

Markets in ancient societies?

the structural analysis of Babylonian price data

Péter Földvári and Bas van Leeuwen

Abstract

Over the past decades, there has been a fierce discussion about the existence of markets in ancient economies. Especially Moses Finley has argued that it is a mistake to project market mechanisms of our day back to the past. However, recently his primitivist approach has given way to the recognition that also past economies exhibit clear market behaviour. Yet, even though this recognition has gained ground, the methods employed to test this hypothesis are still dominated by the Finleyan, non-quantitative, methods. In this paper we test whether working markets existed using both economic theory and methods. The Babylonian price data used for the analysis, although very abundant, contain a lot of missing observations. Therefore, applying several missing data techniques, we find that there is clear evidence for a weak form of market efficiency. Furthermore, we find evidence for demand and supply relations in Babylon.

1. Introduction

From the 1900s onwards, economic historians tended to view economic development as “modern” during and after the industrial revolution and as largely Malthusian before. The differences between the two kinds of societies can be defined in several ways. The demographic difference is probably the most apparent: in a traditional society both fertility and mortality are related to per capita income, while modernization gradually changes this relationship. The difference in the distributional mechanisms of these societies has been subject of a debate for about five decades now. Modernists believe that markets played important role even in the ancient societies, while

primitivists, among whom Finley is most prominent, believe that it is a mistake to project market mechanisms of our days back into the past, and we should rather look on the ancient economies as mainly consisting of small economic units engaged in sustenance agriculture (Andreau 2002, 33-49).¹

For more recent periods, like the early modern ages and the middle ages, it has been shown that that modern economic growth had probably started before the industrial revolution took off (De Vries and Van der Woude 1997; Van Zanden and Van Leeuwen (2011); Broadberry et al. (2011)), and that market mechanisms were major source of the distribution of goods and services already in the middle ages (Britnell and Campbell 1995). Temin (2002), followed by Romero et al. (2010) reached a similar conclusion regarding Babylonian prices for the period BC 472-62.

Yet, although the primitivist point of view seems untenable nowadays, this has not led to the implementation of modern economic models to ancient economies on a large scale since, as argued by Andreau (2002, 34), “[t]he questions are Finleyan, the methods and ways of thinking bear the stamp of his influence, but the answers are moving farther and farther away [...]”. Indeed, although the available Babylonian price series have been analyzed in several studies, most of them relied heavily on different statistical methods (mostly descriptive statistics or smoothing and trend fitting). The main exponent of this is Slotsky (1997) whose work is further criticized for not taking proper care of the historical background (Van der Spek and Mandemakers 2003) and contains basically no economic analysis. Formal models, however, are rarely used in the economic history of ancient societies. One notable exception (when statistics is driven by economic theory) is Temin (2002) who focuses on the question whether the observed prices in ancient Babylonia were governed by market mechanisms and not by some regulations resulting from a central redistribution of goods. He analyzed the prices from three perspectives.

¹ The distinction primitivism-modernism (Finley vs Rostovtzeff) can also be viewed in terms of substantivism-formalism. The formalists apply economic theories to ancient economies based on the idea that profit maximization was also applicable in ancient economies, a famous example being John Hicks. Substantivists, like Finley, following Karl Polanyi, embrace cross-cultural comparisons, claiming that social relations were more important than profit.

Firstly, he checks if the price series exhibit autocorrelation. Based on a first-order autoregressive model, he concludes that the data are either random-walk or very close to it. If the Babylonian prices indeed prove to be random-walk processes, then we can conclude that agents were not able to systematically predict the change of prices (and achieve a profit by doing so) based on prices of the past, which fulfills the condition of the weak form of the efficient market hypothesis. Secondly, he attempts to justify that there is a long-term trend present in the data, which is based on his assumption that market prices are subject to inflation or deflation. Similarly to Slotsky (1997) he finds that a third degree polynomial trend fits the data relatively well. Thirdly he assesses whether political events caused changes in price and finds that for example the death of Alexander the Great affected prices in a statistically significant way. Over-all, he finds enough evidence to conclude that the prices in Ancient Babylon were governed by market mechanisms.

Temin's method, however useful, was indirect since he neither accounted for missing observations nor did he apply a formalized unit-root test. In this paper we wish to extend the analysis of the Babylonian prices in two ways. First, the available price series contain a lot of missing observations, but the possible effect of this on statistical inference has not yet been explored. Afterwards we apply three formal tests, based on economic theory, to find out more about the market mechanisms in Babylon. This is done in Section 3 by using a unit root test to see if the Babylonian price series are indeed random-walk processes, suggesting a weak form of efficient market hypotheses applicable. In Sections 4 and 5 we analyze the price series within a simple partial equilibrium model of agricultural prices. With only information on prices, one can only estimate a reduced-form equation but the supply and demand equations themselves cannot be identified. However, even with such limitation, we can formally test the effect of external factors like political events and seasonality which are also indicative for the presence of a working market mechanism. We end with a brief conclusion.

2. The problem of missing data

The data for Babylon are based on the Astronomical Diaries (ADs). These diaries are best described as ‘record(s) of observed phenomena carefully chosen from the realms of the celestial, the atmospheric, and the terrestrial,’ (Hunger and Pingree 1999, 141).² These diaries contain a notation of celestial phenomena followed (in an increasing degree over time) by information on other ominous events like monstrous births, direction of the wind, the weather, the level of the Euphrates, but also more mundane things like deeds of kings, important events in Babylon and the level of the prices of six commodities: barley, dates, cuscuta³, water cress, sesame and wool. Typically, one such diary covers the period of half a Babylonian year, i.e. six lunar months.

The very fact that the prices are recorded in the astronomical diaries proves that they were regarded as volatile and unpredictable. One might question the reliability of the price data but the very fact that the additional information of the diaries on checkable events (e.g. on the orbit of the planets and the death of kings) is correct, gives an indication that the information on the prices is also basically correct. The data was recorded by the astronomers simply by visiting the markets of Babylon and observing what they could buy for one shekel of silver. They did not check all shops, though, and hence when we have duplicated observations we find price differences, but these turn out to be very small.

