Fiduciary Coinage and Its Collapse in Ancient Rome

Jorge Alonso Ortiz* José-María Da-Rocha†

September 12, 2026

Abstract

Roman Egypt's closed tetradrachm circulated for decades above its metal value. Four series make this observable: Greenland ice-core lead for silver production, papyrus grain prices, coin fineness, and silver in circulation, from 253,417 hoard coins. Inflation reached Egypt as two coordination shocks. A two-equilibrium model predicts the 296 CE jump as the release of the coin's premium, and the jump matches. Over seven centuries the silver price of wheat mean-reverts while nominal prices rise five to six orders of magnitude. Debasement reached circulating coin a generation after the mint, and Gresham's law fails inside the same hoards.

JEL codes: E31, E42, E51, N13, N23.
Keywords: quantity theory of money; commodity money; debasement; Gresham's law; Roman Egypt.

1 Introduction

Monetary economics has been organized for two and a half centuries around one claim. In the long run, the price level is governed by the quantity of money relative to the volume of trade. Hume (1752) stated the claim and at once qualified it. He observed that an inflow of specie stimulates activity before it raises prices, so that money is neutral in the long run but not in the short run. Fisher (1911) formalized the equation of exchange, and Friedman and Schwartz (1963) rested the modern case on a century of United States history in which money and prices moved together at low frequency while velocity drifted slowly. The empirical content of the proposition lies in this long-run relationship, and so does the difficulty of testing it. There are three difficulties. First, a central bank that accommodates prices or leans against output makes the money stock respond to what it is meant to explain. Second, velocity moves with financial innovation and with the opportunity cost of holding money, so proportionality holds only at the longest horizons. Third, a century of one monetary regime is also a century of change in what money is. As a result, the most convincing evidence has come from settings in which money moved massively for reasons unrelated to the current state of the economy. These are the hyperinflations, where fiscal collapse drives note

*Instituto Tecnológico Autónomo de México (ITAM). Email: jorge.alonso@itam.mx
†Universidade de Vigo. Email: darocha@uvigo.es

issue (Cagan, 1956; Sargent, 1982), and the cross-country record, where decades-long averages of money growth and inflation line up almost one for one (Lucas, 1980; McCandless and Weber, 1995). All three difficulties would disappear in an economy whose money supply was set outside it, and measured rather than inferred.

The Roman Empire is such an economy, and ice-core geochemistry has made its money supply observable. Roman silver was won almost entirely from argentiferous galena, a lead sulfide, by cupellation, a smelting process that released a fine lead aerosol. The westerlies carried this aerosol to the Arctic, where it was deposited year by year in the Greenland ice sheet, and it survives there as a datable layer of lead pollution (McConnell et al., 2018, 2019). As a result, the intensity of Roman silver mining is recorded in the ice at annual resolution over the whole life of the Empire. Greenland lead flux is an annual proxy for the flow of new silver bullion into the Roman economy.

Under a commodity standard the money supply answers to geology and to the treasury, not to a stabilizing authority. New silver enters when a vein is found or a mine reopens, and nominal money enters when the mint is pressed by war or by an accession donative. Neither margin responds to the price of wheat in the Nile valley, and this is the exogenous variation that modern tests find so hard to obtain. However, the series that measures the money base also measures something else. Newly mined silver was struck into coin and expanded the stock of money, but it was at the same time real output of mining, one of the largest industrial activities of the ancient world. The two roles have opposite implications for prices. If a mining boom is a real supply shock, it raises output and lowers the price level. If it is a monetary expansion, it raises the price level. This is why the lead series resists a simple classification, and why Alonso Ortiz and Da-Rocha (2026) can read it as a measure of economic activity where a monetary historian would read it as a measure of the money supply. Because the two readings imply opposite signs, the price data can decide between them.

The decision matters because there are two competing explanations of the great Roman inflation and no way so far to choose between them. The fineness of the denarius fell from ∼97% under Augustus, through Nero's reform of 64 CE and the Severan halving, to ∼5% under Gallienus and ∼2.5% under Claudius II. Over the same period came the steep price rises of the third-century crisis and the empire-wide price ceilings of Diocletian's Edict of 301. The first explanation, from Harl (1996) and Duncan-Jones (1994) to Katsari (2011), is monetary. In this view the inflation was driven by the supply of silver and by a debasement that stretched a shrinking metal base over a growing nominal coinage. The second explanation emphasizes real forces, such as population and supply shocks and the demographic catastrophes of the Antonine and Cyprianic plagues (Scheidel, 2010; Harper, 2017). In its sharpest revisionist form, it raises the possibility that what looks like inflation was monetization, that is, coin spreading into transactions previously settled in kind (Rathbone, 1996).

The debate has stayed open because of measurement. There has been no annual series for the Roman money supply. In addition, the only dense price record survives in Egypt, a province whose closed Alexandrian currency was deliberately insulated from the imperial coinage, so Egyptian prices need not track debasement elsewhere. The two gaps compound each other. Without a money series one cannot regress prices on money, and with prices quoted in a separate currency one cannot read imperial debasement off Egyptian quotations at all. Lead flux closes the first gap. Harper's DARMC record, which reports each dated Egyptian quotation both in the circulating coin and in grams of silver (Harper, 2016), closes the second. It does so on the premise, supported by the evidence on price convergence, that Rome was an integrated market economy in which relative prices carried information (Temin, 2013; Bowman and Wilson, 2009).

The debate is also a question in public finance. Debasement is the pre-modern inflation tax. A state under fiscal stress, facing war on the frontiers, donatives to the army, and the loss of revenue caused by plague, could raise revenue by minting more nominal coin from the same quantity of silver. It could thereby tax the holders of money through the resulting rise in prices. In this reading, the Roman episode is a natural experiment in seigniorage, in which fiscal shocks, plausibly exogenous to the grain economy, drive money creation and hence the price level. The mechanism is the ancient counterpart of the fiscal theory of hyperinflation (Sargent, 1982), and it turns the identification problem in our favor. In modern data the money supply responds to the state of the economy, and this endogeneity makes tests of the quantity theory difficult. In Rome the same feature is a source of exogenous variation instead, because the fiscal and military shocks that drove debasement had little to do with the harvest in the Nile valley.

Debasement raises a further question. Why did a coin whose metal had fallen to a twentieth of its former value continue to change hands at a stable price? Such a coin is a token. It is held above its intrinsic value because each agent expects others to accept it tomorrow. Money of this kind is a self-sustaining convention (Kiyotaki and Wright, 1989), and it is exposed to the self-fulfilling expectations that can destroy such a convention (Obstfeld, 1986). It is also the central theme of the economic history of small change. Sargent and Velde (2002) show that European monetary systems struggled for five centuries to sustain a token coinage without either chronic shortage or runaway depreciation, and that the formula for a stable token coinage (limited issue, convertibility, and a government monopoly) was learned slowly. The medieval debasement puzzle has its own formal literature (Rolnick et al., 1996; Redish, 2000).

Rome offers this literature a natural experiment of unusual clarity. Egypt's closed tetradrachm and, later, the imperial antoninianus were token coins in the precise sense. Their metal value lay far below their exchange value, and the gap was held open for decades by expected acceptance alone. The experiment has a clear treatment, the two reforms of 274/5 and 296 that changed the coin of account and with it the coordinating expectation. It also has a clear outcome, the price explosions that the papyri date to those reforms. Sargent and Velde (2002) reconstruct the European experience from mint records and ordinances. By comparison, the Egyptian price record allows us to observe a single fiduciary equilibrium form under debasement, hold through a plateau, and break on a datable shock.

The paper makes five contributions, which we list in decreasing order of the confidence we place in them. First, we document a fiduciary coinage over its whole life cycle, a token held for decades above its metal content and released at two datable coordination shocks, and we test the two-equilibrium model that describes it. The model predicts that the price jump at a reform equals the premium the coin had accumulated. At 296 the predicted jump is 1.7–2.3× and the observed jump is 2.1×, where the prediction uses the coin's measured content and an assumed metal-anchored price $\pi \approx 19$–25 g/hl. The pre-reform price cancels from this comparison, so the test checks whether the reform returned the silver price of wheat to its seven-century band, and it did. We report the re-tariffing of 274/5 as a dated regime switch rather than a second test. It confounds the release of the premium with a re-tariffing of face value that the model does not price, and the direction of its change in content is contested (Section 5). The logic of the discharged premium still explains why the second jump was proportionally smaller.

Second, we measure the coin that was actually in circulation. Every quantity in the monetary history of Rome that divides by the silver content of the coin has divided by the mint's standard, that is, the content of coin leaving the mint in year t, and not by the content of the coin that people held. We price each of 253,417 dated silver coins in 901 hoards at the fineness in force when it was struck, and recover the circulating series $\bar{c}_t$ directly. After the Severan halving of 194, $\bar{c}_t$ stood 29 percentage points above the mint standard, and it closed half of this gap only after 36 years. Aggregates based on the mint standard are therefore wrong in timing and not only in level (Section 6).

Third, we construct the debasement natural experiment. The silver content of the coin varies enormously while the silver price of grain does not. Comparing the nominal price of grain with its price in silver separates money from real scarcity without a structural model. In the fourth century only ~2% of the variance of nominal prices survives the conversion to silver, and the converted price returns to the band it occupied nearly three centuries earlier. The fourth century is the only period in which the content of the coin is both independently measured and genuinely variable. Fourth, we provide a consistency check of the quantity theory at generational resolution. Across six centuries, nothing in the binned record shows the bullion stock and Mediterranean trade diverging, although the exercise identifies no elasticity (Section 7). Fifth, and underlying all of the above, we assemble the data that a quantitative monetary history of Rome requires and has so far lacked, from a continuous fineness series for both coinages to a machine-extracted panel of exact-dated papyrus prices (Table 1).

Our main findings are as follows. First, the flow of bullion is disinflationary. The correlation between the bullion flow and prices is −0.45 on the published panel, and it does not depend on the leakage rate because leakage enters only the stock. The sign is negative in each of the six panel variants we construct, but the magnitude is not robust, running from −0.04 to −0.45, so the argument rests on the sign. A mining boom coincides with cheaper grain, so the real role of bullion dominates in the short run and lead flux is first of all a series of real output. Second, the relation between the stock of bullion and prices is not identified at generational resolution, and Section 7 reports the whole range of estimates. Third, the great inflation was overwhelmingly monetary. The silver price of grain mean-reverts across seven centuries, while nominal prices rise by five to six orders of magnitude.

1.1 Related literature

This paper is related to four literatures. The first is the quantitative economic history of antiquity (Temin, 2013; Duncan-Jones, 1994; Katsari, 2011), including the work on prices and quantities of the Oxford Roman Economy Project (Bowman and Wilson, 2009; Allen, 2009). Relative to this literature, we provide the annual series for the money supply that it has lacked, and a direct measurement of the coin in circulation where the numismatic literature has had to assert one. In addition, we test the two standing readings of the third century on evidence that neither had. These are the monetization hypothesis of Rathbone (1996), which we test against the payment record of the papyri, and the demographic reading of Harper (2016) and Scheidel (2010), which we test against a path of coin content whose decline begins before the Antonine Plague reached Egypt.

The second literature is the theory of commodity and token money (Kiyotaki and Wright, 1989; Obstfeld, 1986; Rolnick et al., 1996; Redish, 2000; Sargent and Velde, 2002). We contribute a natural experiment in fiduciary collapse, namely the whole life cycle of a token coin, with a premium computed from the coin's own metal and a datable price jump against which to test it. Sargent and Velde (2002) reconstruct five centuries of European experience with token coinage from mint records and ordinances. In our case, the token is observed directly in prices, from its formation through the plateau to the break.

The literature closest to this paper is the empirical work on Gresham's law. Rolnick and Weber (1986) show that the law in its usual form fails repeatedly in the historical record, and that bad money drives out good only when the good money cannot circulate at a premium. Velde et al. (1999) formalize the distinction in a model with light and heavy coins, imperfect information, and equilibria in which coin passes by weight or by tale. Which regime held has had to be inferred from ordinances and from anecdotes about circulation. Section 6 measures it. Within the same hoards, and across the largest break in fineness in the Roman record, the finer cohort does not leave circulation faster. On this coinage and in these decades, coin passed by tale, which is also the condition that the fiduciary equilibrium of Section 5 requires.

Finally, Alonso Ortiz and Da-Rocha (2026) read the same lead series as a proxy for output and find that the Roman business cycle was overwhelmingly endogenous. The fact that supports that reading is that the bullion flow is disinflationary and hence predominantly real. This paper establishes that fact, and it shows that the great inflation layered on top of the real cycle was nonetheless monetary.

The rest of the paper is organized as follows. Section 2 describes the monetary system and sets out a framework with two margins. Section 3 describes the data. Section 4 presents the debasement natural experiment. Section 5 develops the model of the fiduciary equilibrium and its dated collapse. Section 6 measures the coin that was actually in circulation. Section 7 closes the quantity identity at the scale of the empire. Section 8 concludes. Appendix A gives the details of data construction, and Appendix B reports the evidence on the money stock, the coin counts, and the fiscal origins of debasement.

2 The monetary system and the bullion framework

2.1 Institutions

The Roman coinage was a tri-metallic system (Harl, 1996; Duncan-Jones, 1994). The unit of account and the main coin for large payments was the silver *denarius*, struck at ∼97% fineness and ∼3.9 g under Augustus. Above it stood the gold *aureus*, tariffed at 25 denarii throughout the Principate. Below it circulated a token bronze coinage (the *sestertius*, *dupondius*, and *as*) whose face value already exceeded the value of its metal. The mint was an arm of the treasury, not an independent authority. It struck coin to meet fiscal needs, such as army pay, accession donatives, and public works. When silver was short, its main instrument was to cut the fineness of the denarius while holding its tariff, so that more nominal money could be minted from the same metal. The mint used this instrument with increasing force from Nero's reform of 64 CE, through the Severan halving, to the collapse of the third century, when the main silver coin, by then the double-denarius *antoninianus*, fell below 5% fine.

Egypt was a separate monetary area. From the Ptolemies through the third century it operated a closed currency. The imperial denarius did not circulate there, and internal payments used the billon Alexandrian *tetradrachm*, nominally four drachmas but with its own, separately managed silver content. Coin could not cross the boundary freely, so Egyptian prices in drachmas need not track the debasement of the denarius elsewhere. This feature insulated Egyptian nominal prices for two centuries and then released them abruptly when the boundary was dismantled. The dismantling came in two steps that structure much of the analysis below. The first was Aurelian's re-tariffing of the coinage in 274/5. The second was Diocletian's abolition of Egypt's separate coinage in 296, followed by the empire-wide maximum prices of the Edict of 301. The two series at the center of the paper should be read against this background. The flow of new silver, measured by lead flux, is the metal, and it is set by the mines. The fineness of the coin is the money per unit of metal, and it is set by the mint.

2.2 Two margins of bullion

Let b_t denote bullion production, proxied by lead flux. Bullion enters the economy along two margins. As real output, b_t is value added in mining. A mining boom raises real output Y_t, which lowers the price level for a given money stock. In addition, booms tend to coincide with good harvests and peace, which independently depress grain prices. As the monetary base, b_t adds to the stock of silver money. The relevant object is then the stock M_t, an accumulation of past flows net of leakage,

$$M_t = \sum_{s \le t} (1-\delta)^{t-s} b_s, \tag{1}$$

where δ captures silver leaving circulation. The channels are the documented drain of Roman silver to India and China in exchange for silk and spices, hoarding, physical loss, and re-minting under debasement. Under the quantity theory, $\log P_t = \log M_t + \log V_t - \log Y_t$, so the money stock should raise the price level in the long run.

The framework predicts a *reversal of sign by horizon*. Through the real margin the flow is disinflationary in the short run, and through the monetary margin the stock is inflationary in the long run. The rest of the paper tests both margins. Where the stock margin proves hard to identify, we isolate the monetary channel through the fineness of the coinage instead. Debasement raises the *nominal* money supply for a given quantity of silver. It is a pure expansion of the money supply, which the stock in equation (1) does not by itself capture.

