of joint significance use the Newey-West covariance matrix estimator. **, * indicate statistical significance at 5% and 10% levels, respectively. R^2 values in parentheses correspond to the pure trend model. The test of a Quadratic Null against the Semiparametric Alternative is one-sided.

TABLE 2B. Macro variable estimates—metals 1901–2014

	Copper		**Iron**		**Lead**	
	Coef	SE	Coef	SE	Coef	SE
OECD Growth	0.0085	0.0033**	−0.0026	0.0026	0.0101	0.0043**
EMERGE Growth	−0.0035	0.0070	−0.0071	0.0044	−0.0029	0.0079
U.S./U.K. Unemployment	−0.0358	0.0104**	−0.0143	0.0068**	−0.0206	0.0084**
R^2	0.81	(0.77)	0.93	(0.92)	0.49	(0.42)
Test of Joint Significance	χ^2_3	p-val	χ^2_3	p-val	χ^2_3	p-val
	20.6	<0.01	12.0	<0.01	14.0	<0.01
	Nickel		**Tin**		**Aluminium**	
	Coef	SE	Coef	SE	Coef	SE
OECD Growth	−0.0012	0.0038	0.0133	0.0052**	0.00097	0.0035
EMERGE Growth	0.0024	0.0075	0.0034	0.0065	−0.00021	0.0057
U.S./U.K. Unemployment	0.0056	0.0157	−0.0224	0.0122*	0.00693	0.009
R^2	0.83	(0.83)	0.83	(0.80)	0.93	(0.93)
Test of Joint Significance	χ^2_3	p-val	χ^2_3	p-val	χ^2_3	p-val
	0.27	0.97	16.0	<0.01	0.81	0.85
	Zinc		**Silver**			
	Coef	SE	Coef	SE		
OECD Growth	0.0044	0.0079	−0.0046	0.0061		
EMERGE Growth	−0.0092	0.0072	0.0105	0.0070		
U.S./U.K. Unemployment	−0.0181	0.0063**	−0.0220	0.0100**		
R^2	0.26	(0.23)	0.91	(0.90)		
Test of Joint Significance	χ^2_3	p-val	χ^2_3	p-val		
	10.7	0.01	6.9	0.07		

Standard errors and tests of joint significance use the Newey-West covariance matrix estimator; **, * indicate statistical significance at 5% and 10% levels, respectively. R^2 values in parentheses correspond to the pure trend model.

data period, none of the parametric effects are significant. For 1993–2014, U.S. growth is a strongly significant driver of continental natural gas prices. Note that in this case, we include the OPEC market share which is moderately significant (an OPEC dummy would not be identified).

We estimate the model for U.S. coal prices from the beginning of the 20th century, 1901–2014 and for the post-war period, 1946–2014. U.S. economic growth is not a significant factor in explaining the price series over either period. Over the post-war period, unemployment is marginally significant with a negative impact on prices. The OPEC effect is large and strongly statistically significant, increasing

https://doi.org/10.1017/S1365100516000511 Published online by Cambridge University Press

TABLE 3A. Hydrocarbons price correlations

1919–2014	Oil	Gas	Coal
Oil	1.00	0.85	0.81
Gas	0.85	1.00	0.68
Coal	0.81	0.68	1.00

coal prices in excess of 25% on average relative to trend. The "Regulatory Dummy" which is one after 1970, is intended to capture the impacts of coal mine safety legislation passed in 1969 and the Clean Air Act Amendments passed in 1970, which established the U.S. Environmental Protection Agency (EPA). The net effect is positive and strongly statistically significant, increasing coal prices by an estimated 9% (using the full data) and 15% (using post-war data) relative to trend.

The parametric variables—OECD and emerging economy growth rates, and the U.S./U.K. unemployment rate—are jointly significant in 5 out of 8 metals. For this group of models, there is substantial variation in the data-driven smoothing parameter, leading to considerable variation in the goodness-of-fit across models. However, as with hydrocarbons, the inclusion of macro variables increases R^2 modestly, with a somewhat larger increase in the lead equation.

OECD growth rates are also significant in the lead equation and the unemployment rate is significant in explaining zinc, iron, tin, silver, and lead prices. Growth in emerging economies is statistically insignificant in all metals models.