More worrisome is the large number of missing observations in the prices. On an estimated total duration of the diaries of 600 years, this means a survival rate of ca. 5.5%. However, survival rates are still relatively small. Fortunately, there is a quite large literature on the statistical consequences and treatment of missing data (Allison 2001, Little and Rubin 2002, McKnight 2007). Based on this literature, we can conclude that even though there are numerous methods to take care of the problem, one has quite limited options. First of all, one needs to establish if the

²The name ‘Diary’ was coined by Abraham Sachs on the basis of the colophon-title *našāru* (EN.NUN) *šá ginê*, ‘regular observation’.

³There is actually a discussion what this product is. Most authors nowadays argue that it is cuscuta, but some others have argued for mustard. In sum, there is no agreement in the literature.

data is missing completely at random (MCAR), missing at random (MAR) or missing not at random (MNAR). The first concept means that the probability that an observation is missing is independent of the value of the variable, and of any other variables. In order to have the data to be MAR, it is required that the probability that an observation is missing should not be related to the value of the variable once we control for other variables. In the first case, listwise deletion (simply deleting the missing observations) will not lead to a bias. If the data is missing not at random (MNAR), one needs to use some more sophisticated techniques, and model the missingness of the data itself.

At first sight it can be argued that the occurrence of missing observations in the data are not random, there is clear pattern of serial correlation: this can be easily shown with a Wald-

Table 1
Results from the runs test

	barley	Cress	dates	cuscuta	sesame	wool
Number of missing observations	3557	3733	3603	3764	3704	3757
number of available observations	535	359	489	328	388	335
p-value of the test	0.000	0.000	0.000	0.000	0.000	0.000

Note: the null hypothesis is that the series are not serially correlated

Wolfowitz or runs-test:⁴ As the test results suggest in Table 1, we must reject the idea that the missingness of the data is completely random: it seems that the tablets containing the data survived in groups as indicated by the serial correlation. In other words: a tablet is more likely to have survived if the previous table also survived. This is partly caused, of course, because several periods with prices can be found on one tablet. However, such relationship may also have existed if the tablets from successive periods were stored close to each other. In case of any event, the observations for successive periods were more likely to perish. A simple linear probability model

⁴ One should bear in mind that the runs-test is by no means a test of any of the above mentioned hypotheses (MCAR, MAR, MNAR). The runs-test null hypothesis is simply that the series occur in a purely random manner. But even if that is the case, theoretically, they could be correlated with any other variables we use in the further analysis making them MAR or MNAR.

confirms this finding: for all our series we find that the probability of the availability of an observation is increased by 50-60% if the previous period's observation was not missing either.

Still, the missing data becomes a real problem only if any of the variables we use in the analysis is correlated with the probability of an observation missing (that is with the serial correlation we found the data can still be MCAR). In other words, it is only a problem if we want to test whether a high river level causes famines, but we only have price observations when the river level is high. In order to test this, we carry out the Little's MCAR (Little 1988) test on our series to find out if we can reject the null hypothesis that they are MCAR. The test statistics is 294.664, the 1% critical value of the chi-square distribution with 175 degrees of freedom is 221.438. Hence we can confirm with high probability that Babylonian data is not MCAR. Unfortunately, the MAR hypothesis is basically non-testable.

There are several methods to address the missing data problem: the simplest is the list wise deletion that is simply removing the observations where there is a missing data. The second option is pairwise deletion, when all available observations are made use of: the means, standard deviation and the covariance of the series is estimated from all available data and these statistics are used to arrive at the parameters of the regression analysis. With a Monte Carlo simulation we can illustrate that in case of data generated to be very similar to the price data that listwise deletion leads to higher standard error estimates and therefore lower t-statistics (Appendix B). We also experiment with two data augmentation (imputation) techniques: mean imputation means that the missing observations are replaced by the sample mean of the respective series, while multiple imputation relies on a Markov chain Monte Carlo process to impute the missing data, which process is carried out five times. The pooled estimates (mean of the five successive estimations) have good properties even when the covariances and variances are estimated (Schaffer,1997).⁵

⁵ The missing observations in the dependent variable can also be treated by the Heckman method (Heckman, 1979). In an earlier draft of this paper we have carried out regression with that method as well, but since it has not led to

These methods will be applied to analyze the effect of missing data on the tests for the presence of market mechanism. In the next section, we start with a unit root test for the presence of a random walk, which, when existing, proves the existence of the weakest form of an efficient market hypothesis since it indicates that prices are non-predictable. The final two sections discuss a structural model which not only tests for the presence of a demand and supply mechanism, but also allows us to distinguish external effects on prices and the presence of seasonality, all of which indicate the presence of a working market.

3. Unit-root tests

With unit root tests we formally test if the logarithms of the price series are random walk processes. Since they usually have a trend, we assume that they are random-walk processes, generally described for product i as:

$$\ln p_{i,t} = \ln p_{i,t-1} + \varepsilon_{i,t} \quad (1.)$$

where we assume that the error-term ε is a white noise process, independently and identically distributed with zero mean. An important property of random walk series is that their change is completely random, as can be seen if take the first difference of equation (1):

$$\Delta \ln p_{i,t} = \varepsilon_{i,t} \quad (2.)$$

We apply two types of formal tests: the Dickey-Fuller test and the KPSS test (Kwiatkowski et al 1992). Since we have a lot of missing observations, we do not include lags of the dependent variable in the Dickey-Fuller test (that would make it an Augmented Dickey-Fuller test). We also use the KPSS test. Since the two tests have different null-hypothesis (DF – nonstationarity, KPSS-stationarity) their simultaneous use can be advantageous.

significant changes in the coefficients (and the coefficient of the inverse Mills ratio remained insignificant) we decided to omit it.