2.3 Debasement and the inflation tax

Debasement is best seen as a third margin, and its origin is fiscal. Let c_t^m be the mint standard, the silver content of coin struck in year t, and b_t the flow of new bullion. The nominal coin stock is then $M_t^{\text{nom}} = \sum_{s \le t} (1-\delta)^{t-s} b_s / c_s^m$. For a given path of bullion, a fall in c_t^m raises nominal money one for one, $M^{\text{nom}} = M^{\text{Ag}}/c$. The real revenue that the state extracts by cutting c_t^m is the classical seigniorage of the inflation tax. It consists of the resources commanded by newly minted nominal money, and it is paid for by the fall in the real value of the money already held. A state under fiscal stress, facing war, donatives, and the loss of revenue caused by plague, will therefore lower c_t^m, and it will do so for reasons that have nothing to do with the Egyptian harvest. In this sense the great debasement is a natural experiment in seigniorage, and the fineness series, not the bullion stock, isolates the monetary channel. Fineness moves with fiscal pressure and bullion with mining geology, and the two need not move together except at the lowest frequencies. Section 6 measures how far apart they run.

The notation carries a distinction that the literature has generally ignored, and Section 6 shows that the distinction is large. The mint standard c_t^m applies to coin issued in year t. Prices, however, are quoted in the coin that people actually hold, and the wedge $M^{\text{nom}}/M^{\text{Ag}}$ depends on the average content of that coin, which we denote $\bar{c}_t$. The two coincide only in an economy that recoins instantly. Rome did not recoin instantly. The series $\bar{c}_t$ is an average over decades of past issues, and after a debasement it falls toward the new standard with a lag that we measure at 36 years. Where an argument below depends on the coin in circulation rather than on mint policy, above all for the fiduciary premium of Section 5, the relevant object is $\bar{c}_t$, and we indicate where the available series is c_t^m instead.

3 Data

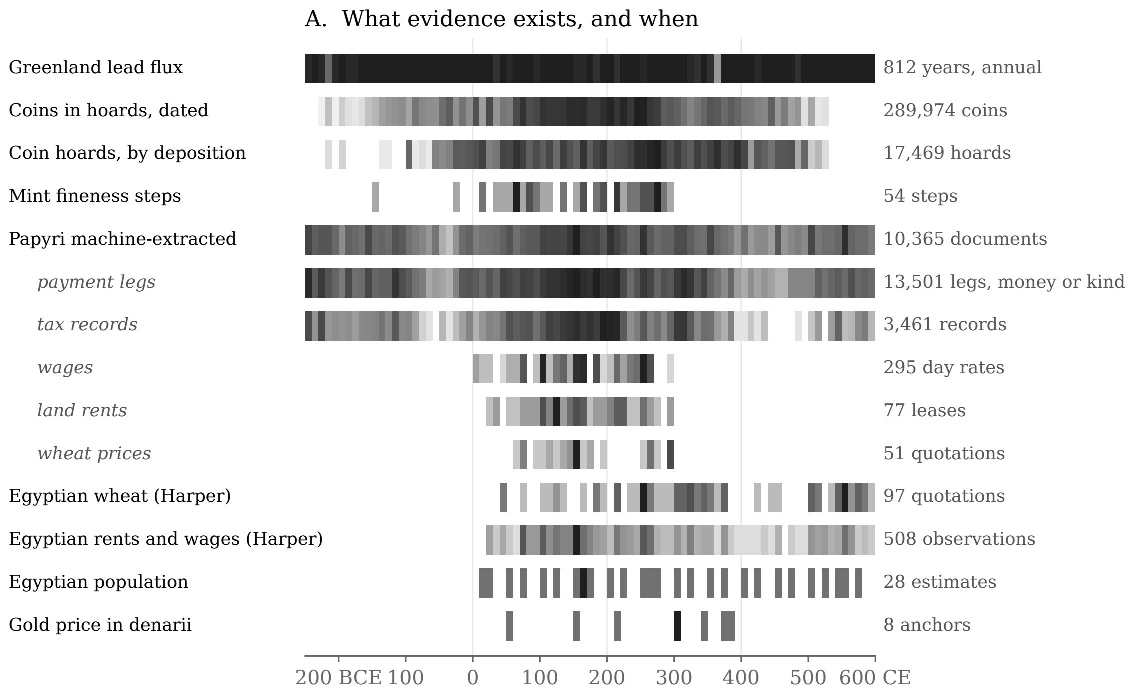
We draw on five bodies of evidence. Table 1 summarizes the assembled series, and Appendix A gives the details of construction, sources, and coverage.

The money base is Greenland lead flux, annual, 200 BCE–2014 CE (McConnell et al., 2018, 2019). The compilation of 25 ice-core sites behind it supplies the falsification test of Appendix A. The price level is the silver price of grain. It comes from Babylonian barley (van der Spek and Mandemakers, 2003) and from the 104 dated Egyptian wheat quotations of Harper's DARMC

Table 1: The assembled series.

Series	Role	Coverage	N
Greenland lead flux	bullion flow / money base	200 BCE–2014 CE	annual
Egyptian wheat, nominal & silver	price level P_t	45–650 CE	104 (102 Ag)
Babylonian barley (silver)	price level (check)	Hellenistic–Roman	132
Egyptian wages / rents / land (Harper)	factor prices	1–7C CE	258 / 214 / 67
DDbDP papyri, extracted here	source corpus	1–400 CE	12,482 docs
wheat / barley prices	price level (annual)	68–294 CE	51 / 17
land rents (artabas/aroura)	real factor price	1–296 CE	77
wages (drachmas/day)	factor price	1–296 CE	295
tax register rolls	fiscal base	8–300 CE	15,954 entries
flood-failure declarations	harvest shock controls	1–250 CE	85
Heroninos estate accounts	intra-crisis factor prices	249–268 CE	234 records
Egyptian population	real activity Y_t	14–575 CE	28
Mediterranean shipwrecks	real activity Y_t (empire)	Republic–Late Antiquity	series
CHRR+CHRE coin hoards	circulation / hoarding (δ)	Republic–Late Antiquity	17,989
CHRE coin-level layer, added here	circulating content $\bar{c}_t$	30 BCE–310 CE	253,417 coins
Denarius/antoninianus fineness	mint standard c_t^m	Augustus–Diocletian	25
fineness step table, added here	c_t^m as a step function	150 BCE–305 CE	54 steps
Alexandrian tetradrachm fineness	mint standard c_t^m (Egypt)	Tiberius–Diocletian	14
Gold price in denarii	second numeraire	50–385 CE	8
Military *stipendium*	nominal-wage anchor	Augustus–Maximinus	dated
Diocletian's Price Edict	administered ceiling	301 CE	basket

geodatabase (DARMC, 2016; Harper, 2016). The quotations are reported in the circulating coin and, for 102 of them, in grams of silver, and the debasement experiment rests on this distinction. The grain entries of Diocletian's Price Edict (Allen, 2009) give an anchor outside Egypt at the end of the crisis. Circulation and debasement are measured separately. A set of 17,989 coin hoards (CHRR, 2026; CHRE, 2026) serves as a proxy for hoarding, and a reign-level fineness series for the denarius and for Egypt's tetradrachm (Harl, 1996; Butcher and Ponting, 2015; Christiansen, 2004; Howgego et al., 2010) records mint policy. Beneath the hoard counts, the coin-level layer of Coin Hoards of the Roman Empire, used here for the first time, supplies 253,417 dated silver coins in 901 hoards. These coins are the raw material of Section 6. Figure 1 shows the whole record. The anchors from the price of gold, military pay, and the grain dole, and all the details on sources and coverage, are in Appendix A.



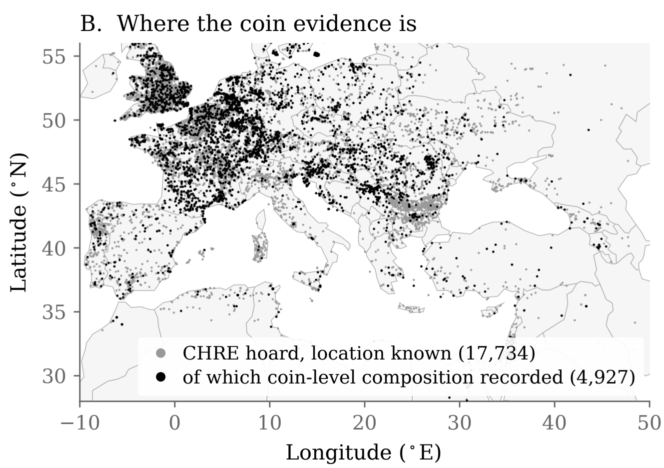


Figure 1: **The assembled record.** Panel A shows each body of evidence used in the paper, by decade. Because the rows differ in size by four orders of magnitude, the shading is the share of the row's observations that fall in each decade, normalized within the row, and the count at the right gives the scale. Counts are restricted to the plotted window, 200 BCE–600 CE, and can therefore fall short of a total quoted in the text. Indented rows are drawn from the machine-extracted papyrus corpus above them. Panel B shows the geography of the coin evidence. Gray marks every hoard in Coin Hoards of the Roman Empire with a known location. Black marks the subset whose composition is recorded coin group by coin group. This subset is the layer measured in Section 6, and it is the source of the imbalance in coverage discussed in Appendix B. A further 312 located hoards lie outside the frame, among them the Indo-Roman finds that bear on the eastern drain discussed in Section 7.

4 The debasement natural experiment

Harper's panel reports 104 dated quotations of the price of Egyptian wheat. It gives the nominal price for every quotation and, for 102 of them, the silver equivalent, that is, the nominal price converted at the actual silver content of the coin. Over 45–650 CE the nominal price rises five to six orders of magnitude, while the silver-equivalent price mean-reverts around a median of ∼19 g of silver per hectoliter. The monetary channel is thus clearest in the fineness of the coinage, not in the bullion stock. Debasement raises the nominal money supply for a given quantity of silver, a pure expansion of the money supply. The silver-equivalent price nets this expansion out and is therefore the real price, while the nominal price includes it. Comparing the two series separates money from real scarcity (Figure 2).

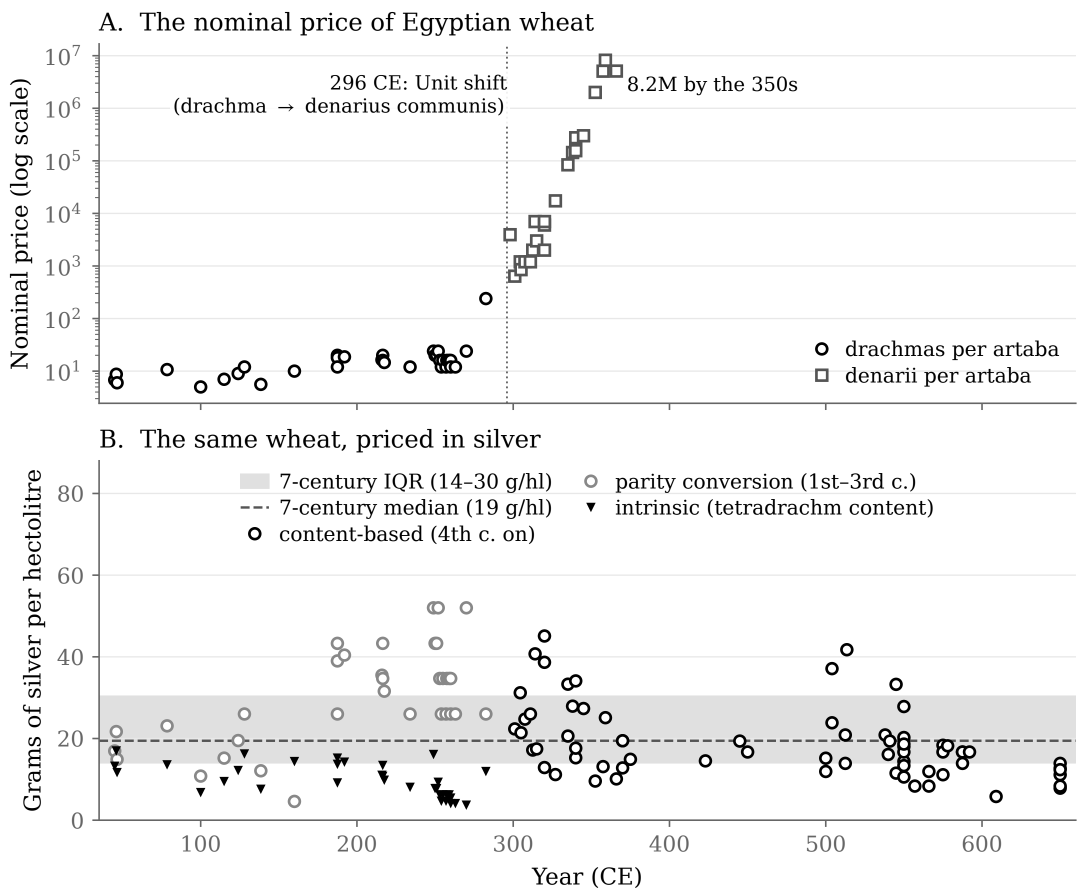


Figure 2: **The debasement natural experiment.** Panel A shows the nominal price of Egyptian wheat in the circulating coin of account (log scale; drachmas per artaba to the 270s, denarii per artaba thereafter). It rises five to six orders of magnitude. Panel B shows the silver-equivalent price (linear scale; grams of silver per hectoliter), which mean-reverts around a median of ∼19 across seven centuries. The gap between the two series is debasement. Source: Harper (2016) DARMC panel, $n = 104$ dated quotations (102 with a silver conversion).

4.1 The fourth century

The silver price of Egyptian wheat mean-reverts across seven centuries. Its median is ∼19 g of silver per hectoliter (interquartile range 14–30), and it shows a mild secular decline rather than an explosion. The nominal price rises five to six orders of magnitude, from ∼15 drachmas per artaba in the early Empire (nearer ∼8 in the first century itself) to over eight million denarii per artaba in the 360s. The cleanest window is the fourth century, when the silver content of the coin varies so much that the conversion is genuinely informative. Between 301 and 366 CE the nominal price rose ∼13,000-fold ($n = 22$ quotations). Over the same years the silver-equivalent price stayed within a 4.7-fold band and, if anything, moved mildly against the nominal price ($r = -0.35$). A decomposition of the variance of the log nominal price over this window leaves only ∼2% in the real (silver) component. By this measure, the fourth-century inflation was ∼98% monetary.

This decomposition is an accounting identity before it is a test. The silver price is the nominal price converted at the content of the coin, so $\log P^{\text{nom}} = \log P^{\text{Ag}} - \log c$ holds by construction. A variance share must assign most of the variance to whichever term moves more, and the ∼98% is a share net of a −14% covariance term rather than a clean partition. The informative result is a different one. Converting fourth-century prices at the independently measured content of the coin returns a silver price of grain inside the same 14–30 g/hl band as in the first century, some 270 years and five orders of magnitude of nominal inflation earlier. This agreement is not true by definition, and an account of the great inflation based on real scarcity has no ready explanation for it.

The source of the monetary explosion is the collapse of the silver content of the coinage, shown in Figure 3. The imperial denarius fell from ∼97% fine under Augustus to ∼2.5% by the 270s. Egypt's tetradrachm followed its own path, and it insulated Egyptian prices until its own content crashed, abruptly from 164/5 CE, on the eve of the Antonine Plague rather than at its onset, and terminally after 259/60.

4.2 The third-century crisis

The same comparison places the third-century crisis. The Plague of Cyprian and the military anarchy raised the real price of wheat by about half at its peak in 200–270, from ∼23 to ∼35 g/hl. This rise is an overlay on a nominal inflation of many orders of magnitude. In addition, it is an upper bound on the real movement rather than a measurement of it. Harper's silver conversion is nearly constant through the first three centuries (Appendix B), so his third-century real series is the nominal series rescaled, and the +50% is close to the nominal rise carried through. What the record supports is the ordering. The crisis moved real prices by tens of percent, while debasement moved nominal prices by orders of magnitude.