For all 11 commodities, the addition of macroeconomic variables results in modest improvement in R^2 (see Tables 2A and 2B where the bracketed value is obtained from the pure nonparametric regressions depicted in Figure 1). This, however, should not be interpreted as diminishing the role of macroeconomic effects within these models as the pure nonparametric model is indirectly picking up the omitted macro effects.

4.3. Shape Similarity

Table 3A summarizes price correlations among hydrocarbons beginning in 1919 (the first year for which natural gas price data are available). Not surprisingly, the historical price correlations between oil and both gas and coal are high, exceeding 0.8. The correlation between gas and coal is also strong at 0.68.

Table 3B summarizes price correlations among the 8 metals for the period 1901–2014. Copper and nickel exhibit a strong correlation of 0.66. Tin exhibits a 0.65 correlation with silver; iron and tin correlation is 0.61. For nickel and silver it is 0.60.

Based on these correlations, we consider more formally whether certain commodities bear shape similarity in their trends. Figure 2 illustrates data for pairs of commodities. In each case, the estimated parametric effects have been removed

https://doi.org/10.1017/S1365100516000511 Published online by Cambridge University Press

https://doi.org/10.1017/S1365100516000511 Published online by Cambridge University Press

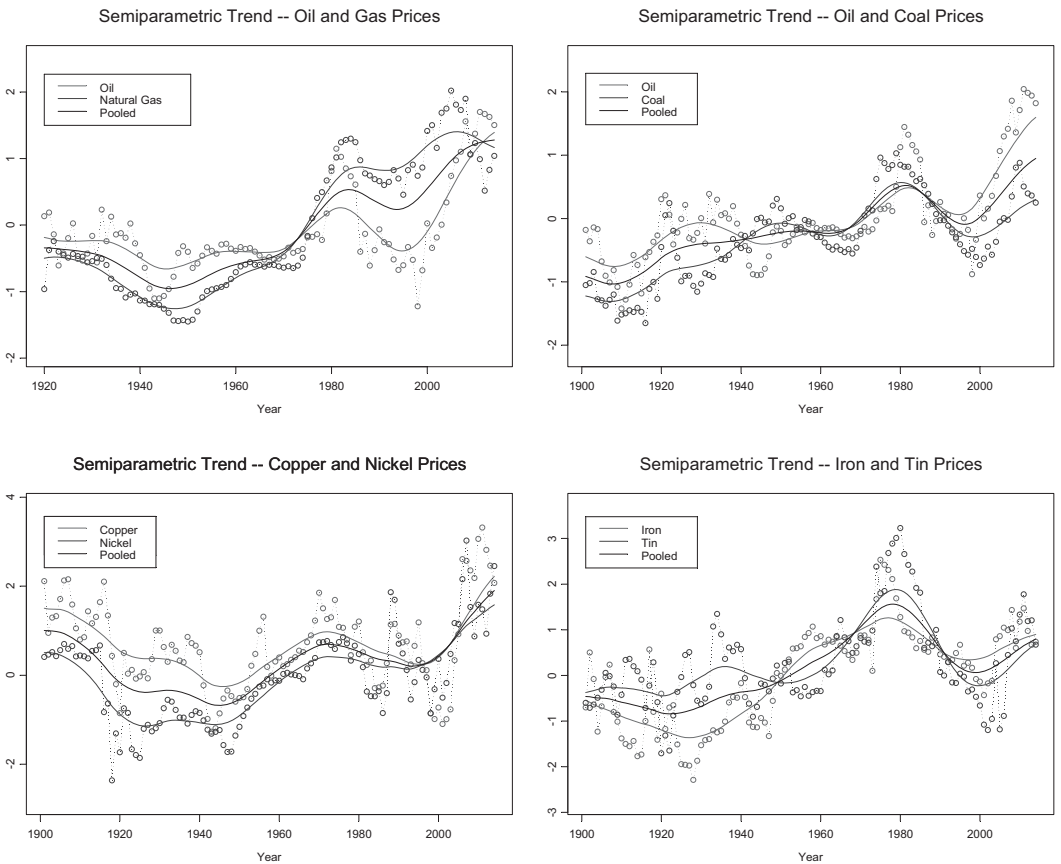


FIGURE 2. Shape similarity.