Table 3
Unit root tests on the level of the log of the price series (p-values in parentheses)

	barley	cross	dates	cuscuta	sesame	wool
D-F	-2.931 (0.043)	-2.143 (0.228)	-1.817 (0.371)	-0.141 (0.942)	-2.758 (0.066)	-0.681 (0.848)
KPSS	1.027 (<0.01)	0.813 (<0.01)	1.237 (<0.01)	1.453 (<0.01)	1.449 (<0.01)	2.671 (<0.01)

Note: <0.01 means that the null hypothesis can be rejected at 1% level of significance,
Null-hypotheses: D-F: nonstationarity, KPSS: stationarity

Table 4
Unit root tests on the level of the first difference of the log of the price series (p-values in parentheses)

	barley	cross	Dates	cuscuta	sesame	wool
D-F	-15.80 (0.000)	-8.483 (0.000)	-12.86 (0.000)	-7.413 (0.000)	-12.93 (0.000)	-9.222 (0.000)
KPSS	0.057 (>0.1)	0.047 (>0.1)	0.052 (>0.1)	0.402 (>0.05)	0.104 (>0.1)	0.163 (>0.05)

Note: >0.1 (>0.05) means that the null hypothesis cannot be rejected at 10% (5%) level of significance,
Null-hypotheses: D-F: nonstationarity, KPSS: stationarity

The D-F and the KPSS tests indicate that the all price series are I(1), that is they are random-walk processes. From an economic point of view, this means that no agents could have possibly predicted the change in prices, and the rational expectation regarding prices is the primitive expectation, that is, that prices do not change from one period to the other.

Using the expectation operator this can be expressed as follows:

$$E_{t-1} \ln p_{i,t} = \ln p_{i,t} \quad (3.)$$

where $E_{t-1} \ln p_{i,t}$ denotes the expectation in period t-1 regarding the price in the following period. That these expectations are rational are easy to see, since the deviations from the expected price (the forecasting error) has a zero expected value. This is indicative of efficient markets since, if people could predict price change in the future, they could make a profit at low risk. Still, it may be questioned whether with such high number of missing observations we still can trust the unit-root test results. In order to decide this question, we have carried out a Monte Carlo simulation (see Appendix A). The results suggest that we can rely on the unit root tests but we should rather prefer the KPSS.

4. Two partial equilibrium models of agricultural prices

Another way to study the existence of markets is the existence of a working demand and supply system, which implies also that political and other factors may affect prices as well as the existence of seasonality in prices. However, such a structural equation needs to rely on a sound theoretical basis. In the followings we introduce two partial equilibrium models for the agricultural prices: the first, short model, is simplified, but finally results in a specification that can be estimated with the available data. The second model is superior theoretically, but because lagged values of the observations are involved, results in a smaller sample.

a. a short model

Let us assume that the supply and demand of agricultural goods in Ancient Babylon can be described with two equations:

$$\ln Q_{i,t}^S = \alpha_0 + u_t \quad \text{Supply (4.)}$$

$$\ln Q_{i,t}^D = \beta_0 + \gamma \ln P_{i,t} + \sum_{k=1}^{n_p} \beta_k \ln P_{k,t} + \omega Y_t + v_{i,t} \quad \text{Demand (5.)}$$

$$\ln Q_{i,t}^S = \ln Q_{i,t}^D \quad \text{Condition of equilibrium in market } i$$

Where Q_i denotes the quantity of i supplied or demanded (superscripts S and D), $u_{i,t}$ denotes the supply shocks (locusts, natural disasters, destructions of war, weather conditions). In equation (4) we assume that the supply is not affected by price, which is an obvious simplification but we will show that this version can be estimated with the available data. In the demand equation P_i denotes the price of product i , while P_j is the price of product j (there are n_p number of other goods) that is either substitutes or complement of product i . Y denotes aggregate income, and $v_{i,t}$ is a random variable capturing the effect of demand shocks (e.g. king or satrap visiting Babylon, wars, plagues). The aggregate income is not observed, so we assume that most of its effects are captured

by a time-trend. The latter is included because it is believed that Babylon underwent a loss of its importance and some of its population in this period. If we accept that the per capita income remained roughly the same in the period (as it can be derived from the basic Malthusian model), lower population means lower aggregate income. So the trend should generally yield a negative coefficient in case of normal goods, and a positive coefficient in case of inferior goods.

Since we have no data on quantities, it is not possible to identify both equations. What one can estimate with knowing prices only is a reduced form equation for the price of product i .

$$\ln P_{i,t} = \frac{\alpha_0 - \beta_0}{\gamma} - \sum_{k=1}^{n_p} \frac{\beta_k}{\gamma} \ln P_{k,t} - \frac{\omega}{\gamma} Y_t + \frac{u_t - v_t}{\gamma} \quad (6.)$$

Where we can assume that $\gamma < 0$, that is the quantity of product i demanded is negatively related to price of product i 's price that is the signs of the coefficients in the reduced equation will have the opposite sign than in the supply equation (4.). Since γ is unknown, however, we cannot determine the coefficients of the original demand and supply equations. The system of equations is unidentified.

b. a more complex model

One may argue that the previous model is far too simple, and expectations regarding future prices should play a role in the market mechanisms. We can easily rewrite the model to have expectations included:

$$\ln Q_{i,t}^S = \alpha_0 + \alpha_1 \ln(E_{t-1} P_{i,t}) + \sum_{k=1}^{n_p} \alpha_{k+1} \ln(E_{t-1} P_{k,t}) + u_t \text{ Supply} \quad (7.)$$

$$\ln Q_{i,t}^D = \beta_0 + \gamma \ln P_{i,t} + \sum_{j=1}^{n_p} \beta_j \ln P_{j,t} + \omega Y_t + v_t \quad \text{Demand} \quad (8.)$$

$$\ln Q_{i,t}^S = \ln Q_{i,t}^D \text{ Condition of equilibrium in market } i$$

In equilibrium:

$$\ln P_{i,t} = \frac{\alpha_0 - \beta_0}{\gamma} - \sum_{j=1}^{n_p} \frac{\beta_j}{\gamma} \ln P_{j,t} - \frac{\omega}{\gamma} Y_t + \frac{\alpha_1}{\gamma} \ln(E_{t-1} P_{i,t}) + \sum_{k=1}^{n_p} \frac{\alpha_{k+1}}{\gamma} \ln(E_{t-1} P_{k,t}) + \frac{u_t - v_t}{\gamma} \quad (9.)$$

Where we can make use of assuming rational expectations, i.e., prices are expected not to change from one period to the next.

$$\ln P_{i,t} = \frac{\alpha_0 - \beta_0}{\gamma} - \sum_{j=1}^{n_p} \frac{\beta_j}{\gamma} \ln P_{j,t} - \frac{\omega}{\gamma} Y_t + \frac{\alpha_1}{\gamma} \ln P_{i,t-1} + \sum_{k=1}^{n_p} \frac{\alpha_{k+1}}{\gamma} \ln P_{k,t-1} + \sum_{j=1}^{n_z} \frac{\alpha_{j+k+1}}{\gamma} Z_{jt} + \frac{u_t - v_t}{\gamma} \quad (10.)$$

which is basically an ARDL(1,1) or ARX(1,1) model, where we explain the prices not only by contemporary prices of other products, but the lagged prices as well. Unfortunately this specification would leave us with too few observations and as such cannot be estimated with the current number of missing observations and the usual methods of data imputation result in bias in dynamic specifications.