4.3 The onset of inflation in Egypt

Nominal prices in drachmas held a median of ∼13 (range 5–24) and were stable across 45–263 CE, even as the denarius fell from 97% to 20% fineness. Egypt's closed Alexandrian tetradrachm insu-

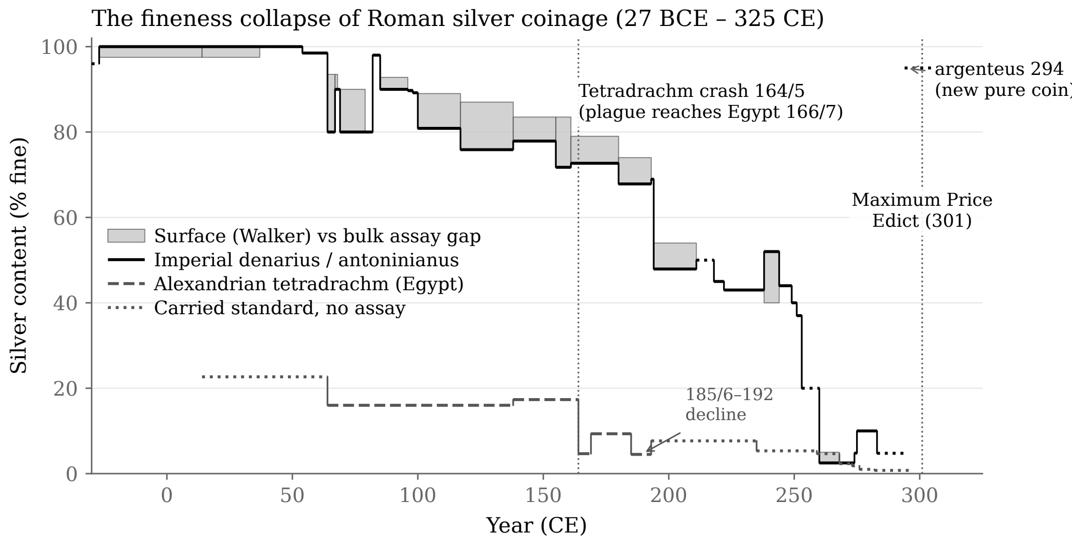


Figure 3: **The fineness collapse.** Silver content (% fine) of the imperial silver coin (the denarius to 237, the antoninianus from 238, and the argenteus from 294) and of Egypt's closed Alexandrian tetradrachm, drawn from the fineness step table of Appendix A. Each standard is held constant between reforms and is never interpolated across one. Dotted segments are carried standards rather than assays, and the band on the imperial line is the surface (Walker) reading above the bulk assay. The tetradrachm crashes from ∼14–20% to ∼4–5% beginning in 164/5 CE (vertical dotted line), shortly before the Antonine Plague reached Egypt (166/7–167/8). It recovers to a half standard from 169/70 and declines again from 185/6. Sources: Harl (1996), Walker's *Metrology*, and Butcher & Ponting (2015) for the imperial coin; Christiansen (2004) and Howgego et al. (2010) for the tetradrachm.

lated the province from the imperial coinage. The Heroninos estate accounts confirm this stability from inside the crisis. Across 234 dated transactions of 249–268 CE, nominal day wages held at 2–4 drachmas a day (median 2.3) while the silver in the coin was being stripped away. The inflation reached Egypt abruptly in the 270s, when Aurelian's reforms and the collapse of the tetradrachm integrated Egypt into the imperial currency, and by 282 CE nominal prices had jumped ∼18-fold. The inflationary effect of debasement thus appears as a delayed release at a regime switch, not as a smooth annual pass-through. The numismatic literature has long noted this pattern (Harl, 1996; Duncan-Jones, 1994), and the price record now makes it quantitative.

The onset can be dated from the papyrus record of the crisis years, assembled from Duncan-Jones (1976) and Rathbone (1996). This record fills the interval between Harper's last drachma quotation (282 CE) and the first price in the reformed coinage (301 CE). Egyptian wheat, at ∼16 drachmas per artaba through the mid-third century, jumps roughly twelvefold at Aurelian's reform of 274/5, to ∼200 drachmas by the late 270s. This half of the evidence rests on a single quotation of debated date, O.Mich. 1.157, and is a datum rather than a series. Prices then hold a plateau of 200–300 drachmas for two decades (240 in 282, 300 in 293, and ∼220 in 294). A second and larger move follows, to 640 by 301 and to the Edict's ceiling of 1,200–1,333, coincident with Diocletian's abolition of Egypt's separate coinage in 296. The great Egyptian inflation was a regime switch in two steps, tied to the reforms of 274/5 and 296, and not a smooth pass-through.

The plateau is the clearest evidence of the mechanism. By 279 the reformed tetradrachm held only ∼0.08–0.20 grams of silver, yet wheat stayed at 200–300 drachmas for twenty years. The coin circulated far above its metal content until Diocletian's reform broke the arrangement. This is the fiduciary equilibrium modeled in Section 5 (Figure 4).

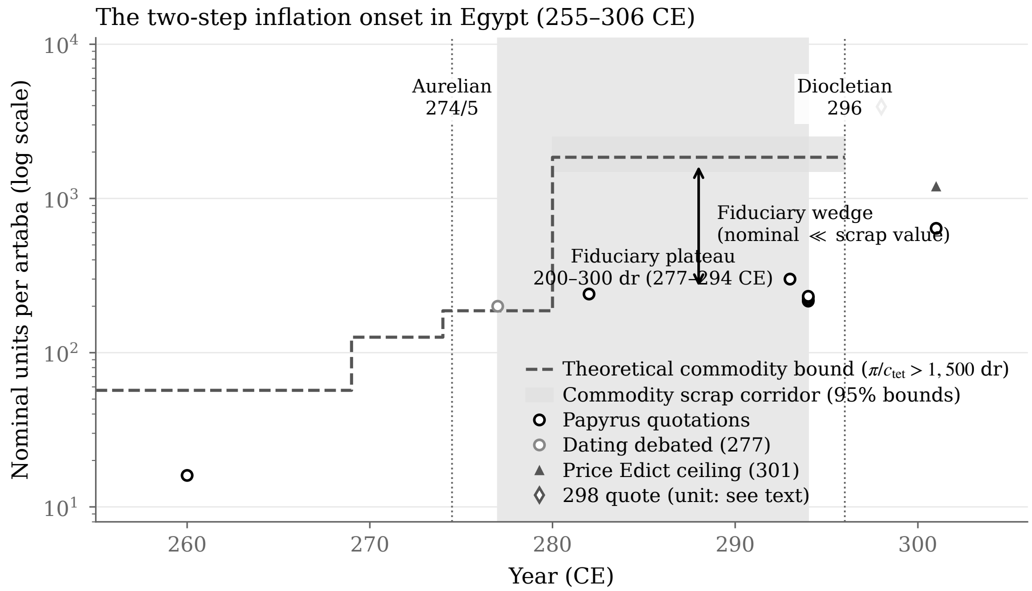


Figure 4: **Egypt's dated two-step onset.** Wheat price in drachmas per artaba (log scale) across the 282–301 gap, from the papyrus record inside the crisis. A stable price of ∼16 drachmas through the mid-third century jumps roughly twelvefold at Aurelian's reform of 274/5. It then holds a plateau of 200–300 drachmas for two decades. The plateau is itself a fiduciary equilibrium, since the reformed coin held only ∼0.08–0.20 g of silver. A second and larger move comes at Diocletian's abolition of Egypt's separate coinage in 296. Sources: Duncan-Jones (1976), Rathbone (1996), Harper (2016).

4.4 Two anchors across the 282–301 gap

Diocletian's Price Edict confirms the monetary reading from outside Egypt. The Edict is an empire-wide maximum in the reformed coinage, not an Egyptian papyrus price, and it caps wheat at 100 denarii per *castrensis* modius. At Allen's silver equivalent of the denarius (0.032 g), this is ∼24.8 g of silver per hectoliter. The figure is close to Harper's contemporary Egyptian silver price (∼22 g/hl in 301 CE) and inside the seven-century band (median ∼19, IQR 14–30). Two independent sources, a western administered ceiling and Egyptian market papyri, agree that the real price of wheat was normal in 301 CE, at the moment when nominal prices had run away. Because the Edict is a ceiling, the agreement also shows that the authorities pegged it to the prevailing real price expressed in the debased coin. In other words, they tried to freeze the nominal price at its normal silver value, and this is itself a diagnosis of the inflation as monetary. The wage entries of the Edict extend the check to labor. In real terms, goods prices and wages passed through the great inflation roughly intact

(Appendix A).

Gold provides a second numeraire that crosses the same gap, and it gives the same result. A pound of gold goes from ∼1,125 denarii under the Principate to 60,000 in 300 CE and $\sim 5\times10^8$ by the 380s, the same monetary explosion measured in a second numeraire. Meanwhile the gold–silver ratio implied by the mint's tariff drifted from 11.7 : 1 to 6.4 : 1, because the aureus stayed tariffed in a debasing denarius. Gresham's law requires a posted mispricing of this kind. As Section 6 shows, the blanched silver coinage never offered such a mispricing (Appendix B, Figure 8).

4.5 The price rise of the Antonine period

The first three centuries are harder to decompose. The standard silver conversions rescale the nominal series at a nearly constant ∼0.97 g of silver per drachma and ignore the tetradrachm's own debasement. We convert instead at the measured content of the tetradrachm (Appendix B). With this conversion, the best-known price event in Egypt can be decomposed. Wheat doubles between the 150s and the 190s, as its median rises from 7.9 to 17.2 drachmas per artaba ($\Delta \log = +0.78$). Harper (2016) reads this rise as the effect of the Antonine Plague, and Rathbone (1996) as the effect of debasement. Without the documented decline from 185/6 that Howgego et al. (2010) report, the coin accounts for 74–96% of the rise. Restoring that decline raises the share to 139–187%, which means that the intrinsic real price fell across the window. We report the decomposition as at least ∼75% coin and, on the corrected content path, most likely more than 100%. Appendix A sets out the content path and its uncertainty.

The chronology is the sharper result. The fineness of the tetradrachm begins to crash in 164/5 CE. The plague is first attested at Seleucia in the winter of 165/6 and reaches Egypt only in 166/7–167/8. The debasement therefore precedes the epidemic, and it cannot have been a response to fiscal distress caused by the epidemic. The Parthian war of 161–166 is the natural trigger. In addition, the corrected content path keeps falling until the end of the reign of Commodus, two decades after the epidemic arrived. The decomposition splits the price rise, not its ultimate cause. The chronology excludes the plague as the trigger and leaves it as a candidate only for the later phase.

The harvest is a candidate for the onset but not for the later phase. The corpus holds 85 formal flood-failure declarations dated 1–250 CE. These are *apographai abrochiās*, by which a landholder registered land that the Nile had not reached, and they cluster at the start of the window. Of these, 19 fall in the 160s, nine of them in 163 alone and from two nomes, against four in the 150s and four in 170–189. A failed inundation therefore stands beside the Parthian war at the onset of the doubling. It cannot account for the rise across the 170s and 180s, when declarations are rare, and the decomposition leaves it at most a quarter of the doubling.

Two further pieces of evidence bear on the experiment, and Appendix B develops both. First, the nominal coin stock, separated from bullion by the fineness series, moves with Egyptian prices during the era of physical debasement, and stops doing so during the fourth-century inflation of the unit of account. The correlations are +0.89 before 296 ($n = 42$) and −0.57 after ($n = 22$, 95%

CI $[-0.80, -0.20]$), and both are sensitive to the leakage rate and to the activity series. Second, an index of imperial political instability is negatively related to coin fineness (-0.48). The index runs at 2.6 events a year during the crisis of the soldier emperors, against 0.2 under the Principate. The debasement was driven by fiscal pressure rather than by harvests. Neither series is a clean instrument, but both corroborate what the fineness record establishes directly.

4.6 The monetization hypothesis

The sharpest objection from the revisionist side is due to Rathbone (1996), who argues that the third-century price rise reflects monetization, that is, coin spreading into transactions previously settled in kind, rather than a fall in the value of money. The hypothesis has a testable implication, and the papyri can test it, because Egyptian contracts record the form in which each leg of a payment was settled. In 12,370 dated documents we classify every payment leg as money or kind. This gives 15,053 legs dated to within twenty-five years, after excluding restored and low-confidence readings. Monetization predicts a rising money share through the third century. The share does not rise. It is flat across the Roman period and falls into the crisis, from 93.2% in 150–199 to 82.2% in 200–299 ($z = 11.3$). The fall survives every pre-200 baseline, an alternative cut of 150–234 against 235–284, and collapsing the data to one observation per document (Appendix B). A second test on 1,480 agricultural leases, in which rent is fixed either in drachmas or in wheat, does not corroborate the fall. In the leases the fall appears only against the local peak of 150–199. The evidence supports the negative claim, that monetization did not rise and so cannot be the source of the price rise. It does not support the stronger claim that the economy demonetized.

4.7 Fiscal evidence

The seigniorage argument of Section 2 has a fiscal counterpart in the same papyri, in 3,694 dated tax records from 1,310 documents. These records show two things. First, the nominal claims of the state were not indexed. The poll tax is the cleanest case. It was a fixed charge per head, levied at customary amounts of 4, 8, 12, or 16 drachmas that recur unchanged across the record. Across 179 receipts from 110 documents, the nominal payment shows no trend (-0.08 log points per century, $t = -0.38$, with place fixed effects), while its silver value falls significantly (-0.53, $t = -2.44$). The median receipt was worth 4.45 g of silver before 165 CE and 2.08 g after. Debasement thus eroded real tax revenue, the ancient counterpart of the Olivera–Tanzi effect, at the same time as it generated seigniorage from re-minting.

Second, the state stopped accepting money. The money share of tax records falls from 95.8% in 1–224 CE to 79.6% in 225–324 ($z = 12.9$). The fall is robust to every pre- and post-window we tried and to collapsing the data by document. It has no counterpart in market sales, where the share stays near 100%, and it survives the exclusion of the grain taxes. A state whose nominal receipts were fixed by custom and whose currency was depreciating had exactly this incentive. The village tax rolls behind the receipts record the tax base itself. Of 15,954 entries in 334 rolls, 117 taxpayers are entered as dead and 110 as fled. The Mendesian arrears roll of the 170s (*P.Thmouis* 1) reports villages that were once populous reduced to a handful of men by plague and flight.

4.8 Factor prices and the return on land

A natural objection is that all of this is specific to grain. The same documents answer the objection with factor prices. Nominal land rent rises +1.79 log points per century ($t = 4.9$) and nominal wages rise +0.59 ($t = 3.8$), while their silver-converted series show no significant trend. The most informative series is rent fixed in kind, which needs no silver conversion, no fineness series, and no assay. Across the Antonine window it falls from a median of 5.50 artabas per aroura in 45–163 CE (130 observations, 82 documents) to 3.60 in 175–230 (55 observations, 45 documents), a change of −34%. Harper's independently compiled file gives −29% over the same window.

The same construction can be applied to land as an asset. The gross yield on land is rent times the price of wheat, divided by the sale price per aroura, and the drachma cancels from this ratio. The yield runs at a pre-plague median of 14.5% (26 sales, IQR 9.9–21.9) and roughly doubles to 31% after the plague. The result is robust to matching tolerances from ten to twenty-five years and to dropping the two highest post-plague values. The mechanism is visible in the data. Land held its price in drachmas (the median rose from 300 to 325 per aroura) while wheat doubled, so land halved in terms of grain while its rent in grain fell by only 29%. These are real movements, and their size is what matters. They amount to tens of percent, against nominal movements of five to six orders of magnitude.

5 A model of the fiduciary equilibrium and its collapse

The third-century debasement opened a large wedge between the bullion base and the nominal money stock. The Egyptian evidence shows how the Roman economy sustained this wedge. Nominal prices were roughly stable through the debasement of the tetradrachm and then jumped abruptly when confidence broke in 274/5. This is the signature of a token currency held above its metal value by self-fulfilling beliefs. A simple model makes the mechanism precise, and it turns the regime switch into a result of comparative statics rather than a narrative.

5.1 Setup

A coin carries an intrinsic silver content c_t (grams), which declines with debasement, and buys g_t units of grain. The observed silver price of grain is $P_t = c_t/g_t$. The value of the coinage g_t can settle at one of two equilibria. In the commodity (fundamental) equilibrium, the coin is valued at its metal, $g_t = c_t/\pi$, where π is the metal-anchored real price, a fundamental with $\pi \approx 19$–25 g of silver per hectoliter. Then $P_t = \pi$. The silver price is constant regardless of debasement, and the nominal price $1/g_t = \pi/c_t$ rises one for one as c_t falls, so that debasement passes through fully to prices. In the fiduciary (token) equilibrium, the coin circulates above its metal value, $g_t = \bar{g} > c_t/\pi$. This equilibrium is sustained because each agent accepts the coin at $\bar{g}$ expecting others to do the same next period. The nominal price $1/\bar{g}$ is then nearly constant and does not rise with debasement, as in the observed Egyptian stability. The silver price $P_t = c_t/\bar{g}$ falls as c_t falls, because the coin trades above its shrinking metal content.