TABLE 3B. Metals price correlations

1901–2014	Copper	Nickel	Zinc	Iron	Tin	Silver	Lead	Aluminum
Copper	1.00	0.66	0.36	0.30	0.28	0.52	0.56	0.18
Nickel	0.66	1.00	0.43	0.51	0.33	0.60	0.24	−0.15
Zinc	0.36	0.43	1.00	0.24	0.29	0.24	0.33	−0.06
Iron	0.30	0.51	0.24	1.00	0.61	0.47	0.16	−0.52
Tin	0.28	0.33	0.29	0.61	1.00	0.65	0.21	−0.33
Silver	0.52	0.60	0.24	0.47	0.65	1.00	0.18	−0.13
Lead	0.56	0.24	0.33	0.16	0.21	0.18	1.00	0.00
Aluminum	0.18	−0.15	−0.06	−0.52	−0.33	−0.13	0.00	1.00

and the resulting data have been standardized by subtracting the mean and dividing by the standard deviation. More formally, let $y_A = (y_{A1}, ..., y_{AT})^\intercal$ and $z_A = (z_{A1}, ..., z_{AT})^\intercal$, $y_A^* = [y_A - z_A\hat{\beta}_A - \overline{(y_A - z_A\hat{\beta}_A)}]/s_{y_A - z_A\hat{\beta}_A}$ and y_B^* in a similar fashion.[12] Let $\hat{f}_A$ and $\hat{f}_B$ be the nonparametric estimates of the individual trend effects applied to these standardized data (in this case, we have applied a simple kernel smoother). Now, concatenate the y_B^* and y_B^* vectors and regress on t to obtain a pooled estimate of the trend effect, say $\hat{f}_{pooled}$.

The upper left panel of Figure 2 displays the results for oil and gas. Not surprisingly, the individual estimates display a high degree of shape similarity, mainly because of rising hydrocarbon prices in the 1970s, receding prices in the 1980s and then again rising prices in the 21st century. In the most recent years, beginning around 2009, the individual price trends diverge. During this period, U.S. natural gas prices plummet as a result of the fracking revolution, while oil prices remain high until late 2014, at which time they also plunge. The upper right hand panel of Figure 2 illustrates trends in oil and coal prices. There are some differences in timing, but the two individual estimates follow similar paths. The bottom left and right panels of Figure 2 display standardized trend curves for copper and nickel, and iron and tin. These metals prices exhibit considerable volatility, but the estimated trend curves for each respective pair show a degree of shape similarity.

5. CONCLUSIONS

This paper proposes models for commodity prices where trend effects are incorporated nonparametrically and macroeconomic variables enter parametrically. We seek to identify the trend component using data-driven techniques; in estimating the smoothing parameter using cross-validation, we exclude 5-year data windows to mitigate the risk that serially correlated residuals will lead to over-fitting. Relative to the pure nonparametric models, the introduction of macroeconomic variables improves fit modestly.

https://doi.org/10.1017/S1365100516000511 Published online by Cambridge University Press

The change in strategic behavior by OPEC in 1973 initiated a new era during which the market was dominated by OPEC, and other participants made strategic choices in response (for example, developing new supplies at costs which significantly exceeded costs of production in the Arabian Peninsula). The structural break is arguably attributable to OPEC, but the value of the coefficient—which indicates a price increase in excess of 50%—is a consequence of complex strategic interactions. In the absence of mitigating responses by suppliers, prices could very well have been higher (for example during the period 1986–2000).

The OPEC impact on coal markets is also strongly statistically significant, increasing coal prices by an estimated 25% on average. U.S. legislation regulating coal mine health and safety contributes to a 9% to 15% increase in the price of coal.

Quadratic specifications of price trends are rejected for all commodities except zinc. Band pass estimates of long term trend (which consider cycles of 70 years or more) are rejected for all commodities except zinc and iron. Richer band pass models which incorporate "super cycles" lead to results that in many cases track our nonparametric estimates closely, the difference being that our models are founded on data based identification of the time effect.

Some commodity pairs exhibit a significant degree of shape similarity, for example, oil and natural gas, and oil and coal. Given the potential for substitution, this is not surprising. Copper and nickel, and tin and iron also exhibit degrees of shape similarity. These results suggest that joint semiparametric estimation with common trends across equations may lead to further fruitful results. We leave this strand for future research.

Shape similarity can also inform the predictive process. Consider the hydrocarbon triad. From 1946 to 2008 the correlations of the (log) price of oil with gas and coal are 97% and 98% respectively. Then gas and coal prices turn down (see Figure 1). In this context, the recent drop in oil prices is perhaps not entirely surprising.