5. Estimation results

In this section we present the results of estimating the reduced equation (6.) from the partial equilibrium model. We report the parsimonious specifications only (except for the multiple imputation), that is after sequentially removing regressors whose coefficient were not significantly different from zero at at least 10%. Also, we report all specifications with four missing data treatment methods: listwise deletion means that the regression is carried out on observations where all regressors were observed. This approach strongly reduces the size of our sample. As an alternative we report results form a pairwise deletion approach, when the main statistics of each variables are calculated using all available observations, and the parameters of the regression are estimated from those statistics. It has been found that if the correlation is high among the variables, listwise deletion has better properties (Allison, 1992: 9.). The mean imputation is the third method, where the missing observations in the regressors are replaced by their sample mean

(unconditional mean imputation). The missing values of the dependent variable are not imputed in this case, since that would lead to serious bias. This method is usually not recommended (see Little and Rubin, 2002: 62) since it results in a serious underestimation of the variances and covariances. Nevertheless as our Monte Carlo simulations (Appendix B) show, this method leads to inefficient (high standard errors) but unbiased parameter estimates. And finally we report the Multiple Imputation method based on a Fully Conditional Specification carried out 5 times.

Table 5
The partial equilibrium model results for barley with four missing data treatment method

	listwise deletion			pairwise deletion			mean imputation			multiple imputation		
	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value
Constant	-1.052	-1.318	.195	.048	.134	.894	.550	5.619	.000	-.949	-3.708	.016
lnpdates	.181	1.243	.221	.185	3.278	.001	.320	21.794	.000	.232	6.325	.002
lnpmus	-.162	-1.263	.214	-.260	-3.639	.000	-.105	-4.513	.000	-.216	-6.721	.001
lnpcres	.477	4.273	.000	.519	8.747	.000	.226	12.130	.000	.504	11.986	.000
lnpses	.629	3.335	.002	.362	4.902	.000	.235	11.141	.000	.413	5.892	.003
lnpwool	-.009	-.054	.957	.154	1.897	.060	.150	6.583	.000	-.005	-.065	.951
nameters	-.057	-.444	.659	-.177	-3.607	.000	-.034	-2.622	.009	.062	1.235	.280
t	-.0001	-1.241	.222	-.00004	-1.256	.211	.000005	1.650	.099	-.00005	-1.335	.247
February	-.051	-.231	.819	-.139	-.969	.334	-.020	-1.194	.232	.015	.113	.915
March	-.073	-.311	.758	-.195	-1.340	.183	-.024	-1.426	.154	.011	.098	.926
April	.085	.259	.797	-.164	-1.089	.278	-.012	-.690	.490	.082	.884	.419
May	-.858	-2.495	.017	-.474	-3.234	.002	-.047	-2.788	.005	-.228	-1.681	.162
June	-.935	-2.748	.009	-.332	-2.302	.023	-.046	-2.721	.007	-.408	-2.178	.092
July	-.685	-2.350	.024	-.203	-1.404	.163	-.030	-1.773	.076	-.287	-1.984	.114
August	-.677	-2.546	.015	-.218	-1.486	.140	-.042	-2.495	.013	-.306	-1.737	.154
September	-.489	-1.894	.066	-.293	-1.999	.048	-.042	-2.458	.014	-.262	-1.990	.112
October	-.472	-1.902	.065	-.149	-1.014	.312	-.032	-1.904	.057	-.228	-4.014	.006
November	-.057	-.211	.834	-.187	-1.296	.197	-.034	-1.990	.047	-.112	-1.619	.161
December	-.126	-.510	.613	.087	.606	.546	-.003	-.169	.866	.068	.712	.511
N	57			145			4091			4091		
R ²	0.860			0.799			0.304			0.805*		
DW	1.414			0.949			1.047			1.917*		

Note: Listwise deletion: observations with any of the variables having a missing value are excluded (default treatment method in most software packages). Pairwise deletion: summary statistics are calculated for each variable using all available observations. The coefficients are calculated from these statistics. Mean imputation: the missing values are replaced by the sample mean of the series. Multiple imputation: the missing values are imputed with a Fully Conditional Specification method 5 times, using all available variables. The regression is run on each set of imputations, pooled estimates are reported.

* calculated as the mean of the respective statistics of the five set of estimates

Table 6
The partial equilibrium model results for cress with four missing data treatment method

	listwise deletion			pairwise deletion			mean imputation			multiple imputation		
	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value
Constant	.572	.595	.556	.916	2.180	.031	1.508	19.381	.000	1.524	3.753	.017
lnpbar	.669	4.273	.000	.724	8.747	.000	.154	12.130	.000	.791	9.247	.001
lnpdates	.322	1.909	.064	.023	.331	.741	.061	4.798	.000	.007	.084	.937
lnpmus	.335	2.312	.026	.136	1.554	.123	.175	9.180	.000	.227	4.276	.008
lnpses	-.234	-.933	.357	.022	.230	.818	.294	17.211	.000	-.130	-1.429	.221
lnpwool	-.067	-.341	.735	.092	.954	.342	.173	9.189	.000	.155	1.390	.234
nameters	.064	.424	.674	.153	2.572	.011	.036	3.313	.001	.005	.090	.933
t	.0002	1.920	.062	-.00003	-.627	.532	-.000003	-1.247	.213	.0001	1.265	.272
February	.130	.497	.622	.105	.622	.535	.006	.415	.678	-.046	-.282	.791
March	.188	.683	.499	.064	.373	.710	-.003	-.244	.807	-.070	-.795	.460
April	-.044	-.114	.910	.049	.275	.784	-.004	-.317	.751	-.092	-.946	.386
May	1.261	3.238	.002	.365	2.063	.041	.010	.726	.468	.187	.906	.413
June	1.067	2.627	.012	.221	1.278	.203	.004	.264	.792	.248	1.200	.293
July	.610	1.715	.094	.111	.644	.521	.003	.186	.853	.149	.594	.583
August	.374	1.116	.271	.025	.144	.886	-.006	-.456	.648	.048	.172	.872
September	.275	.869	.390	.129	.736	.463	.004	.274	.784	.079	.324	.762
October	.344	1.139	.262	.112	.642	.522	.007	.498	.618	.146	1.078	.335
November	.145	.457	.650	.077	.451	.652	-.004	-.316	.752	.040	.333	.754
December	.055	.188	.852	-.164	-.968	.335	-.012	-.819	.413	-.164	-1.064	.342
N	57			145			4091			4091		
R ²	0.832			0.725			0.302			0.722*		
DW	1.164			0.701			0.999			1.942*		