Both are equilibria of an economy in which money has value only because it is expected to have value (Kiyotaki and Wright, 1989). The fiduciary equilibrium requires a confidence (transversality) condition, namely that the coin will still be accepted tomorrow. Debasement erodes the metal floor but not the fiduciary value, so the token can drift ever further above its content. The third-century record shows exactly this pattern. The intrinsic silver price implied by nominal quotations falls to ∼4–6 g/hl while the metal in the coin collapses, and the coin circulates far above its content. A token coinage sustained above its intrinsic value by convention, and always exposed to the failure of that convention, is the ancient instance of the problem that Sargent and Velde (2002) trace across medieval and early modern Europe.

5.2 The premium and the price jump

The two equilibria can be compared through a single object. Let the fundamental (melt) value of the coin be $g_t^F = c_t/\pi$ and the fiduciary value be $\bar{g}$. The coin then trades at a *premium* over its metal of

$$\mu_t = \frac{\bar{g}}{g_t^F} = \frac{\bar{g}\,\pi}{c_t} > 1. \tag{2}$$

The fiduciary equilibrium exists whenever agents expect the coin to be accepted at $\bar{g}$ next period. This condition is self-referential, and a coordinating device can sustain it even as μ_t grows. Examples of such devices are the tariff and the willingness of the fiscal authority to receive the coin at face value in taxes. Debasement raises the premium mechanically. With $\bar{g}$ pinned by convention, a fall in the metal content c_t raises μ_t one for one, so the token drifts ever further above its metal. When confidence fails and expectations re-anchor on the fundamental, $\bar{g} \to c_t/\pi$, and the nominal price jumps discretely from $1/\bar{g}$ to π/c_t, that is, by exactly the factor μ_t. The price jump at a monetary reform should therefore equal the premium that the coin had accumulated before the reform.

With $\mu_t = \pi\bar{g}_t/c_t$ and $\bar{g}_t = A/P_t^{\text{before}}$, where A is hectoliters per artaba, the predicted jump is $\pi A/(c_t P_t^{\text{before}})$ and the observed jump is $P_t^{\text{after}}/P_t^{\text{before}}$. The pre-reform price cancels. Whether the prediction matches the observation does not depend on P_t^{before} at all, since halving or doubling it moves both sides in the same way. The comparison reduces to a single question. Does $c_t P_t^{\text{after}}/A$, the silver price of wheat after the reform, fall in the fundamental band π? It does at both dates, with 20.6 and 23.3 g/hl against $\pi \approx 19$–25. The post-reform price was free to land anywhere, and it landed inside the band. The test is therefore whether the reform restored the metal-anchored price, not an independent prediction of the size of the jump. The test also inherits the uncertainty about π, which we take from the literature and do not measure here. The Egyptian record provides two realizations, and we compute μ_t for both from the tetradrachm's own content and the fundamental $\pi \approx 19$–25 g/hl.

Reform	coin content c_t (g Ag/dr)	predicted μ_t	observed jump
Aurelian 274/5	0.040	11.5–15.1×	12.5×
Diocletian 296	0.014	1.7–2.3×	2.1×

The 296 jump falls inside the predicted band, near its center. The 274/5 row is reported for documentation and is discussed below. The ranges in the table hold the fineness of the coin at its mid estimate and vary only π across the band that the price record independently supports. If fineness is also allowed to vary over its full range, from minimum to maximum, the predictions widen to 9.2–30.3× and 1.3–3.5×. The 296 jump satisfies these wider bands too, but the test is then far less demanding. At both reform dates the fineness figures are inferred from the parallel antoninianus rather than assayed, so the denominator of μ_t is reasoned rather than measured. The 296 comparison involves no dispute over readings, and it is the test. A single episode cannot reject a model. It can only find the model compatible or not, and at 296 the model is compatible, near the center of the band. The smaller second jump is what the logic of the discharged premium predicts, since Aurelian's reform had already released most of the premium.

5.3 Why 274/5 is not a second test

Read in the same way, Aurelian's reform looks like a success, with 11.5–15.1× predicted against 12.5× observed. We do not accept this reading, because the event confounds the release of the premium, which the model captures, with two things that the model does not capture. First, the direction of the change in content is contested. Christiansen (2004), the source of the table, has the reformed tetradrachm more debased. However, the wet-chemical assays that Haklai-Rotenberg (2011) survey, following Cope et al. (1997), record a slight *rise* in silver in 274. Such a rise would cut the prediction to 4.2–6.6×, a factor of two to three below the observed jump. Second, the XXI/KA mark re-tariffed the unit of account, which is a fiat devaluation that the premium accounting does not price. We therefore report 274/5 as a dated regime switch whose magnitude the framework does not explain, and rest the quantitative claim on 296 alone. That the price explosion of the 270s was larger than a collapse of confidence can deliver is itself evidence that the re-tariffing raised prices on top of the collapse of confidence (Appendix B).

5.4 Mint standard and circulating coin

The c_t in the table is a mint standard, while μ_t is a premium over the metal in the coin that a buyer actually hands over. Section 6 measures the distance between the two for the coinage in which both can be observed, and the distance is large in both directions. The imperial circulating stock held about twice the fine silver of the mint standard in the 250s and 14.6 times it in 270, but only 0.39 times it in 280, six years after Aurelian *raised* the standard. Egypt admits no such correction. Its closed currency leaves about a hundred tetradrachms in a corpus of half a million coins, so only the direction of the error can be stated. At the ratio of 280, the post-reform silver price of wheat would be roughly 8 g/hl instead of 20.6, below the fundamental band. At the ratio of 270, it would be far above the band. Either substitution rejects the model. That the test passes at the mint standard is thus evidence that Egypt's reformed coin displaced the old stock quickly. Such a quick displacement is plausible in a closed currency area whose boundary the reform was dismantling, and the Egyptian hoard evidence assembled by Christiansen (2004) could settle the question.

5.5 What contracts say about the coin

The premium in equation (2) is not observed, but we can observe whether the parties to contracts saw it. We read every dated document in the Duke Databank that names a sum of money (12,482 documents; Appendix A) for clauses that qualify the coin. Among documents that state a money sum, the share that specify imperial coin (Σεβαστοῦ νομίσματος) had run at 3–7% under the Principate. It then rose from 3.5% in 150–199 to 7.2% in 200–249, 9.3% in 250–299, and 12.9% in 300–349. Clauses that distinguish old coin from new, rare before, appeared in 5.3% of documents in 250–299 and 2.3% in 300–349. In contrast, the letters and petitions are nearly silent. Between Augustus and the end of the fifth century, six documents refuse or distrust a coin. Only one of them falls in the third century. It is the order of 260 CE to the money-changers who had shut their banks rather than accept "the divine coin of the emperors" (*P.Oxy.* XII 1411). This sequence is the one that the model implies. Parties wrote the quality of the coin into their contracts at a rising rate through the half-century before the two coordination events, while open refusal stayed rare until the banks shut.

5.6 The collapse in two steps

The transition came in two dated steps. Aurelian's re-tariffing of 274/5 devalued the coinage and moved wheat about twelvefold to a new plateau, from 16 to ∼200 drachmas per artaba. However, the reformed tetradrachm, which by 279 held only ∼0.16 g of silver at mid-band, still circulated far above its metal, so the plateau was itself a fiduciary equilibrium. It held for two decades (200–300 drachmas in 277–294). Diocletian's abolition of Egypt's separate coinage in 296 was the terminal coordination shock. It made the metal fundamental salient and re-anchored expectations on the imperial denarius. Formally, $\bar{g} \to c_t/\pi$ raises the nominal price from $1/\bar{g}$ to π/c_t. With c_t minuscule, this is the second move, from 300 to 640 drachmas and to the Edict's ceiling of 1,200–1,333. The second step is larger than Aurelian's twelvefold jump in absolute drachmas, but smaller as a ratio (∼2–4×), and equation (2) is a statement about the ratio. The premium μ_t had been partly discharged in 274/5, so the premium accumulated by 296 was smaller, and a smaller premium implies a smaller proportional release. The release restores the silver price to the fundamental $\pi \approx 19$–25 g/hl (median ∼19, up to the Edict end of the range), on which the Edict and Harper agree. The inflation thus came as two coordination events, twenty-two years apart, rather than as one smooth pass-through.

5.7 Implications

The model has three implications. First, the great Egyptian inflation was not a smooth monetary pass-through but a *coordination* event. Decades of latent inflation, stored in the gap between the token value and the metal content, were released at once when confidence failed. Second, the model bounds how far debasement can run before the fiduciary equilibrium becomes fragile. This gives a testable link between the depth of debasement and the hazard of a collapse of confidence, and it connects the episode directly to the literature on self-fulfilling crises (Obstfeld, 1986).

The third implication is the sharpest. The two equilibria disagree about sorting, and this disagreement can be tested on coins. Take two coins of the same denomination struck to different standards, for example a denarius struck before 194 and one struck after. In the commodity equilibrium each coin is valued at its metal. The finer coin is then worth more than the debased one at the same face value, and it is withdrawn. This is Gresham's law, and it implies that the finer cohort should disappear from circulation faster. In the fiduciary equilibrium both coins pass at $\bar{g}$ because both are accepted at $\bar{g}$. The premium μ_t is simply larger on the debased coin, and as long as the premium on the finer coin is also positive, no one gains by sorting. A gradient by fineness in the rate at which cohorts leave circulation is therefore a signature of the commodity equilibrium, and its absence is a signature of the fiduciary one. Velde et al. (1999) formalize this distinction as circulation by weight against circulation by tale. It is one of the few propositions in monetary theory that the coins themselves can settle, because it is a prediction about coins and not about prices. Section 6 tests it on the largest fineness break in the imperial record and finds no gradient. If anything, the finer cohort gains. We read this as evidence that the coinage passed by tale. An alternative reading points in the same direction. The mint blanched debased blanks precisely so that they could not be told apart, so agents may have been unable to sort rather than unwilling to. The evidence cannot separate the two readings, but both imply that the metal in the coin had stopped governing its acceptance. The evidence does exclude one interpretation of the third-century hoarding spike documented in Appendix B, namely that it was a flight to good metal. The burials were large and concentrated in the decades of invasion, but they were not sorted by fineness.

6 What was actually in circulation

Everything so far has divided by the silver content of the coin, and the series available for this purpose is a chronology of *mint policy*. It records what the mint struck in a given year. The arguments need a different object. A price is quoted in the coin that someone hands over, the inflation tax falls on the coin that someone holds, and the fiduciary premium of Section 5 is a premium over the metal in that coin. In an economy that recoins instantly, the two objects coincide. Rome did not recoin instantly, and this section measures how far apart they were.

6.1 Construction

Beneath its record of hoards, Coin Hoards of the Roman Empire publishes the composition of a subset of deposits as coin groups, each with a denomination and a range of striking dates. Flattened, this layer has 139,494 rows covering 521,950 coins. We map each row to a silver denomination, take the midpoint of its striking range, and assign the fineness in force in that year from a step table of 54 steps built from Harl (1996), Walker's *Metrology*, and Butcher and Ponting (2015) (Appendix A). Fineness is a step function between reforms and is never interpolated across one. Joined to the terminal dates of the hoards, the layer yields an analysis panel of 253,417 dated silver coins in 901 hoards, 77.2% of the in-scope silver coins. The rest are mostly undated generic "AR" rows in catalogs, which carry no striking date and cannot enter any series by vintage. The bootstrap estimation frame behind the estimates below retains 253,301 of these coins. We define

$$\bar{c}_t \;=\; \frac{\sum_{h:\,|t_h - t| \le 10} \;\sum_v q_{hv}\, c^m_v}{\sum_{h:\,|t_h - t| \le 10} \;\sum_v q_{hv}}, \tag{3}$$

the quantity-weighted mean of the mint content of each coin's striking vintage, over the coins in hoards buried within ten years of t, on five-year centers from 30 BCE to 310 CE. We report it in two ways, weighted by coins as written and with equal weights on hoards. In the second version each hoard contributes its own mean once, so that a single enormous deposit cannot dominate a bin. The 95% intervals come from 999 bootstrap resamples of the 901 hoards.

6.2 The mint standard and the circulating stock

Figure 5 and Table 2 report the results. The circulating stock stood at 79.5% fine in 160 and was still at 74.9% in 200, when the mint had been striking at 47.9% for six years. At the Severan halving of 194 the gap was 29 percentage points, and half of it closed only after 36 years [33, 41], around 230 CE. Measured as fine silver per denarius of account, the two series diverge. Between 160 and 270 the mint standard falls by a factor of 70, and the circulating stock by a factor of 5.4. The collapse of the stock itself happens in 250–280.

The lag works in both directions. Aurelian's reform of 274/5 *raised* the standard, and the stock was as slow to follow upward as it had been downward. In 280 the coin in circulation held 0.076 g of fine silver per denarius of account against a mint standard of 0.195 g. The ratio to the standard was 0.39, whereas ten years earlier it had been 14.6. Whether a content based on the mint standard overstates or understates the metal in circulation thus depends on the side of a reform one looks at, and the error is of order ten in both directions.

6.3 Why the stock lags

Panel A of Figure 6 shows the hoarded silver stock as a matrix of striking decade against deposition decade. Until the 230s the mass lies far to the left of the diagonal, since the coin being buried had been struck decades earlier, and only after 240 does it collapse onto the diagonal. The mean age of

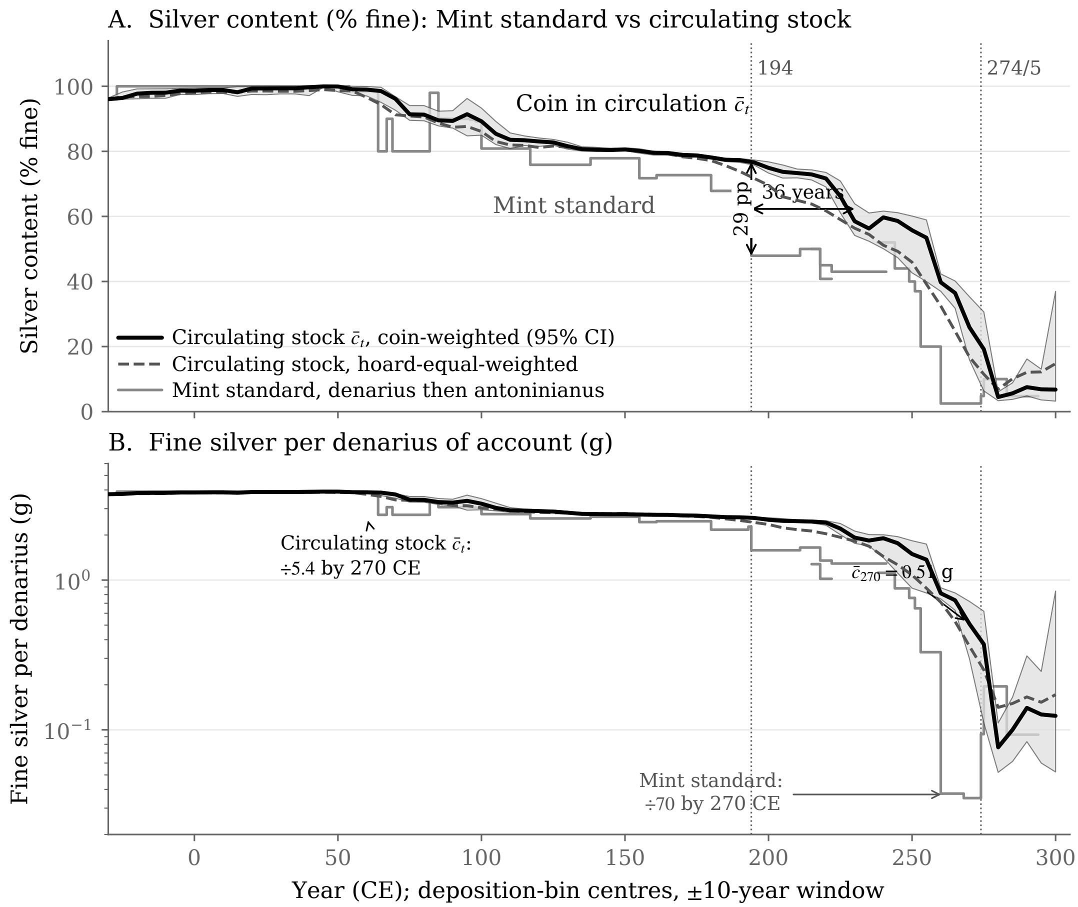


Figure 5: **The mint standard and the coin in circulation.** Panel A shows silver content in percent fine, comparing the mint's step function with $\bar{c}_t$ from equation (3), with the 95% hoard-bootstrap band. At the Severan halving of 194 the circulating stock stood 29 points above the new standard and closed half the distance only after 36 years. Panel B shows the same series as fine silver per denarius of account, on a log scale. Between 160 and 270 the mint standard falls by a factor of 70 and the circulating stock by a factor of 5.4. The lines cross after Aurelian's restoration of 274/5, and for a decade the circulating stock is the *coarser* of the two. Source: CHRE coin-level layer, 253,417 dated silver coins in 901 hoards.

a circulating coin was 51.6 years [44.9, 56.3] in 160 and 59.0 [52.8, 68.1] in 200, and it fell to 16.7 [10.6, 25.1] by 270. In 200 CE, six years after the Severan halving, the share of the circulating stock struck on the standard then in force was 2.9% [0.8, 8.9] weighted by coins and 23.5% [17.0, 31.0] with equal weights on hoards, measured on hoards buried within ten years of that date. The typical hoard held about a quarter Severan coin, while the mass of coins stayed old. The two measures describe the same fact with different weights. New coin is a flow into a large inherited stock, and a flow takes a generation to change an average.