NOTES

1. There is rich literature which considers structural models, incorporating demand and supply side drivers, and investigating the relationship between oil prices in particular and macroeconomic performance. Such models often rely on data over shorter time spans and higher frequencies. See for example, Baumeister and Kilian (2015), Hamilton (2009), and references therein. Pindyck and Rotemberg (1990) study the co-movement of changes in commodity prices. Another important strand of the literature estimates stochastic trend models, possibly with unit roots, as well as their cointegrated relatives. There is also a literature which tests deterministic against stochastic trends [see Ghoshray (2011)], without, as yet, a definitive conclusion.

2. Some papers consider whether firm behavior is consistent with Hotelling's Rule. See e.g. Young (1992).

3. For this paper, estimation and testing is conducted in R/S-Plus which incorporate a range of nonparametric estimation routines such as ''ksmooth'', ''lo'', and ''s'' corresponding to kernel, loess and spline smoothing respectively. We also make use of ''gam'' which estimates the generalized additive model using back-fitting.

https://doi.org/10.1017/S1365100516000511 Published online by Cambridge University Press

4. Results based on the omission of 3 or 7 observations were similar.

5. This version is for the box kernel. More general kernels may of course be implemented. Diagonal entries are zero as this vastly simplifies the derivation of the large sample distribution of the statistic.

6. See e.g. Fan and Li (1996) and Li (1999). For the rationale and intuition underlying this statistic, [see Yatchew (2003), pp. 115–125]. The essential idea is that if the modeling of trend in the restricted regression is correct, then the resulting residuals should not retain any significant time effect.

7. Note that Manthy (1978) updates the work done by Potter and Christy.

8. Note that the BLS Wholesale Price Index was renamed the Producer Price Index after 1978, but the name change was not accompanied by any alteration to the index methodology.

9. For all countries, GDP is measured in terms of constant 1990 Geary-Khamis "international" dollars.

10. We use the generalized additive model (`gam` function) in S-Plus with a smoothing spline (`s` function).

11. For example, as a result of the oil price shocks of the 1970's there was a massive shift from oil to coal in electricity generation. Recent declines in North American natural gas prices, combined with environmental constraints, have led to increased substitution of natural gas for coal in electricity generation.

12. The notation $^{\intercal}$ denotes transpose; the overbar denotes the mean.

REFERENCES

Ahrens, W. Ashley and Vijay R. Sharma (1997) Trends in natural resource commodity prices: Deterministic or stochastic? *Journal of Environmental Economics and Management* 33(1), 59–74.

American Metal Market (1979) *Metal Statistics, 1979: Technical Report*. New York: Fairchild.

Aneiros-Pérez, Germán and Alejandro Quintela-del-Río (2001) Modified cross-validation in semiparametric regression models with dependent variables. *Communications in Statistics - Theory and Methods* 30(2), 289–307.

Barnett, H. J. and C. Morse (1963) *Scarcity and Growth: The Economics of Natural Resource Availability*. Baltimore: Johns Hopkins University Press.

Berck, Peter and Michael Roberts (1996) Natural resource prices: Will they ever turn up? *Journal of Environmental Economics and Management* 31(1), 65–78.

Christiano, Lawrence and Terry Fitzgerald (2003) The band pass filter. *International Economic Review* 44(2), 435–465.

Cuddington, T. John and Daniel Jerrett (2008) Super cycles in metals prices? *IMF Staff Papers* 55(4), 541–565.

Dasgupta, P., G. Heal, and J. Stiglitz (1980) The taxation of exhaustible resources. In G. Hughes and G. Heal (eds.), *Public Policy and the Tax System*, pp. 150–172. London: George Allen Unwin.

Denman, James and Paul McDonald (1996) Unemployment statistics from 1881 to the present day. *Labour Market Trends* 104(1), 5–18.

Erten, Bilge and José Antonio Ocampo (2013) Super cycles of commodity prices since the mid-nineteenth century. *World Development* 44, 14–30.

Fan, Yanqin and Qi Li (1996) Consistent model specification tests: Omitted variables and semiparametric functional forms. *Econometrica* 64(4), 865–890.

Gao, Jiti and Kim Hawthorne (2006) Semiparametric estimation and testing of the trend of temperature series. *Econometrics Journal* 9(2), 332–355.