Note: Listwise deletion: observations with any of the variables having a missing value are excluded (default treatment method in most software packages). Pairwise deletion: summary statistics are calculated for each variable using all available observations. The coefficients are calculated from these statistics. Mean imputation: the missing values are replaced by the sample mean of the series. Multiple imputation: the missing values are imputed with a Fully Conditional Specification method 5 times, using all available variables. The regression is run on each set of imputations, pooled estimates are reported.

* calculated as the mean of the respective statistics of the five set of estimates

Table 7
The partial equilibrium model results for dates with four missing data treatment method

	listwise deletion			pairwise deletion			mean imputation			multiple imputation		
	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value
Constant	.310	.354	.725	.814	1.501	.136	.175	1.765	.078	1.396	8.838	.000
lnpbar	.210	1.243	.221	.422	3.278	.001	.326	21.794	.000	.518	5.485	.004
lnpmus	.110	.789	.435	.349	3.211	.002	.438	19.433	.000	.250	2.466	.065
lnpcres	.265	1.909	.064	.038	.331	.741	.091	4.798	.000	.0003	.002	.998
lnpses	.300	1.336	.189	.069	.570	.570	.111	5.137	.000	.008	.117	.912
lnpwool	.254	1.468	.150	.238	1.947	.054	.076	3.289	.001	.281	4.063	.010
nameters	.026	.188	.852	-.092	-1.188	.237	-.016	-1.207	.227	-.148	-2.274	.079
t	-0.0004	-4.425	.000	-0.0001	-2.759	.007	-0.00001	-3.617	.000	-0.0002	-8.222	.000
February	-.092	-.386	.701	.132	.608	.545	.012	.683	.495	.090	.663	.538
March	.056	.222	.825	.129	.585	.559	.018	1.022	.307	.125	1.210	.276
April	.262	.748	.459	-.007	-.031	.976	.011	.650	.516	.028	.509	.616
May	-.088	-.222	.826	.039	.171	.865	.013	.765	.444	.207	1.011	.365
June	.264	.664	.510	.152	.684	.495	.016	.931	.352	.256	1.513	.197
July	.433	1.322	.194	.088	.400	.690	.007	.413	.680	.181	1.300	.253
August	.569	1.929	.061	.396	1.792	.076	.041	2.367	.018	.449	3.423	.019
September	.418	1.481	.147	.321	1.437	.153	.035	2.023	.043	.438	3.783	.012
October	.156	.561	.578	.110	.494	.622	.006	.376	.707	.217	1.594	.174
November	-.023	-.081	.936	.015	.068	.946	-.007	-.394	.694	.138	1.872	.098
December	.035	.132	.896	.037	.170	.865	.00003	.002	.999	.057	.425	.689
N	57			145			4091			4091		
R ²	0.550			0.592			0.308			0.597*		
DW	0.944			0.714			0.904			1.901*		

Note: Listwise deletion: observations with any of the variables having a missing value are excluded (default treatment method in most software packages). Pairwise deletion: summary statistics are calculated for each variable using all available observations. The coefficients are calculated from these statistics. Mean imputation: the missing values are replaced by the sample mean of the series. Multiple imputation: the missing values are imputed with a Fully Conditional Specification method 5 times, using all available variables. The regression is run on each set of imputations, pooled estimates are reported.

* calculated as the mean of the respective statistics of the five set of estimates

Table 8
The partial equilibrium model results for cuscuta with four missing data treatment method

	listwise deletion			pairwise deletion			mean imputation			multiple imputation		
	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value
Constant	-.434	-.435	.666	.995	2.369	.019	.296	4.497	.000	.540	.798	.468
lnpbar	-.243	-1.263	.214	-.364	-3.639	.000	-.047	-4.513	.000	-.297	-3.865	.015
lnpdates	.143	.789	.435	.215	3.211	.002	.194	19.433	.000	.155	2.273	.081
lnpcres	.359	2.312	.026	.137	1.554	.123	.116	9.180	.000	.199	3.343	.024
lnpses	.343	1.338	.189	.346	3.832	.000	.065	4.572	.000	.422	5.017	.006
lnpwool	-.104	-.517	.608	.330	3.557	.001	.031	2.035	.042	.137	.804	.466
nameters	-.183	-1.181	.245	-.276	-4.921	.000	-.047	-5.288	.000	-.284	-5.664	.004
t	-.0002	-1.476	.148	-.0002	-5.657	.000	-.00001	-3.730	.000	-.0003	-4.708	.009
February	-.069	-.253	.801	-.089	-.523	.602	-.010	-.849	.396	.032	.233	.826
March	-.183	-.641	.526	-.244	-1.422	.158	-.014	-1.187	.235	-.214	-2.222	.077
April	-.155	-.386	.701	-.240	-1.345	.181	-.005	-.397	.691	-.227	-2.434	.058
May	-.527	-1.182	.244	-.131	-.729	.467	.008	.696	.487	-.117	-1.714	.134
June	-.095	-.208	.836	.089	.513	.609	.010	.916	.359	.011	.091	.931
July	.011	.029	.977	.051	.293	.770	.003	.224	.823	.031	.322	.760
August	.164	.468	.642	-.026	-.150	.881	-.008	-.724	.469	.035	.343	.746
September	.291	.889	.380	-.098	-.554	.581	-.011	-1.000	.317	.062	.920	.390
October	.141	.443	.660	.172	.989	.325	.007	.633	.527	.208	2.192	.080
November	.333	1.028	.310	.097	.562	.575	.008	.681	.496	.256	1.797	.140
December	-.136	-.448	.657	.100	.585	.559	-.00003	-.002	.998	.115	.981	.375
N	57			145			4091			4091		
R2	0.550			0.550			0.180			0.590*		
DW	0.944			0.673			0.986			1.954*		