This result gives a quantitative basis to a proposition that the numismatic literature has long asserted but never measured (e.g., Howgego et al., 2010), namely that debasement took decades to reach the coin in use. It also gives the revisionist case of Section 4 a mechanism that does not

Table 2: Mint standard against circulating stock. Values are fine silver per denarius of account, so the antoninianus, which was tariffed at two denarii but never contained two denarii of silver, is counted at its value in account. The mint column refers to the coin that supplied most new silver in each year, the denarius through 237 and the antoninianus from 238. Values at 194 are interpolated between five-year bin centers (the interpolation applies to the estimate, never to the fineness). 95% hoard-bootstrap intervals in brackets.

	mint standard		circulating stock $\bar{c}_t$		stock /
Year	% fine	g/den.	% fine, coin-wtd [CI]	g/den.	mint
160	71.7	2.44	79.5 [78.9, 80.0]	2.73	1.12
194	47.9	1.58	76.7 [75.9, 77.4]	2.61	1.65
200	47.9	1.58	74.9 [73.3, 76.8]	2.53	1.60
215	50.0	1.65	72.9 [71.2, 74.3]	2.46	1.49
250	40.0	0.76	55.7 [42.7, 60.1]	1.49	1.96
270	2.5	0.035	25.9 [15.9, 35.3]	0.51	14.55
280	10.0	0.195	4.5 [3.4, 6.1]	0.076	0.39

require monetization. For a generation after a cut in the mint standard, most of the coin used in payments had been struck before the cut. That alone is a sufficient reason for prices not to respond, and it requires no change in the extent of the money economy. Nothing here revives the hypothesis of Rathbone (1996), which the evidence on payment legs rejected. The measurement is imperial. Egypt's insulation rests on its closed currency rather than on this lag, and the corpus cannot tell whether the tetradrachm stock lagged its own mint in the same way.

6.4 A test of Gresham's law

The same coins test the sorting prediction of Section 5. Old fine coin certainly *left* circulation. On 628 hoards and 241,418 dated coins, the share of the pre-194 vintages in the hoarded stock falls from 0.852 in the 190s to 0.165 in the 230s and 0.053 in the 250s. The question is whether it left *because* it was fine.

One natural test compares this decay with a counterfactual in which new coin merely dilutes old coin. This test identifies nothing. Under a simulated null with no preference for fineness, the same estimator returns an apparent excess decay of +0.7 to +2.0 points a year. The reason is that an accumulated hoard is a lagged average of past circulation rather than a snapshot of it (Appendix B).

The prediction of the theory can instead be tested without any benchmark, because two cohorts inside one hoard share a denominator. The reform of 194 cut the fine silver per coin from 2.36 to 1.59 g, the largest discontinuity in fineness in the window. Across this reform, the ratio of the finer cohort to the debased one, measured across deposition decades, moves at +0.0088 per year. On a narrow comparison of 184–193 against 194–203 it moves at +0.0082. Gresham's law predicts a negative number, but the finer cohort gains (Figure 6, Panel B). The estimate rests on six deposition decades and is imprecise. The standard error is 0.0116 and the 95% interval runs from −0.024 to +0.041, so the test excludes a finer cohort leaving more than about two percent a year faster than the debased one. It cannot detect a weaker gradient. What the test does establish is that the

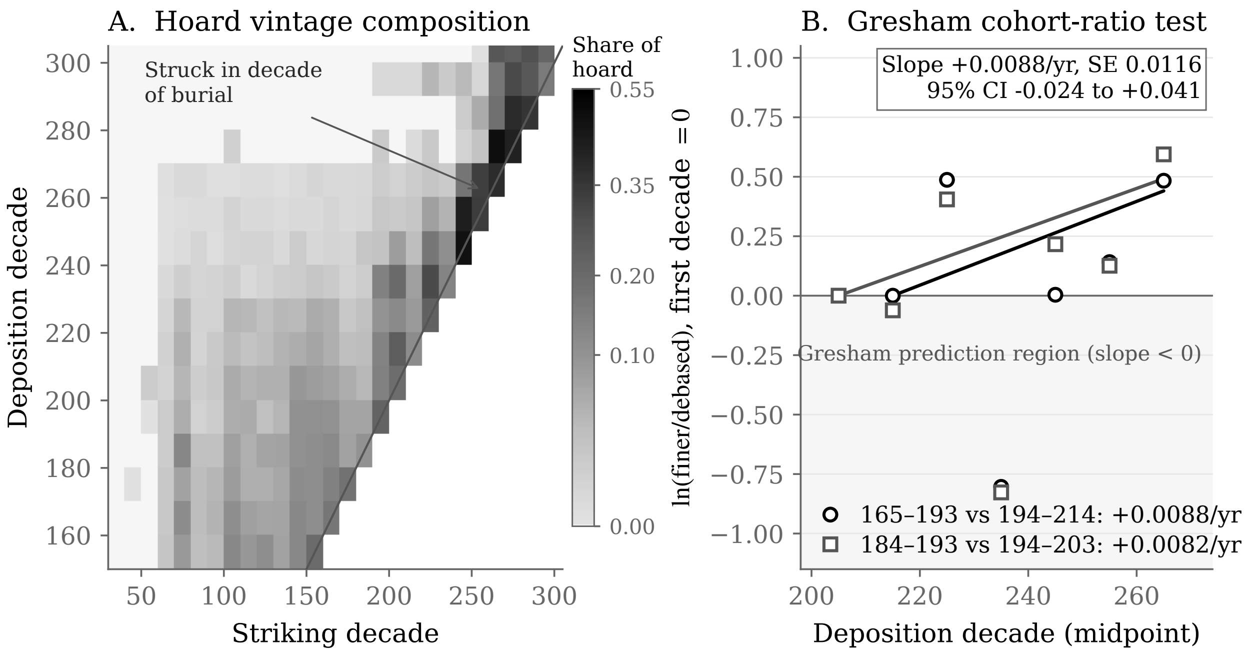


Figure 6: **Why the stock lags, and the test that Gresham's law fails.** Panel A shows the hoarded silver stock by striking decade and deposition decade (shares with equal weights on hoards; each deposition decade's row sums to one across all striking decades, and the display trims issues before 30 CE, at most a tenth of any row). Mass to the left of the diagonal is coin struck before the decade in which it was buried. Until the 230s almost all of the mass lies there, and only after 240 does the stock collapse onto the diagonal. Panel B shows the sorting test, which needs no benchmark. It plots the log ratio of a finer to a debased cohort held in the *same* hoards, normalized to zero in the first deposition decade in which both are present. Gresham's law predicts a downward path, with the finer coin withdrawn faster, and the shaded half is the region it predicts. Across the reform of 194, the largest fineness break in the record, the fitted slope is +0.0088 per year with a standard error of 0.0116 on six deposition decades. The interval excludes a strong gradient and cannot detect a weak one. Source: CHRE coin-level layer; 628 hoards and 241,418 dated silver coins in Panel B.

pre-194 stock decays as a block, at a rate uncorrelated with its own fineness, and that it is removed by replacement rather than by selection. After 238 the denarius is effectively no longer struck, and the share of the antoninianus in newly struck silver entering hoards goes from 0% in 222–237 to 98.6% in 238–259.

The within-silver version of Gresham's law therefore finds no support, and the model that motivated the test survives. The result is less surprising in light of Rolnick and Weber (1986), who found the law failing across the historical record wherever good money could not command a premium. It is also a direct observation of the equilibrium by tale of Velde et al. (1999). A coinage in which the metal no longer governs acceptance should show no gradient by fineness in its withdrawal, and this coinage shows none. Two qualifications apply. First, "would not sort" and "could not sort" are observationally equivalent here, since the mint blanched debased blanks precisely so that they could not be told apart. Both readings imply that the coin passed by tale. Second, the Caracallan window of 215–255 retains roughly 0.9 to 1.3 points a year of excess decay after the correction for the bias under the null, and this remains when one bin at a time is left out. Nothing below rests on this window.

6.5 Mines and mints

Section 2 assumed that bullion and fineness move for unrelated reasons. The circulating series allows us to check this assumption against lead flux at annual frequency ($N = 330$, 30 BCE–310 CE). The correlation is $+0.43$ in levels, where both series fall across the third century and no test can separate a common collapse into its causes. In first differences it is $+0.01$, and no lead or lag from -20 to $+20$ years exceeds 0.03 in absolute value (Appendix B). Mines and mints are unrelated at every frequency at which this can be checked.

6.6 Scope

Two limits apply to the series. First, it is an imperial measurement. Egypt's closed currency leaves about a hundred tetradrachms among half a million coins, so $\bar{c}_t$ is a series for the denarius and the antoninianus. The Egyptian counterpart that Sections 4 and 5 would most like to have does not exist in this corpus. Second, hoards record the coin available to hoarders rather than the stock used in transactions. If hoarders held back good coin, $\bar{c}_t$ overstates the fineness of the coin that circulated. The cohort test above bears directly on this question, and it finds no such selection at the precision it has. Appendix A sets out the coverage, the two weightings, the convention for striking dates, and the weight of the Reka Devnia deposit.

7 The quantity theory and velocity in the very long run

The quantity identity can be closed at the scale of the empire, although this settles less than one would like. The money stock here is bullion, the perpetual inventory of equation (1), which never divides by a fineness. As a result, nothing in this section changes when Section 6 replaces the mint standard with the circulating one. The nominal coin stock, which does change, is treated in Appendix B. The sample consists of 21 twenty-five-year bins spanning 200 BCE–400 CE, and it is thin in two ways that limit what can be asked of it. First, on 21 points the cointegration tests (Engle–Granger, Phillips–Ouliaris, and the error-correction coefficient of a VECM) have almost no power. A failure to reject the null of no cointegration is therefore uninformative rather than a puzzle for time aggregation to explain, and an error-correction speed estimated after that failure presupposes what the test does not support. In their place we report the two descriptive quantities that the identity actually contains, each with a moving-block bootstrap interval (block length 4, 5,000 resamples). Second, the price series is a splice. Of the bins, 15 carry Egyptian wheat and 6 carry Babylonian barley from the Seleucid record (van der Spek and Mandemakers, 2003). No bin carries both, and a gap of a hundred years with no price evidence at all separates the two blocks.

On Egyptian evidence alone the exercise spans the 15 bins from 25 to 400 CE, which is under four centuries rather than six. Table 3 reports both scopes.

Table 3: Money, prices, and velocity: bootstrap intervals over the leakage rate and the scope of the price series. 95% moving-block bootstrap in brackets.

	All bins ($n = 21$)		Egypt only ($n = 15$)	
δ	corr($\log P, \log M/Y$)	V trend/century	corr	V trend/century
0.000	+0.29 [−0.20, +0.71]	−0.17 [−0.40, +0.14]	−0.02 [−0.64, +0.48]	−0.37 [−1.10, +0.95]
0.005	+0.07 [−0.45, +0.51]	+0.03 [−0.21, +0.33]	−0.15 [−0.57, +0.22]	−0.13 [−0.80, +1.23]
0.010	−0.25 [−0.57, +0.24]	+0.12 [−0.14, +0.51]	−0.33 [−0.74, −0.03]	+0.00 [−0.75, +1.42]
0.020	−0.58 [−0.73, −0.20]	+0.19 [−0.09, +0.64]	−0.56 [−0.75, −0.26]	+0.09 [−0.85, +1.75]
0.050	−0.69 [−0.80, −0.42]	+0.22 [−0.09, +0.73]	−0.68 [−0.75, −0.37]	+0.12 [−1.09, +2.02]

Three results follow, and only the first is favorable to the framework of Section 2. First, the flow is disinflationary. The correlation between the bullion flow and prices is −0.45 on this panel, and it does not depend on δ because leakage enters only the stock. The magnitude is not robust to the choice of panel. The correlation falls to −0.27 without the Babylonian splice, to −0.17 and −0.19 on the papyrus wheat panels, to −0.10 on a four-commodity index, and to −0.04 once wages are added. The sign is negative throughout, and the argument rests on the sign. Second, the stock margin points the wrong way. Over the range of leakage rates implied by our own calibration of the drain ($\delta \approx 0.005$–0.02), corr($\log P, \log M/Y$) runs from +0.07 [−0.45, +0.51] to −0.58 [−0.73, −0.20]. It is indistinguishable from zero at one end of that range and significantly wrong-signed at the other, on both scopes. The +0.29 at $\delta = 0$ is not a competing estimate. It is an artifact of assuming that silver never left circulation, which the Indo-Roman and hoard evidence rules out. Time aggregation will not recover the sign that the theory predicts either, since attenuation shrinks an elasticity toward zero rather than driving it through zero to −1.2. Third, velocity is trendless. The bootstrap interval on the trend contains zero at every rate on the grid and on both scopes, so the δ that minimizes the point estimate of the trend is minimizing noise. What the binned record does show is that the bullion stock and Mediterranean trade did not visibly diverge across six centuries. The record shows consistency with the quantity theory, not an elasticity, and 21 bins of a spliced price series will not yield an elasticity.

The leakage rate is calibrated from Pliny's eastern drain against estimates of the imperial silver stock. Once wear and permanently buried hoards are added, the calibration gives $\delta \approx 0.005$–0.02 a year. The hoard record shows that the rate was not constant, since 57% of all coin hoarded in 0–400 CE falls in 225–300. This hoarding is the fall in velocity during the crisis. Entered as a structural leakage, it breaks the long-run consistency, which is what should happen if hoarding responded to inflation rather than drove it. The Indo-Roman coin finds cannot bound the drain during the crisis (Appendix A).

7.1 Output and velocity

Closing $MV = PY$ requires an output series at the same scope as the money stock. Lead is empire-wide, so real activity Y comes from the Mediterranean shipwreck series, the standard empire-wide proxy for trade, which is independent of both lead and prices. We do not use Egyptian population, which would mismatch the scope and leave velocity uninformative. When we back out $\log V = \log P + \log Y - \log M$ across the 21 bins, the trend in velocity is smallest in absolute value at $\delta \approx 0.005$ (0.03 per century), and this rate falls inside the range that the eastern drain independently implies. A finer grid puts the minimum at $\delta \approx 0.0025$, below the calibrated floor. This coincidence is not identification. Here $\log V$ is an accounting identity rather than an estimate, its only empirical content is whether the implied series drifts, and on 21 bins that question cannot be answered.

This consistency result is the within-economy analog of the cross-country regularity documented by Lucas (1980) and McCandless and Weber (1995). Our design follows one economy across six centuries, rather than many economies at a point in time. The two designs share a limitation, since neither identifies short-run neutrality. They differ in what they recover. The cross-section recovers a coefficient of proportionality, and this design recovers none. It establishes that the money base and the volume of trade did not diverge structurally, not that prices responded to money one for one.