Gaudet, Gérard (2007) Natural resource economics under the rule of Hotelling. *Canadian Journal of Economics* 40(4), 1033–1059.

Ghoshray, Atanu (2011) A reexamination of trends in primary commodity prices. *Journal of Development Economics* 95(2), 242–251.

Gilbert, J. Richard (1978) Dominant firm pricing policy in a market for an exhaustible resource. *Bell Journal of Economics* 9(2) 385–395.

https://doi.org/10.1017/S1365100516000511 Published online by Cambridge University Press

Gilbert, J. Richard (1979) Optimal depletion of an uncertain stock. *Review of Economic Studies* 46(1), 47–57.

Hamilton, D. James (2009) Understanding crude oil prices. *The Energy Journal* 30(2), 179–206.

Heal, Geoffrey and Michael Barrow (1980) The relationship between interest rates and metal price movements. *Review of Economic Studies* 47(1), 161–181.

Hotelling, Harold (1931) The economics of exhaustible resources. *Journal of Political Economy* 39(2), 137–175.

Jacks, D. S. (2013) From Boom to Bust: A Typology of Real Commodity Prices in the Long Run. Working paper No. w18874, National Bureau of Economic Research.

Lebergott, S. (1957) Annual estimates of unemployment in the United States, 1900–1954. In Universities-National Bureau (ed.), *The Measurement and Behavior of Unemployment*, pp. 211–242. Cambridge, MA: National Bureau of Economic Research.

Lee, Junsoo, John A. List, and Mark C. Strazicich (2006) Non-renewable resource prices: Deterministic or stochastic trends? *Journal of Environmental Economics and Management* 51(3), 354–370.

Li, Qi (1999) Consistent model specification tests for time series econometric models. *Journal of Econometrics* 92(1), 101–147.

Livernois, John (2009) On the empirical significance of the Hotelling rule. *Review of Environmental Economics and Policy* 3(1), 22–41.

Livernois, R. John and Russell S. Uhler (1987) Extraction costs and the economics of nonrenewable resources. *Journal of Political Economy* 95(1), 195–203.

Maddison Project (2013) Statistics on World Population, GDP and Per Capita GDP 1-2010 [Data file]. Retrieved from http://www.ggdc.net/maddison/maddison-project/data.htm.

Manthy, R. S. (1978) *Natural Resource Commodities: A Century of Statistics*. Baltimore: Johns Hopkins University Press.

Moazzami, Bakhtiar and F. J. Anderson (1994) Modelling natural resource scarcity using the 'error-correction' approach. *Canadian Journal of Economics* 27(4), 801–812.

Newey, K. Whitney and Kenneth D. West (1987) A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55(3), 703–708.

Perron, Pierre (1989) The great crash, the oil price shock, and the unit root hypothesis. *Econometrica* 57(6), 1361–1401.

Pindyck, S. Robert (1978) The optimal exploration and production of nonrenewable resources. *Journal of Political Economy* 86(5), 841–861.

Pindyck, S. Robert (1980) Uncertainty and exhaustible resource markets. *Journal of Political Economy* 88(6), 1203–1225.

Pindyck, S. Robert (1999) The long-run evolution of energy prices. *Energy Journal* 20(2), 1–27.

Pindyck, S. Robert and Julio J. Rotemberg (1990) The excess co-movement of commodity prices. *The Economic Journal* 100(403), 1173–1189.

Potter, N. and F. T. Christy (1962) *Trends in Natural Resource Commodities: Statistics of Prices, Output, Consumption, Foreign Trade, and Employment in the United States*, 1870–1957. Washington: Resources for the Future.

Robinson, M. Peter (1988) Root-N-Consistent semiparametric regression. *Econometrica* 56(4), 931–954.

Salant, W. Stephen. (1982) Imperfect competition in the international energy market: A computerized nash-cournot model. *Operations Research* 30(2), 252–280.

Schurr, S. H. (1960) *Historical Statistics of Minerals in the United States*. Washington: Resources for the Future.

Slade, E. Margaret (1982) Trends in natural-resource commodity prices: An analysis of the time domain. *Journal of Environmental Economics and Management* 9(2), 122–137.

Slade, E. Margaret and Henry Thille (2009) Whither hotelling: Tests of the theory of exhaustible resources. *Annual Review of Resource Economics* 1(1), 239–260.