Note: Listwise deletion: observations with any of the variables having a missing value are excluded (default treatment method in most software packages). Pairwise deletion: summary statistics are calculated for each variable using all available observations. The coefficients are calculated from these statistics. Mean imputation: the missing values are replaced by the sample mean of the series. Multiple imputation: the missing values are imputed with a Fully Conditional Specification method 5 times, using all available variables. The regression is run on each set of imputations, pooled estimates are reported.

* calculated as the mean of the respective statistics of the five set of estimates

Table 9
The partial equilibrium model results for sesame with four missing data treatment method

	listwise deletion			pairwise deletion			mean imputation			multiple imputation		
	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value
Constant	2.736	6.428	.000	1.838	5.042	.000	2.887	51.517	.000	2.524	15.931	.000
lnpbar	.353	3.335	.002	.440	4.902	.000	.126	11.141	.000	.471	6.149	.003
lnpdates	.146	1.336	.189	.037	.570	.570	.058	5.137	.000	.002	.055	.958
lnpmus	.128	1.338	.189	.300	3.832	.000	.078	4.572	.000	.355	5.556	.004
lnpcres	-.093	-.933	.357	.019	.230	.818	.231	17.211	.000	-.095	-1.416	.224
lnpwool	.328	2.930	.006	.125	1.389	.167	.282	17.400	.000	.290	6.873	.001
nameters	-.014	-.148	.883	.178	3.256	.001	.033	3.443	.001	.026	.378	.724
t	.0002	2.962	.005	.0001	3.501	.001	.00001	2.981	.003	.0002	11.595	.000
February	.050	.303	.764	.025	.159	.874	-.002	-.147	.883	-.083	-.618	.567
March	.033	.189	.851	.120	.748	.456	.002	.129	.897	-.006	-.138	.892
April	-.145	-.593	.556	.062	.374	.709	-.007	-.571	.568	-.041	-.427	.688
May	.242	.883	.383	.239	1.436	.153	.008	.666	.505	.094	1.421	.204
June	-.156	-.561	.578	.153	.943	.347	.011	.884	.377	.270	1.392	.233
July	-.074	-.320	.751	.113	.703	.483	.011	.884	.377	.192	1.984	.106
August	.132	.619	.540	.111	.682	.497	.011	.921	.357	.186	1.440	.216
September	-.034	-.167	.868	.152	.929	.354	.013	1.020	.308	.094	.595	.582
October	.145	.753	.456	-.037	-.225	.822	-.002	-.136	.892	.070	.734	.497
November	-.077	-.384	.703	.137	.859	.392	.013	1.041	.298	.026	.326	.757
December	.081	.436	.665	-.009	-.059	.953	.004	.291	.771	-.040	-.600	.571
N	57			145			4091			4091		
R2	0.855			0.638			0.327			0.726*		
DW	1.636			1.019			1.030			1.918*		

Note: Listwise deletion: observations with any of the variables having a missing value are excluded (default treatment method in most software packages). Pairwise deletion: summary statistics are calculated for each variable using all available observations. The coefficients are calculated from these statistics. Mean imputation: the missing values are replaced by the sample mean of the series. Multiple imputation: the missing values are imputed with a Fully Conditional Specification method 5 times, using all available variables. The regression is run on each set of imputations, pooled estimates are reported.

* calculated as the mean of the respective statistics of the five set of estimates

Table 10
The partial equilibrium model results for wool with four missing data treatment method

	listwise deletion			pairwise deletion			mean imputation			multiple imputation		
	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value
Constant	-4.016	-8.742	.000	-3.497	-14.685	.000	-2.625	-49.455	.000	-3.876	-41.025	.000
lnpbar	-.008	-.054	.957	.179	1.897	.060	.070	6.583	.000	-.007	-.066	.950
lnpdates	.206	1.468	.150	.122	1.947	.054	.035	3.289	.001	.153	3.763	.015
lnpmus	-.065	-.517	.608	.274	3.557	.001	.032	2.035	.042	.121	.776	.480
lnpcres	-.045	-.341	.735	.077	.954	.342	.118	9.189	.000	.123	1.320	.255
lnpses	.550	2.930	.006	.120	1.389	.167	.245	17.400	.000	.305	6.822	.001
nameters	.214	1.788	.082	.186	3.493	.001	.020	2.276	.023	.203	2.882	.043
t	.0002	2.079	.044	.0002	6.925	.000	.00001	5.237	.000	.0002	9.767	.000
February	-.085	-.395	.695	.015	.100	.921	.005	.402	.687	-.029	-.274	.796
March	-.223	-.994	.326	.157	.999	.319	.010	.856	.392	.045	.313	.768
April	-.047	-.149	.882	.350	2.178	.031	.025	2.159	.031	.176	1.572	.181
May	.045	.125	.901	.222	1.358	.177	.017	1.453	.146	.059	.322	.763
June	.490	1.394	.171	-.017	-.107	.915	.007	.582	.561	-.103	-.605	.576
July	.034	.111	.912	.011	.070	.944	.010	.844	.399	-.070	-.448	.676
August	-.394	-1.453	.154	-.055	-.346	.730	.006	.498	.619	-.186	-.898	.418
September	-.249	-.963	.341	.051	.318	.751	.014	1.190	.234	-.150	-.978	.379
October	-.454	-1.885	.067	-.042	-.262	.793	.010	.825	.409	-.209	-1.027	.360
November	-.398	-1.582	.122	-.058	-.371	.711	.002	.166	.868	-.234	-1.698	.158
December	-.297	-1.266	.213	-.072	-.463	.644	.002	.164	.870	-.035	-.184	.863
N	57			145			4091			4091		
R ²	0.793			0.654			0.239			0.661*		
DW	1.165			0.504			0.899			1.925*		

Note: Listwise deletion: observations with any of the variables having a missing value are excluded (default treatment method in most software packages). Pairwise deletion: summary statistics are calculated for each variable using all available observations. The coefficients are calculated from these statistics. Mean imputation: the missing values are replaced by the sample mean of the series. Multiple imputation: the missing values are imputed with a Fully Conditional Specification method 5 times, using all available variables. The regression is run on each set of imputations, pooled estimates are reported.