8 Conclusion

This paper studies money and prices in the Roman economy with four series: ice-core lead, papyrus prices, coin fineness, and the silver in circulating coin. The strongest of our four results is the fiduciary reading of the Egyptian inflation. A token coinage was held for decades above its metal value and released at two datable coordination shocks. The two-equilibrium model of Section 5 predicts that the price jump at the release should equal the premium that the coin had accumulated. Computed from the tetradrachm's own content, the prediction for 296 is 1.7–2.3×, against 2.1× observed. The pre-reform price cancels from this comparison (Section 5), so the test checks whether the reform restored the silver price of wheat to the assumed fundamental band, and it did. The re-tariffing of 274/5, whose change in content is contested and whose change in face value the model does not price, is reported as a dated regime switch rather than a second test. The logic of the discharged premium still explains why the second jump was proportionally smaller. The result relies, however, on a mint standard standing in for the metal in circulating coin. Section 6 shows that this substitution is off by a factor of ten in either direction for the coinage in which it can be checked. The result should therefore be read as conditional on Egypt's reformed coin having displaced its old stock quickly.

The second result is that the coin in circulation is not the coin the mint struck, and the difference is large enough to change the timing of the whole account. Measured on 253,417 dated silver coins, the circulating stock stood 29 points of fineness above the Severan standard of 194, and it closed half of that gap only after 36 years. Between 160 and 270 the mint standard fell by a factor of 70, and the money in use by a factor of 5.4. Every aggregate in the literature that divides by the mint's fineness is therefore wrong in timing and not only in level. The same coins test the sharpest prediction of the theory, that fine coin was driven out because it was fine, without any benchmark. They find no such gradient at the precision the hoards allow. This is what the model predicts. A coinage that passes by tale should show no gradient by fineness in its withdrawal, and this one shows none.

The third result is that the great inflation was overwhelmingly monetary. Across seven centuries the silver price of grain mean-reverts while nominal prices rise five to six orders of magnitude. In the fourth century, when the content of the coinage genuinely varies, the conversion returns a real price inside the band it occupied nearly three centuries earlier. The fourth and weakest result is that the bullion flow is disinflationary, so lead is predominantly a series of real bullion output. This answers the question of whether lead measures output or money, and it is the fact on which the reading of Alonso Ortiz and Da-Rocha (2026) rests. The relation between the stock and prices remains the fragile part of the picture. At 25-year bins, the elasticity of prices with respect to money is insignificant at the leakage rate that our own calibration prefers, and significantly negative at the other end of that range. The stock margin supports consistency, in the form of a trendless velocity at a plausible drain, but not an elasticity.

This paper takes a first step toward an annual test of the quantity theory over several centuries in an economy with commodity money, with a debasement natural experiment that separates the real and monetary sides of money and a model of fiduciary collapse for the form the inflation took. Three pieces are still missing. The first is a papyrus price panel that is exact-dated and dense enough to estimate an elasticity at annual frequency. The second is a series of mint output to replace the lead-flux proxy for the money base, which the die studies of Appendix B are beginning to make possible. The third is an Egyptian counterpart to $\bar{c}_t$, which the tetradrachm hoard record could supply and the imperial corpus used here cannot. We do regard three facts as established. First, the bullion flow is disinflationary, and hence predominantly real. Second, the great inflation layered on top of it was overwhelmingly monetary in the one century in which the content of the coin can be measured independently. Third, the money in which that inflation was denominated reached the debased standard a generation after the mint did.

These results have three broader implications. For method, they show that a paleo-archive can be read as an economic time series with the tools one would bring to a modern national account. Doing so brings quantitative evidence to the debate over whether the inflation was monetary or real, which the written record alone left open for a century. For monetary economics, they add to the record on the quantity theory the observation it most conspicuously lacked, an economy with commodity money in which the money base is measured rather than inferred, even though at generational resolution the long-run relation between money and prices is not identified. For the reading of the Roman business cycle in Alonso Ortiz and Da-Rocha (2026), they supply the fact on which that reading rests, namely that the bullion flow is disinflationary and hence real. The boundary between the two is equally clear. The inflation layered on top of that real cycle was monetary through and through. The Roman economy generated its fluctuations endogenously and paid for its wars by debasing its coin. The ice-core record measures the first of these facts, and the papyri measure the second.

A Data construction

A.1 The assembled series

The money base is Greenland lead flux, annual, 200 BCE–2014 CE (McConnell et al., 2018, 2019). Behind it sits a compilation of ice-core lead records from 25 sites, drawn from twelve publications. The compilation supplies both the regional variants and the falsification test below, with four Arctic cores and one core in the Bolivian Andes that no plausible transport connects to an Iberian smelter. The price level is the silver price of grain. It comes from Babylonian barley ($n = 132$; van der Spek and Mandemakers, 2003) and from Egyptian wheat in Harper's full DARMC geodatabase "Prices, Wages, and Rents for Roman Egypt, AD 1–700" (DARMC, 2016), which we assemble in full. The geodatabase gives 104 dated quotations of the price of wheat (102 with a silver conversion) for 45–650 CE. The quotations report the nominal price (drachmas, then denarii, per artaba) and, for 102 of them, the silver-equivalent price (grams of silver per hectoliter), and the debasement experiment turns on the distinction between the two. Companion series for land prices ($n = 67$), wages ($n = 258$), rents ($n = 214$), and Egyptian population ($n = 28$, 14–575 CE) extend the record to factor prices and supply Y_t for the quantity identity. The grain entries of Diocletian's Price Edict (301 CE; Allen, 2009), an empire-wide administered ceiling in the reformed western coinage, give an independent anchor outside Egypt at the end of the crisis.

Circulation and debasement are measured separately. The CHRR+CHRE coin hoards (CHRR, 2026; CHRE, 2026) (17,989 deposits, of which 17,499 are imperial and 490 Republican) serve as a proxy for hoarding and dishoarding, and their spikes during the crisis identify the hoarding component of δ. A sourced reign-level series of the silver fineness of the coinage, from Augustus to Diocletian, combines Harl (1996), Walker's *Metrology*, and the drilled-core bulk assays of Butcher and Ponting (2015). A parallel series covers Egypt's own coin, the billon Alexandrian tetradrachm (Christiansen, 2004; Howgego et al., 2010). Together these series deliver the wedge between the *coin stock* M_t^{nom} and *bullion*, since a debasement raises nominal money without new silver. The fineness series is a chronology of *mint policy*. Beneath the hoard counts there is a further layer, used here for the first time. For a subset of its deposits, Coin Hoards of the Roman Empire also publishes the composition of each hoard, coin group by coin group, with the denomination and striking date of each group. Flattened, this layer has 139,494 coin-group rows covering 521,950 coins. Joined to the fineness step table, it yields 253,417 dated silver coins in 901 hoards, each carrying the fineness in force when it was struck. These coins are the raw material for $\bar{c}_t$, which Section 6 builds. Three further anchors are documented below. They are a series of the price of gold in denarii from the Principate through the fourth century (Bagnall and Bransbourg, 2019), the military *stipendium* as a dated anchor for nominal wages through the third century, and a dated chronology of the grain dole (*frumentationes*) as an annual regime variable for price controls and fiscal outlays.

A.2 An ice-core control from the Andes

The identification in this paper and in Alonso Ortiz and Da-Rocha (2026) rests on a chain of steps. Roman silver was won by cupellation, cupellation released a lead aerosol, the westerlies carried the aerosol to the Arctic, and the ice preserves it in datable layers. The chain has a built-in falsification test. If Arctic lead records European smelting, it should show the Roman arc, and a glacier outside the European airshed should not.

We run this test on a compilation of ice-core lead measurements from 25 sites, assembled from twelve publications (McConnell et al., 2018, 2019; Eichler et al., 2015). Five of the sites carry usable coverage of 0–700 CE. Four of them are Arctic, namely NGRIP2 and NEEM in Greenland, Tunu2013 in the northeast of Greenland, and Akademii Nauk on Severnaya Zemlya, four thousand kilometers to the east. The fifth is Illimani, a Bolivian glacier at 16.6°S and 6,300 m. The series are decadal means of log lead concentration, standardized within site over the common window, and Figure 7 plots them.

The Arctic cores tell one story. Each stands high through the Principate and falls across the second and third centuries. From the Principate (0–164) to the crisis decades (250–299), they drop by 1.9, 1.4, and 2.6 standard deviations at NGRIP2, Akademii Nauk, and NEEM, against 2.2 in the headline series. Their correlations with that series over seventy decades are +0.91, +0.51, and +0.70. Illimani shows neither pattern. Its correlation with the headline series is −0.18 in levels and +0.13 differenced, and across the same comparison of the Principate with the crisis it moves by −0.01 standard deviations in concentration and −0.17 in crustal enrichment. The Roman collapse is absent from the Andes.

Two qualifications keep this result in proportion. First, the four Arctic cores are the constituent records behind the headline series rather than independent confirmations of it, so their agreement is a matter of internal consistency. The independent evidence is Illimani, which comes from a different research group and from the other hemisphere. Second, the Andean core is not simply flat. It rises markedly after 450 CE, and Eichler et al. (2015) attribute that rise to metallurgical processing for *silver* production under the Tiwanaku and Wari cultures. This is what makes the null result informative rather than merely negative. The same core, read in the same way, detects pre-modern silver smelting when smelting happened within its airshed, and it detects nothing during the four centuries when Rome was the largest producer of silver on earth. Antarctica would be a cleaner control still, but EPICA Dome C carries only two observations in the window and cannot serve.

A.3 Fineness series

The reign-level series of silver fineness runs from the denarius (to Caracalla), through the antoninianus (from 215 CE), to the argenteus (from 294). It combines Harl (1996) and Walker's *Metrology* for the continuous series of surface measurements with Butcher and Ponting (2015) for the drilled-core

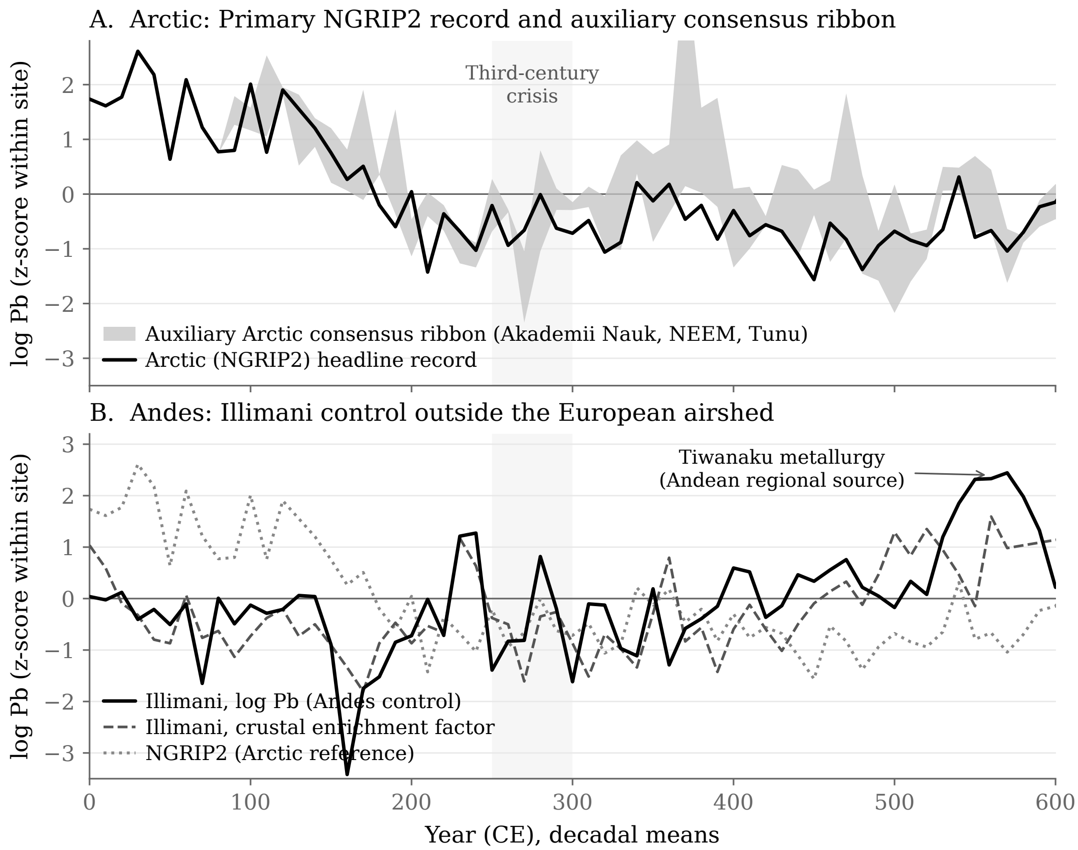


Figure 7: **The Roman signal, and where it is absent.** Decadal means of log lead concentration, standardized within site over 0–700 CE. The band marks the third-century crisis. Panel A shows the four Arctic cores behind the headline series, which all stand high under the Principate and fall together across the second and third centuries. Panel B shows the Illimani core in the Bolivian Andes, in concentration and in crustal enrichment factor, against NGRIP2 for reference. The Roman arc is absent. Illimani's own rise after 450 CE reflects the Tiwanaku and Wari silver metallurgy identified by Eichler et al. (2015). The core detects pre-modern silver smelting in its own airshed, and none from Rome. Sources: McConnell et al. (2018, 2019) for the Arctic, Eichler et al. (2015) for Illimani.

bulk figures. The bulk figures run ∼10–15% below Walker's surface measurements because of deliberate blanching, that is, the surface enrichment of debased blanks. Fineness collapses from ∼97% under Augustus, through Nero's reform of 64 CE and the Severan halving (∼50%), to the nadir of the third century (Gallienus ∼5%, Claudius II ∼2.5%), before the Aurelianic and Diocletianic reforms. The series delivers the mint standard c_t^m of Section 2. It also stands in for the content c_t of the coin in circulation in the model of Section 5, imperfectly, as Section 6 shows.

Egypt's closed currency was the billon tetradrachm (= 4 drachmas). Its measured silver content is ∼23% under Tiberius, 12–20% from Nero to Hadrian, and 14–20% under Pius. It then crashes to ∼4–5%, beginning in *164/5 CE*, before the Antonine Plague reached Egypt, and taking hold by 167/8. A partial recovery to 8–10% (half the old standard) follows from 169/70, and after 259/60 the coin collapses in parallel with the antoninianus (Christiansen, 2004; Howgego et al., 2010).

The figures are based on laboratory analyses up to the Antonines and are inferred thereafter, in parallel with the antoninianus. The post-260 values in particular are inferred rather than assayed, and results that rest on them are characterizations consistent with the numismatic literature, not estimates.

A.4 The fineness step table and the coin-level hoard layer

The reign-level series above is the input to a step table, which is used wherever a coin must be valued by its striking date. The table has 54 intervals [valid from, valid to) for the denarius, antoninianus, argenteus, and tetradrachm. Each interval carries a weight, a percent fineness on the Walker surface line and on the Butcher and Ponting (2015) bulk line, and the implied fine grams. Each coin takes the interval that contains its striking date, and nothing is interpolated across a reform, since linear interpolation between reigns would manufacture a smooth debasement out of a sequence of discrete political decisions. The table restores three corrections that the raw sources omit. These are Trajan's recall at AD 100, the documented decline of the tetradrachm in 185/6–192 (Section 4), and the wedge between the Walker and bulk measurements, which is ∼4 percentage points and is carried as two parallel columns rather than averaged away.

The coin-level layer behind Section 6 is the composition record that Coin Hoards of the Roman Empire publishes for a subset of its deposits. It covers 5,100 hoards, 139,494 coin-group rows, and 521,950 coins. Each row carries a denomination string, a range of striking dates, a quantity, and flags for imitations and intruders. Mapping the rows to a silver series recovers 329,251 coins (63.1%). The largest unmapped block consists of 106,097 coins entered as generic undated "AR" in summary catalogs, which no method based on vintages can place. We then join the layer to the hoard record, restrict it to deposits cataloged as certain or probable hoards, require a deposition date, and exclude imitations (0.24%), intruders (0.003%), and 73,263 coins with no striking date. The resulting analysis panel has 253,417 coins in 901 hoards. Deposition is the midpoint of the range of terminal years (median spread two years), and striking is the midpoint of the striking range. The sensitivities below vary the convention for striking dates. Five rows carrying six coins with impossible striking dates, either after their hoard was buried or before 160 BCE, are dropped as data-entry errors. One of them is a striking date entered as 2001 instead of 201.