Smith, V. Kerry (1978) Measuring natural resource scarcity: Theory and practice. *Journal of Environmental Economics and Management* 5(2), 150–171.

https://doi.org/10.1017/S1365100516000511 Published online by Cambridge University Press

Smith, V. Kerry (1979) Natural resource scarcity: A statistical analysis. *The Review of Economics and Statistics* 61(3), 423–427.
Smith, V. Kerry (1981) The empirical relevance of Hotelling's model for natural resources. *Resources and Energy* 3(2), 105–117.
Solow, M. Robert and Frederic Y. Wan (1976) Extraction costs in the theory of exhaustible resources. *Bell Journal of Economics* 7(2), 359–370.
Stiglitz, E. Joseph (1974) Growth with exhaustible natural resources: Efficient and optimal growth paths. *The Review of Economic Studies* 41(5), 123–137.
Sweeney, L. James (1977) Economics of depletable resources: Market forces and intertemporal bias. *Review of Economic Studies* 44(1), 125–141.
Swierzbinski, E. Joseph and Robert O. Mendelsohn (1989) Exploration and exhaustible resources: The microfoundations of aggregate models. *International Economic Review* 30(1), 175–186.
U.K. Office for National Statistics (2014) *Labor Market Statistics* [Database]. Retrieved from http://www.ons.gov.uk/ons/rel/lms/labour-market-statistics/index.html.
U.S. Bureau of the Census (1975) *Historical Statistics of the United States, Colonial Times to 1970 (No. 93)*. US Department of Commerce, Bureau of the Census.
U.S. Bureau of Labor Statistics (2014a) *Producer Price Index-Commodities* [Data File]. Retrieved from http://data.bls.gov/timeseries/WPU00000000.
U.S. Bureau of Labor Statistics (2014b) *Producer Price Index-Commodities, Steel Mill Products* [Data File]. Retrieved from http://data.bls.gov/timeseries/WPU1017.
U.S. Bureau of Labor Statistics (2014c) *Unemployment Rate* [Data File]. Retrieved from http://data.bls.gov/timeseries/LNS14000000.
U.S. Energy Information Administration (2011) *Annual Energy Review 2011. Table 7.9 Coal Prices, 1949–2011* [Data File]. Retrieved from http://www.eia.gov/totalenergy/data/annual/.
U.S. Energy Information Administration (2013) *U.S. Natural Gas Wellhead Price* [Data file]. Retrieved from http://tonto.eia.gov/dnav/ng/hist/n9190us3a.htm.
U.S. Energy Information Administration (2014a) *Annual Energy Outlook 2014. Coal Supply, Disposition, and Prices.* [Data File]. Retrieved from http://www.eia.gov/oiafaeo/tablebrowser/.
U.S. Energy Information Administration (2014b) *Henry Hub Natural Gas Spot Price* [Data file]. Retrieved from http://tonto.eia.gov/dnav/ng/hist/rngwhhdm.htm.
U.S. Energy Information Administration (2014c) *U.S. Refiner Acquisition Cost of Crude Oil* [Data file]. Retrieved from http://www.eia.gov/dnav/pet/pet_pri_rac2_dcu_nus_m.htm.
U.S. Geological Survey (1999) *Metal Prices in the United States Through 1998*. US government printing office.
U.S. Geological Survey (2013a) *Minerals Yearbook* [Database]. Retrieved from http://minerals.usgs.gov/minerals/pubs/commodity/myb/.
U.S. Geological Survey (2013b) *Historical Statistics for Mineral and Material Commodities in the United States. U.S. Geological Survey Data Series 140* [Database]. Retrieved from http://minerals.usgs.gov/minerals/pubs/historical-statistics/.
World Bank (2014a) *World Bank Cross Country Database. World Indicators: London Metal Exchange* [Database]. Retrieved from https://www.quandl.com/WORLDBANK?keyword=LME.
World Bank (2014b) *World Development Indicators Databank. GDP Growth* [Data File]. Retrieved from http://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG.
Yatchew, A. (2003) *Semiparametric Regression for the Applied Econometrician*. Cambridge: Cambridge University Press.
Young, Denise (1992) Cost specification and firm behaviour in a hotelling model of resource extraction. *Canadian Journal of Economics* XXV, 41–59.

https://doi.org/10.1017/S1365100516000511 Published online by Cambridge University Press