* calculated as the mean of the respective statistics of the five set of estimates

With our assumption that the series are missing at random, we can assume that none of the methods we used result in biased estimates, but the first three methods result in an inefficient estimation so one needs to be critical with the reported t-statistics and the standard hypothesis tests. Seemingly, even though there is difference in magnitude and standard error, there is no difference in the sign of the coefficients. Hence, the sign of the coefficient (if significant) is

generally the same in all models, but the significance is difference may be biased in the first three methods. Therefore, we prefer the multiple imputation results.

The results of the demand and supply model are quite telling. As above regression results show, barley, dates, and cuscuta have relatively strong seasonality, while sesame, cress and wool do not. Equally, the 'na-meters' variable, giving the *na*, the distance in meters from the maximum river level, is negative for dates and cuscuta and positive for wool. In other words, the lower the river level, the lower the price for dates and cuscuta, while the higher the water level, the lower the price for wool. For cress, sesame, and barley the height of the river is not statistically significant.

As argued in Section 4, within the framework of the simple model from above Tables, we can interpret the price coefficients (elasticities) as having the same sign as the price elasticities in the demand equation provided that we assume the products own price elasticity to be negative, which is common in the literature. Positive price elasticity means that when the price of one product (the regressor) rises, the demand for the other product (whose price is the dependent variable) increases as well. This behavior is typical of substitutes. A negative coefficient on the other hand means that when a product gets more expensive, the demand for the product on the left-hand side of the equation reduces, indicating that the products are complements. We warn against the oversimplified interpretation of these relationships. Two products can be substitutes or complements even if they are seemingly not closely related, simply because of the existence of a budget constraint and because they have different utility for the consumer. For example, we find in Table 11 that wool was a substitute for dates. Obviously, by this finding we do not mean that people used wool for the same purpose as dates. What we argue is that if the price of wool rose people rather reduced their consumption of wool and increased that of dates. This is absolutely not illogical, since this is exactly what Engle's Law states: once real income drops (through price

changes for example), the proportion of foodstuffs in the consumption basket should increase and vice versa. Another interesting finding is that cuscuta and barley are complements. This speaks against the argument that, what we call “cuscuta”, was used solely to produce date beer.

Table 11
Signs of the significant coefficients (from the multiple imputation method)

	barley	cress	dates	cuscuta	sesame	wool
Barley	X	+	+	-	+	0
Cress	+	X	0	+	0	0
Dates	+	0	X	+	0	+
cuscuta	-	+	+	X	+	0
Sesame	+	0	0	+	X	+
Wool	0	0	+	0	+	X

6. Conclusion

In this paper we analyze the available Babylonian price series of six commodities (barley, cress, dates, cuscuta, sesame, wool). We wish to extend the already existing results by two new aspects of our research: a more careful treatment of missing data, the use of a structural economic model (partial equilibrium) for analyzing the market behavior. In addition, we use an indicator of weather (river level) to test its effect on prices.

Our results suggest the presence of a unit root, which implies that people could not forecast price changes, and hence, the weakest form of an efficient market hypothesis is fulfilled. Next, we apply a structural economic model to test for the presence of a supply-demand relationship. Although, because of lack of data on quantities, the model remains unidentified, and we therefore cannot interpret the level of the coefficients, still we can establish the presence of a supply-demand relationship. We find that commodity prices were strongly interrelated (through substitution and complementarity) suggesting a working market mechanism.

Within this model, we can include a number of exogenous variables. The water level had clearly a negative effect on dates and cuscuta while it worked positively for wool. Likewise barley

experiences relatively strong seasonality. All results seem to confirm the interpretation of Temin (2002) that during Babylonian times already a form of market mechanism was in operation.

References

Allison, P. D. (2001) *Missing Data*, SAGE

Andreau, J. (2002) Twenty Years after Moses I. Finley's *The Ancient Economy*, in: Scheidel, W. and Von Reden, S. (eds.). *The Ancient Economy*. Routledge, 33-49.

Bollerslev, T. (1986), Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics* 31(3), 307-327.

Britnell, R. H. And Campbell B.M.S. (1995) *A Commercialising Economy: England 1086 to c. 1300*. Manchester University Press

Broadberry, S., Campbell, B., Klein, A., Overton, M., and Van Leeuwen, B. (2011), 'British economic growth, 1300-1850: some preliminary estimates,' paper presented at the workshop Quantifying Long Run Economic Development in Venice, 22-24 March 2011.

De Vries, J. and Van der Woude, A. (1997). *The first modern economy. Success, failure and perseverance of the Dutch economy, 1500-1815*. Cambridge UP.

Gilks, W.R., Richardson, S. & Spiegelhalter, D.J. (Eds.) (1996) *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London.

Heckman, J. J., (1979) Sample Selection Bias as a Specification Error, *Econometrica*, *Econometric Society*, vol. 47(1), 153-61.

Hunger, H. and Pingree, D. 1999. *Astral Sciences in Mesopotamia (HdO I/XLIV)*.

Hunger, H. (2000) Non-mathematical Astronomical Texts and Their relationships. In: Swerdlow, N. M. (ed.). *Ancient astronomy and celestial divination*. MIT Press, 77-96.

Kwiatkowski, D., Phillips, P., Schmidt, P, and Shin, Y. (1992) Testing the null hypothesis of stationarity against the alternative of a unit root, *Journal of Econometrics* 54, 159-178.