A.5 Sensitivities of $\bar{c}_t$

Four sensitivities matter. The first is coverage. A fineness is assigned to 77.2% of in-scope silver coins. Another 22.3% are undated generic "AR" rows in catalogs, which carry no striking date and which no method based on vintages can place. The remaining half of one percent are the imitations, intruders, and out-of-step coins listed above. The second is weighting. The series weighted by coins and the series with equal weights on hoards agree closely before 230 and diverge in the mid-third century. In that period the Reka Devnia deposit contributes 79,766 panel coins and makes up 63% of the 250 bin on its own, where the estimate is 55.7% fine weighted by coins against 45.9% with equal weights on hoards. This is why Section 6 quotes both weightings for every statement about the mid-third century. The third is the striking date. Taking the start or the end of each striking range instead of its midpoint moves the series by at most 1.2 points at 160, 200, 250, and 270, and by at most 2.8 points anywhere. The fourth is the window. At 250 the estimate is 59.4% with a window of ±5 years, 55.7 with ±10, and 52.6 with ±20, as giant hoards move in and out of the window. At 160 and 200 the three windows agree within 1.5 points. Bins after 295 rest on eighteen hoards or fewer, with intervals to match, and nothing in the paper is read off them.

A.6 The decomposition of the 150s–190s in detail

The fourth-century decomposition of Section 4 is clean because the content of the coin varies enormously in that century. The first three centuries are harder. The standard silver conversions in the literature (e.g., Harper 2016) are essentially the nominal series rescaled. They use a nearly constant parity of ∼0.97 g of silver per drachma, that of the Augustan imperial denarius, and ignore the physical debasement of the Alexandrian tetradrachm, so they cannot separate real from monetary shocks. We close this gap from the numismatic side, using the tetradrachm's own measured content (Appendix B). Egyptian wheat prices double between the 150s and the 190s, as the median rises from 7.9 to 17.2 drachmas per artaba ($\Delta \log = +0.78$). Harper (2016) reads the rise as the demographic shock of the Antonine Plague, and Rathbone (1996) as debasement. The decomposition by intrinsic content attributes the great majority of the rise to the coin, and the exact share depends on a detail of the content path. On a tetradrachm series that holds the standard of 169/70 until the end of the reign of Commodus, the share is 74–96% (81% at mid-band fineness), and the intrinsic real price is flat across the plague (from 12.7 to 12.2 g/hl). Such a series, however, omits a documented step. Howgego et al. (2010) report, in their Table 2 and twice in the text, that the 8–10% standard of 169/70 "remained in use until about AD 185/6 when it begins to drop back down to reach the 4–5% standard again by the end of the reign," and that the weight of the coin fell from 12.9 to 12.5 and then to 11.9 g. Four of the eight post-plague observations, at 187/8 and 192, fall inside this decline. Restoring the decline, and using the measured weights, raises the monetary share to 139–187%. The range reflects the pre-167 standard adopted and the interpolation of the decline. A share above 100% means that the intrinsic real price fell across the window, by 37–50%, rather than holding flat. We therefore report the decomposition as a range, at least ∼75% coin and, on the corrected content path, most likely more than 100%.

The correction sharpens the chronology rather than disturbing it. In 164/5 CE the fineness of the tetradrachm begins to crash. The plague is first attested at Seleucia in the winter of 165/6 and reaches Egypt only in 166/7–167/8. The debasement therefore *precedes* the epidemic in Egypt, and it cannot have been a response to fiscal distress caused by the epidemic. The Parthian war of 161–166, whose costs fell before the plague arrived, is the more natural fiscal trigger. In addition, because the corrected content path keeps falling from 185/6 until the end of the reign of Commodus, the debasement extends two decades *beyond* the arrival of the epidemic, as well as beginning before it. The decomposition splits the price rise, not its ultimate cause. The plague is excluded as the trigger and remains a candidate only for the later phase, and the corrected series shows this phase to be longer and deeper than the uncorrected series implies.

A.7 Mediterranean shipwrecks and land use

Real activity Y in Section 7 is the shipwreck census of the Oxford Roman Economy Project (v2.1, March 2017; Strauss et al., 2017). It contains 1,818 wrecks, of which 1,751 carry both an earliest and a latest date. The median wreck is dated to within a century, whereas the bins of Section 7 are 25 years wide, so a wreck cannot be assigned to a year. Instead, each wreck contributes a unit of probability mass spread uniformly across its own date interval, and the annual series is the sum of these masses. Capacity would be a better measure than the number of wrecks, but it is not available. Only 41 of the 1,818 wrecks carry an estimated tonnage, 29 of them in the window, so the proxy for trade counts hulls.

The obvious alternative is land use. HYDE 3.5 (Klein Goldewijk et al., 2017) supplies cropland, grazing, and population for thirty modern territories of the Empire, but it cannot replace the shipwrecks, for two reasons. First, it reports the Roman window in steps of a century, which gives five points against twenty-one bins. Second, its pre-modern population is back-projected from the same demographic literature on which the Egyptian population series draws, so it would not be an independent measurement. Its path is nevertheless worth recording, because it differs from the shipwrecks in a way that bears on the meaning of velocity. With 100 CE set at 100, HYDE cropland, grazing, and population all peak at 200 CE, and they stand at 94, 96, and 96 in 400 and at 67, 78, and 71 in 600. The shipwreck census peaks at 0 CE and stands at 37 and 19 in those years. A proxy for trade and a proxy for population are not the same Y. Over the same span one falls by roughly two-thirds while the other falls by a twentieth, and velocity backed out of the identity inherits whichever proxy is chosen. That is one more reason why Section 7 claims consistency and not an elasticity.

A.8 Gold, military pay, and the grain dole

Three further series serve as dated anchors. The first is the price of gold in denarii, from the Principate (aureus = 25 denarii) through 300–301 CE and the fourth century (Bagnall and Bransbourg, 2019). Gold is the one commodity priced on both sides of the 282–301 gap. The second is military pay. The *stipendium* was 225 denarii under Augustus, 300 from Domitian, ∼500 under Severus, ∼750 under Caracalla, and 1,800 under Maximinus, which gives a dated anchor for nominal wages through the third century. The third is the grain subsidy. We code a dated chronology of the *frumentationes* as an annual regime variable. There was no standing subsidy before 123 BCE, only occasional sales by the aediles. A fixed Gracchan price applied from 123 BCE, and Clodius introduced a *free* dole in 58 BCE. The chronology also codes the Sullan suspension and the caps of Caesar and Augustus. Because the plebs paid a fixed price after 123 BCE, and later a zero price, market grain prices and the cost to the recipient population are decoupled. As a result, the regime enters both as a price control and as a shock to fiscal outlays, and hence as a source of pressure on the money supply. In addition, the 282–301 CE gap in the wheat price series is filled with the papyrus prices from inside the gap reported by Duncan-Jones (1976) and Rathbone (1996), which date the two-step onset (Aurelian in 274/5, Diocletian in 296).

A.9 The leakage rate and the hoarding spike

The silver drain is calibrated to an order of magnitude, as in Section 7, and it varies over time, with a hoarding spike in the third century. Evidence from Indo-Roman coin finds and from the survival of hoards could refine the calibration. How large was the leakage? Pliny puts the eastern drain at 100 million sesterces a year to India, the Seres, and Arabia (*NH* 12.84), of which 50 million went to India alone (*NH* 6.101). Set against scholarly estimates of the imperial silver stock (∼10–20 billion sesterces; Duncan-Jones, 1994), this implies $\delta \approx 0.2$–1.0% a year from the drain alone. The rate rises to ∼0.5–2% once physical wear and permanently buried hoards are added, so that $\delta \approx 0.005$–0.02, the low end of the grid of Section 7. The hoard record shows that the rate was not constant. Of all coin hoarded in 0–400 CE, 57% falls in 225–300, and the annual total peaks in 271 at some 150× the early imperial mean. Counting hoards rather than coins, the figures are 31% and 22×, since the two measures weight a handful of very large deposits differently. This hoarding is the fall in velocity during the crisis. Coin left circulation exactly as debasement flooded the economy with nominal money, which damped inflation during the crisis, while dishoarding afterward helped fuel the runaway inflation of the fourth century. Entering the spike as a structural leakage breaks the long-run consistency, drives the elasticity negative, and makes the trend in velocity explode. That is what hoarding should do if it responded to inflation rather than drove it.

The Indo-Roman coin finds are sometimes read as evidence that the eastern drain stopped altogether during the crisis. The find record as compiled here contains no dated Indian deposit in any decade from the 220s to the 280s. These decades are empty in a corpus biased by survival, however, and they are not measured zeros. The same series carries 97,000 coins in a single decade around 0 CE. The drain is therefore unmeasured during the crisis rather than shown to have stopped, and the finds cannot bound how much of the leakage of the third century was internal.

A.10 Machine extraction of the papyrus record

The price and wage observations drawn from the Duke Databank (DDbDP, 2026) are extracted by a large language model (Claude Sonnet 5, Anthropic), not by hand, under a fixed set of instructions and a fixed record format. The model receives the EpiDoc TEI transcription of each document and must classify every figure it reports as a unit price, a total for a stated quantity, a bare quantity, or a bare sum. Only the first two are admitted. Text supplied by editors is marked in the input, so a numeral restored by a modern editor can be separated from one read on the papyrus. Dates are never taken from the model. They come from the HGV metadata attached to the document identifier. Documents that the model fails to read are counted rather than discarded. The reading is validated document by document against Harper (2016). Of the papyri he cites that the corpus can reach, a wheat price is recovered from 75%, and 12 of these 15 agree with his published figure to within 10%. Where a document yields several prices, we compare the median, which is chosen without reference to Harper. Picking instead the candidate nearest to his published figure would give 13 of 15. That rule uses the answer to choose the guess, so we report the blind number. The comparison is made document by document rather than between medians, since a reading can reproduce a plausible central tendency from the wrong documents. Finally, of 120 labor accounts that state their own arithmetic total, 110 reproduce that total from the rate and the count of worker-days that the model extracted independently of it.

A.11 A second reading by an open-weight model

To measure how much the extraction depends on the model that made it, every candidate document was read a second time by an open-weight model (Qwen3.6-27B at 8-bit precision). The instructions and record format were the same, with one addition, a table of the numeral conventions of the editions. Without this table, the open model read the capital-letter thousands of the Greek numerals as units (Γφ, 3,500, as 35), and it agreed with the primary reading on the money figure in only 55% of a reference set of 40 documents, against 69% with the table. The two readings were aligned record by record on the 12,328 documents that both returned. In all, 21,989 record pairs align, which is 68.6% of the 32,033 records of the primary reading. On aligned pairs the readings agree on the money figure in 69.6% of cases, on the unit priced in 75.5%, on the transaction type in 71.3%, and on the commodity in 66.0%. For the three panel quantities of this appendix (wheat per artaba, rent per aroura, and wage per day), the two readings fall within 10% of each other in two-thirds of the documents that carry them.

The readings differ on 8,005 documents. Of these, 200 were drawn at random and adjudicated by a third reading of the text (Claude Sonnet 5), which returns a verdict on each disputed record, 861 verdicts in all. The primary reading is right in 67.7% of the verdicts, the open-weight reading in 14.2%, both are defensible in 7.1%, and neither is right in 11.0%. Where the two readings give different money figures, the primary reading is right in 78% of cases and the open-weight reading in 4%. Where the open-weight model reports a money record that the primary reading lacks, the adjudicator finds a real record in 33% of cases and none in 49%. Where the primary reading reports a money record that the open-weight model lacks, the record is real in 88% of cases. The "neither" verdicts mostly involve denominators that the text never states, such as a flat rent for a half-share read as a rate per aroura.

A.12 The panels on agreed records

A more demanding use of the second reading is to rebuild the panels on the records on which the two readings agree. A record is agreed if it is aligned and has the same money figure within 1%, the same unit priced, and the same transaction type. This gives 8,898 records, or 27% of the primary reading, and the inclusion rules of each panel are left unchanged. The rent medians do not move. They are 6.0 artabas per aroura in AD 1–165 on 36 of 48 observations and 5.5 over 1–296 on 60 of 77. Nor does the median price of wheat in the first and second centuries move, at 8.0 drachmas per artaba on 16 of 34 quotations. The third-century windows keep one agreed quotation each and cannot be checked in this way. The one series that moves is the day wage. With one median per document, the rate falls from 1.71 to 1.21 drachmas on the 20 of 27 documents that keep an agreed record, back inside the published band of 0.7–1.3. It moves because the money figure is the field that the readings most often dispute in wage accounts. The reported 1.71 is thus the conservative figure, and the wage side of the real wage is unchanged at 4.85 liters when the wheat denominator is held at the main panel.

A.13 Coin-quality clauses

The clause shares reported in Section 5 come from the same open-weight model, asked a separate question about every document in the corpus that names a sum of money (12,482 documents). The question is whether the sum is qualified as imperial coin, as struck or approved coin, as an accounting standard, as old or new coin, or by a weight standard. The shares are counts of the answers by half-century among documents that state a money sum, and they are not adjudicated. A stratified subsample of 3,215 of the documents gives the three third-century shares within a point of the full corpus (4.6, 9.6, and 12.9%). The count of refusals comes from a separate reading of every letter, petition, and official text (13,545 documents). A statement is counted only where the text itself says that a coin was refused, void, or not current.

A.14 An exact-dated papyrus price panel

Sharpening the stock margin requires a price series at annual rather than generational frequency, and the Duke Databank of Documentary Papyri (DDbDP) is the corpus that can supply one. The 12,370 documents extracted from it yield 51 wheat quotations that survive every inclusion rule below. Of these, 42 are dated to the day and 35 rest on assayed rather than inferred coin fineness. Building the panel involves three difficult problems.

The first problem is dating. Documents are keyed to the Heidelberger Gesamtverzeichnis (HGV), whose date record carries either a single date, for a text dated to a day or a year, or a pair of not-before and not-after dates for a range. Of 65,972 indexed HGV documents, 15,619 are dated to the day and only 2,172 to a bare year. The remaining two-thirds are ranges, many of them spanning a century. Two traps in the encoding matter. First, a range-dated text collapsed to its midpoint is indistinguishable in a panel from a text dated to an afternoon, unless the span travels with it. The span in years is therefore carried on every observation, and nothing is imputed. Second, a Trismegistos number can cover many HGV documents with different dates. For example, TM 11921 is an archive of 17 texts spanning 297–301 CE. The unit of record must therefore be the HGV document, not the Trismegistos number.

The second problem is telling a price from a quantity. A papyrus recording "2,427 artabas of wheat" is an account total, not a price, while one recording "τιμῆς πυροῦ ἀρταβ- κ ... δραχμ- τιε" (315 drachmas for 20 artabas) is a price. Any extraction that does not force this distinction at the point of capture will produce a series dominated by quantities, since quantities are far more common in the corpus than quoted rates. We therefore require every extracted figure to be classified as a unit price, a total for a stated quantity, a bare quantity, or a bare sum, and we admit only the first two. Conversion to silver uses the measured content of the Alexandrian tetradrachm for the year in question, with the assay band propagated. It does not use the single fixed parity that, as noted above, would make the silver series a rescaling of the nominal one.

The third problem is that a rate need not have money on top. A record with one money field and one quantity field per figure can represent "wheat at eight drachmas the artaba." It cannot represent the standard Egyptian lease, ἐκφορίου … ἑκάστης ἀρούρης ἀνὰ πυροῦ ἀρτάβην μίαν (one artaba of wheat for each aroura), in which both sides are quantities and one of them would have to be discarded. Such a record would also split a lease paying part in grain and part in cash into two unlinked rows, each reporting a fraction of the obligation. Each record therefore carries a *list* of payment components, each independently in money or in kind, over a single non-monetary denominator, so that a rate can be represented whichever side carries the coin. The candidate pool of 12,489 documents is drawn by folding diacritics and by admitting the vocabulary of rents and wages alongside that of money. Requiring a money word before examining anything else would exclude leases paid wholly in wheat, and a literal match on accented Greek stems would miss the commonest inflected forms of ἀρτάβη, the commonest measure word in the corpus.

The median of 8.0 drachmas per artaba in the first and second centuries, against the 8–12 of Duncan-Jones, rests on 34 quotations. The median of 14.0 in the early third century, against the 12–20 of Rathbone, rests on 2, and the median of 16.0 in the mid-third century rests on 5. The nominal median lies inside the published band in every window we can check, but only the first window is a test. The other two are consistency checks on a handful of documents. The document-level validation against Harper (2016) recovers 12 of 15 of his cited papyri to within 10% on a blind match. Nineteen of the 51 quotations are not market sales. They are 11 tax assessments and 8 valuations. A tax or a valuation assessed *at* a stated price is evidence on prices, so we admit them, but a reader may reasonably want the 32 market sales alone. In addition, the late window is concentrated. One papyrus (HGV 21076, AD 294) supplies eight of the sixteen records in 250–296. A total of 51 observations spread over three centuries, a median of one or two per decade, will not identify a relation between money and prices that 21 generational bins could not identify, so we report no elasticity from the panel. The silver conversion, now correct for the period, still rests on the measured content of the tetradrachm, which is assayed only up to the Antonines. The panel establishes that the extraction is sound and that the series can be built. This is a necessary condition for the work at annual frequency that the conclusion calls for, not a substitute for it.