Little, R. J.A. and Rubin, D. B. (2002) *Statistical Analysis With Missing Data*, John Wiley & Sons

Little, R. J.A. (1988) A Test of Missing Completely at Random for Multivariate Data with Missing Values. *Journal of the American Statistical Association* 83 (404), 1198- 1202

McKnight, P. E., McKnight, K.M., Figueredo, A. J., Sidani, S. (2007) *Missing data: a gentle introduction*, Guilford Press

Romero, N. E., Q. D. Y. Ma, L. S. Liebovitch, C. T. Brown and P. Ch. Ivanov, "Correlated walks down the Babylonian markets," *Europhysics Letters*, Vol. 90 (1) 2010, 18004-p1 - 18004-p6.

Schafer, J.L. (1997) *Analysis of Incomplete Multivariate Data*. Chapman & Hall, London.

Slotsky, A.L. (1997) *The Bourse of Babylon. Market quotations in the Astronomical Diaries of Babylonia*. CDL Press.

Temin, P. (2002), Price behaviour in Ancient Babylon. *Explorations in Economic History* 39 (1), 46-60

Van der Spek, R. J. and Mandemakers, C. (2003) Sense and nonsense in the statistical approach of Babylonian prices. *Bibliotheca Orientalis* 60, 521-537.

Van Zanden, J.L. and Bas van Leeuwen, 'The origins of 'modern economic growth'? Holland between 1347 and 1807,' paper presented at the workshop Quantifying Long Run Economic Development in Venice, 22-24 March 2011.

Appendix A: Monte Carlo simulation of unit root tests with data missing at random

We assume that series y is a random walk process with 3500 observations:

$$x_t = x_{t-1} + \varepsilon_t, \text{ with } x_0=2.5, \text{ and } \varepsilon \sim N(0, 0.2) \text{ (A1.1)}$$

The missingness is simulated based on the results from an LPM regression on the availability dummy of barley prices, which yields (t-statistics in parentheses):

$$D_t^{\text{missing}} = 0.0534 + 0.51 \cdot D_{t-1}^{\text{missing}} + u_t \text{ (A1.2)}$$

(12.3) (27.3)

where we find that standard error of the residual u is about 0.27. This formula simply reflects the autocorrelation in the probability of an observation to survive. Using this relationship we estimate the probability for an observation to survive. If this is less than 0.5, the observation is deleted, otherwise it is kept. This random simulation of the missingness means that the number of available observations changes by each iteration. Hence, we report the mean of the available observations in the summary table. Finally, we carry out DF and KPSS unit root tests just like in the paper and

observed the test statistics (t-statistics reported from the DF auxiliary regression and the KPSS LM statistics). This procedure is repeated 10000 times. The results for the Dickey-Fuller tests are reported in Table A.1.

Table A.1
Results from the Monte Carlo simulation for the DF test with nonstationary data

	level of significance		
	1%	5%	10%
Dickey-Fuller (one-sided test)	-2.73	-2.02	-1.65

average number available of observations=375.7

Kwiatkowski et al (1992) report asymptotic critical values for the KPSS test. In Table A.2 we report the share of those experiments, where the LM test statistics proved to be less than the asymptotic critical values (this is our estimate of the chance of Type-II error).

Table A.2
Results from the Monte Carlo simulation for the KPSS test with nonstationary data

	level of significance		
	1%	5%	10%
asymptotic critical values (Kwiatkowski et al, 1992)	0.739	0.463	0.347
% cases with smaller test statistics than the asymptotic critical value	2.14%	0.56%	0.26%

average number available of observations=375.7

We find that if x is indeed a random walk process, we would erroneously reject the null hypothesis of stationarity in at most 2.14 % of the cases, which is acceptable.

Finally, we may experiment with a specification that differs from A1.1 only in so far the autoregressive coefficient is now set to 0.99. These series are now stationary even though the effect of shocks will fade away very slowly. Now we can use the KPSS to test the null of stationarity (see Table A.3).

Table A.3

Results from the Monte Carlo simulation for the KPSS and DF tests with stationary data

	level of significance		
	1%	5%	10%
KPSS critical values with the autoregressive coefficient being 0.99	0.129	0.205	0.263
% cases when the DF test would have led to false acceptance of the null hypothesis of non-stationarity	95.0%	82.6%	70.6%

average number available of observations=376.3

The critical values are lower than the asymptotic critical values by Kwiatkowsky et al (1992), so we can conclude that if we can reject the null hypothesis of stationarity using their critical values we can safely argue that the series are nonstationary, even with a large number observations missing not at random. On the other hand, if we had used the critical values for the DF test as in Table A.1 (which are lower than suggested by MacKinnon (1996)), we would have not rejected the null hypothesis that y is a random walk process in 70.6 – 95% of the cases. With the critical values suggested by MacKinnon this ratio would have been even higher. This is indicative that we should prefer the KPSS test in this paper.

Appendix B: Monte Carlo simulation of static regression with data missing at random

The Data Generating Process (DGP) is assume dot be very similar to as described in Appendix A. We first create 3500 observations with the same formula as in (A1). Then we create variable y as follows:

$$y_t = 0.5 + 0.3x_t + e_t \quad e \sim N(0, 0.3) \quad (B1.1)$$

That is we create two cointegrated series. This is what we expect in case of related prices when they are determined by market mechanisms. Also, this is what our regression results suggest. That the two series are cointegrated means from an estimation perspective that the coefficients can be estimated with OLS procedure as it is consistent.

The missingness of the two series is simulated independently. In case of series x we used the very same method as described in (A1.2). For series y we used the same method, but with different coefficients (we used our dates price series to derive the coefficients).

$$D_t^{\text{missing}} = 0.0469 + 0.53 \cdot D_{t-1}^{\text{missing}} + v_t \quad (\text{B1.2})$$

(13.1) (20.9)

where we find the standard deviation of v to be 0.25.

This procedure led to a dataset very similar to the observed data, with the 3500 theoretical observation only 145 available on average for regression analysis (overlap of the two series). The estimation has been repeated 10000 times. We report the results in Table B.1

Table B.1
Results from the Monte Carlo Simulation

method	statistics	constant	slope	average number of observations
whole sample	average	0.500	0.300	3500
	std dev	0.009	0.001	
Listwise deletion	average	0.500	0.300	143.8
	std dev	0.049	0.007	
Mean imputation	average	0.500	0.300	342.9
	std dev	0.122	0.007	

The results in Table B.1 suggest neither method should lead to biased parameter estimates but the efficiency of the estimation reduces