A.15 Land rents and wages

The same extraction yields 77 observations of arable land rent, in artabas of wheat per aroura per year, and 295 wage observations, both for 1–296 CE. Neither series is available from the price panel, and both require the record of payment components, which represents a rate whichever side carries the coin. The rent series is reported in the unit that the papyrological literature uses. Cash rents sit beside it rather than in it, since converting a money rent at a wheat price would import the price series that the rents are meant to be independent of. Share rents, vineyards and orchards, public and temple land, and figures covering a whole lease term rather than a year are excluded on recorded fields, not on keywords. Rents stated inclusive of the landlord's seed are adjusted and flagged. The level agrees with the published record. The median is 6.0 artabas per aroura in AD 1–165, against Harper's trimmed mean of 5.8, and 5.5 across the whole window, against a range of 3–10 in the literature.

The pooled rent series falls by 51% after the Antonines. This fall reflects composition, not economics, and we report no post-Antonine decline from the rent panel. Rents differ across nomes by a factor of four. In AD 1–165 they are 11.0 artabas per aroura at Tebtynis, 7.0 in the Hermopolite, 5.6 at Oxyrhynchos, and 2.6 at Bakchias. The nomes represented in the sample change across exactly the window in question, and the high-rent nomes drop out of the later evidence. Holding the nome constant, Oxyrhynchos, the only nome observed on both sides of the window with more than a handful of documents, falls by 9% between AD 1–165 and 166–249 (from 5.59 to 5.11 artabas per aroura, $n = 12$ and $n = 8$), and by 27% if the late window is extended to 166–296 (from 5.59 to 4.10, $n = 12$ on each side). Both the rent and the wage panels admit HGV documents whose date is a range of up to 25 years, with the span carried on every observation. Restricting the sample to exactly dated texts leaves 49 rent observations and does not change the medians materially.

A.16 A real-side control without silver conversion

The rent series bears on the weakest link in the decomposition of Section 4. That decomposition compares the nominal price of grain with its price in silver, and the silver conversion rests on the measured content of the tetradrachm, which is assayed only up to the Antonines. A rent of N artabas of wheat per aroura, in contrast, is denominated in *grain*. It is invariant to debasement by construction, since no coin, assay, fineness band, or conversion enters it. If the great inflation was monetary, this series should be roughly flat across the window in which nominal prices explode.

Holding the nome constant, Oxyrhynchite rent runs at 5.59 artabas per aroura in AD 1–165 ($n = 12$) and at 3.73 in 250–296 ($n = 4$), a fall of a third. Over the same span the nominal price of wheat rises from 8.0 drachmas the artaba in 1–165 ($n = 32$) to 220 in 274–296 ($n = 10$), a factor of 27. A real magnitude that falls by a third, against a nominal price that rises by a factor of 27, is the monetary reading of the inflation. This comparison involves no silver conversion, no assay, and no fineness band.

Four limits qualify the comparison. First, we report no real wage for 250–296, because the wheat prices in that window will not support a ratio. The window is admitted here because a median price *level* is not a ratio denominator matched observation by observation, and the precision comes from the rent side, not from the price side. A reader who rejects this distinction can read the contrast as directional only, which is all we claim from it. Second, the late Oxyrhynchite window rests on four observations. Third, no wheat price in the panel is both Oxyrhynchite and post-Aurelian, so in the late window only the rent figure is measured within one nome, and the late nominal median is itself concentrated, with HGV 21076 (AD 294) supplying eight of the sixteen records in 250–296. Fourth, the comparison draws on the same rent evidence as the previous subsection, so it corroborates rather than tests independently. What it establishes is narrow. The direction and the rough magnitude of the decomposition survive when the assay of the tetradrachm is removed from the argument altogether.

A.17 Real wages

The full basket of the Edict extends the check of Section 4 from grain to labor. A wage of 25 denarii per day plus maintenance (∼11 denarii; Allen, 2009), against the Edict's maximum price of wheat, implies a daily real wage of 3.2–4.7 liters of wheat for an unskilled farm laborer. The figure is below Harper's pre-plague (45–163 CE) Egyptian norm of 4.9–5.8 liters per day. In real terms, goods prices *and* wages passed through the great inflation roughly intact.

Wages in drachmas per day divided by wheat prices in drachmas per artaba give a wage in grain. For 45–163 CE the papyrus real wage is 3.88 liters per day across 93 records, when each day-rate wage is matched to the median wheat price within fifteen years. It is 4.85 when we take one median per document, the rule that we also apply to rents, and this is the figure we report. It lies just below Harper's 4.9–5.8 and above the range of the Edict. The figure is stable, and it moves by 7% across matching windows from ten to forty years.

The denominator of a real wage is the price panel, which contains three wheat quotations in 166–249 CE. The implied post-Antonine real wage ranges from 3.6 to 8.3 liters per day, depending on the matching window, a factor of 2.3 that is larger than any effect one would wish to detect. For this reason we report no real wage for the two later periods. A rise in the real wage in these periods would be the expected Malthusian signature of the Antonine plague, but here it would come from the choice of window, not from the data.

The 33 records that explicitly carry a maintenance clause have a median of 0.40 drachmas per day, against 1.21 for the rest. Cash paid alongside board, lodging, or clothing is not a wage but a supplement to a compensation package whose larger part is unobserved, and pooling the two groups produces a median that describes neither. Recording the clause allows us to exclude these records. A text that explicitly denies maintenance is recorded as such, and exactly one record in 295 does so. The remaining 261 are *silent*, and silence is not evidence that no maintenance was paid. The exclusion removes attested cases rather than identifying pure cash wages, and the residual is a mixture whose composition we do not observe.

Unskilled day rates run at 1.21 drachmas per day across 118 records, inside the band of 0.7–1.3 reported by Drexhage (1991) and Scheidel (2010). However, these records come from only 27 documents, one of which supplies 23 of them, and the effective number of independent documents is about twelve. Taking one median per document gives 1.71 drachmas per day, which is outside the published band. We apply the same rule to rents, because records from one archive are not independent draws. We therefore report the aggregated 1.71. Applying the rule to rents and not to wages would have kept the wage series inside the band through an inconsistency. The gap between 1.21 and 1.71 is the cost of applying the rule to both. Of the 118 observations behind the median day wage, 95 are denominated in obols, so the conversion of obols to drachmas matters more here than anywhere else in the paper. We fix it from the papyri, not by convention. Roman Egypt reckoned the drachma at seven obols, and a single account confirms this twice. P.Oxy. (HGV 11596) totals ἐργάται ϙη ἐξ ὀβολῶν ιβ δραχμαὶ ρξη, that is, 98 worker-days at 12 obols entered as 168 drachmas, and $98 \times 12 = 1{,}176 = 168 \times 7$. The same account enters 31 worker-days at 12 obols as 53 drachmas 1 obol, and $31 \times 12 = 372 = 53 \times 7 + 1$. The second entry is decisive, because a remainder of exactly one obol cannot balance on a drachma of six obols. With six obols to the drachma, the median day wage would be 1.33 rather than 1.21 drachmas, outside the published band.

B The money stock, the coin counts, and the fiscal origins of debasement

B.1 The coin stock and bullion

Write the bullion stock as a perpetual inventory of new silver, $M_t^{\text{Ag}} = \sum_s (1-\delta)^{t-s} b_s$, and the nominal coin stock as the same flows minted at the silver content of the coinage c_s^m, $M_t^{\text{nom}} = \sum_s (1-\delta)^{t-s} b_s / c_s^m$. Debasement makes the coin stock grow faster than bullion. The wedge $M^{\text{nom}}/M^{\text{Ag}}$ rises about fourfold to its peak in the 270s, as the denarius falls from 3.8 to $\sim$0.1 g of fine silver, and this is nominal money created without new silver. Through the first three centuries the nominal price of Egyptian wheat rises with this nominal coin stock, which is an inflation driven by physical debasement.

Two properties of this construction should be read together with Section 6. First, it is correct by vintage in principle, since each cohort of bullion is minted at the standard of its own year. However, it assumes that all new silver was struck at the current standard and that no old coin was melted and re-struck, and the second assumption is hard to sustain. Newly struck coin makes up 43 to 51% of the hoarded stock in each decade from the 240s to the 270s, at a time when lead-flux bullion was falling by some 70%. New metal alone cannot supply this much coin, so old coin was being melted and re-struck. Second, and more simply, the coin stock is a construction, whereas the divergence within hoards reported below is a measurement.

The fourth century differs in kind, and the two reforms mark the divide. Diocletian restored the fineness of the coinage (the argenteus of 294 was again $\sim$95% silver), yet the price of wheat in the denarius communis ran to millions per artaba. The money of account inflated while the metal improved, so the metal-anchored coin stock no longer tracks prices. The great inflation of the fourth century was therefore not a commodity debasement but a fiat inflation of the unit of account, whose value rested on tariffing and confidence, as in the fiduciary regime of Section 5. The construction is illustrative. The content c_s is that of the imperial denarius while the prices are Egyptian, and the co-movement before 296 is partly a common trend. A structural estimate of the coin stock requires data on mint output.

B.2 Die studies and the level of the money supply

The proxy based on flows over fineness cannot fix the absolute scale of the money stock. This scale comes from mint die studies. Coin production is the number of obverse dies behind a coinage, estimated from the dies recorded in large samples, times an output per die that is heavily debated but of the order of tens of thousands of coins (de Callataÿ, 2011). The Roman Republican Die Project (over 300,000 die-identified coins behind Crawford, 1974) supplies die counts for the Republican denarius, and the Reka Devnia hoard is the standard proxy for *relative* imperial production by reign (Spoerri Butcher, 2019). Of the 101,096 coins that the deposit records in the CHRE data, 81,044 denarii are attributed to an issuing reign. The remaining 20,052 are unattributed or illegible, carry no mint date, and cannot enter a production series (Spoerri Butcher, 2019). Assembled into an estimate of the money supply, this evidence puts the imperial silver stock above HS 20 billion around 160 CE (Duncan-Jones, 1994), on the order of 16,000 metric tons of fine silver. The figure is consistent with the 8,000–16,000 metric tons implied independently by our calibration of the eastern drain (Section 7). The die evidence thus anchors the *level* of the money stock that the construction of the coin stock scales.

The die-analysis records of the Roman Republican Die Project show how far a continuous series has come. Of the 309 Crawford types in the project's phase-one window (RRC 336–392, 92–75 BCE), 30% have been transcribed so far, giving 2,371 recorded obverse dies. Interpolating across the untranscribed gaps and assuming a standard output of 30,000 coins per obverse die places the peak of physical production during the Sullan civil wars (82 BCE) at roughly 18.6 million denarii. We present this as a demonstration that such a series can be built, not as a series. With only 30% of types transcribed the gaps are wide, and the linear interpolation across them is crude enough to return a negative die count in one of the eighteen years (76 BCE). A continuous production series requires further transcription, not further interpolation.

B.3 Coin counts and the divergence

The imperial half of the production record is already public. Coin Hoards of the Roman Empire publishes the coin-level composition of Reka Devnia (hoard 3406) as 3,353 coin-type rows whose quantities sum to the full hoard of 101,096 coins. Aggregated to production by mint date, the raw counts appear to deliver the test of coin against bullion directly. Coin abundance at Reka Devnia correlates -0.68 with lead-flux bullion and -0.62 with coin fineness, and abundance rises about 51% into the Severan period as bullion falls 60%. The correction for wastage overturns this result, however. A hoard closing around 251 is a snapshot of *circulation*, and it over-represents recent issues. Deflating the count of each reign by a plausible annual rate of loss (the standard correction of Duncan-Jones, 1994, $\lambda = 1$–3%) inflates the early empire. The corrected production series then *declines* into the Severan period (-27% to -86%) and becomes *positively* correlated with bullion and fineness.

A far broader series shows the same raw pattern and the same confound. In the linked-data corpus of Nomisma.org (OCRE, 2026), with some 211,000 silver specimens, the count of silver coins roughly *doubles* through the crisis, as the flood of antoniniani peaks near 17,000 specimens in 250–274 CE, while lead-flux bullion falls by ∼70%. However, this corpus is weighted by survival and by modern collecting in the same direction as the raw hoard counts. Both datasets of coin counts are therefore *consistent with* the accounting divergence $M^{\text{nom}} = M^{\text{Ag}}/c$, but they do not establish it independently. The divergence rests on the fineness record, not on coin counts.

B.4 Gold as a second numeraire

Gold provides a second numeraire that crosses the 282–301 gap. The aureus stood at 25 denarii (∼1,125 denarii per pound of gold) throughout the Principate. By an official transaction of 300 CE a pound of gold cost 60,000 denarii, and the Edict caps it at 72,000 in 301. The rise is ∼64-fold and concentrated in the late third century, and the price continued to ∼39 million by 348 and to $\sim 5\times 10^8$ by the 380s (Bagnall and Bransbourg, 2019). The path of gold and the path of wheat tell the same story in two numeraires, that of a single monetary explosion with commodity relative prices comparatively stable through it (Figure 8).

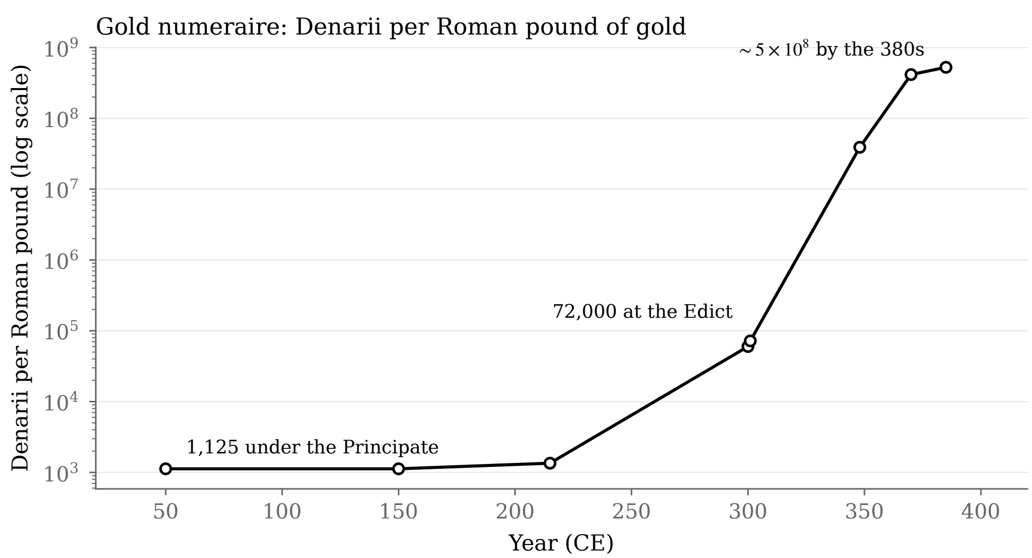


Figure 8: **Gold as a second numeraire.** The price of gold in denarii per Roman pound (log scale) rises from 1,125 under the Principate to $\sim 5\times 10^8$ by the 380s. It is the same monetary explosion that the price of wheat shows, measured in an independent numeraire priced on both sides of the 282–301 gap. Source: Bagnall & Bransbourg (2019).

The two metals also drifted apart at the mint. The classical Roman ratio of the values of gold and silver was near 12 : 1. Because the aureus stayed tariffed in an ever more debased denarius, however, its *implied* ratio fell from 11.7 : 1 around 50 CE to 6.4 : 1 by 215, so gold was progressively undervalued in silver at the mint. Gold undervalued at the mint is the textbook setting for Gresham's law. Undervalued good money is hoarded and exported, and it is driven from circulation by the debased silver. This anticipates the flight to a gold standard in the fourth century, which the solidus completes. Coin counts cannot quantify that flight, because the biases of survival and collection documented above make ratios of abundance unreliable. The divergence in the value ratio rests on the tariff and the fineness record alone. The same reading does *not* carry over to the silver coinage, where the hoard record can test it directly and rejects it (Section 6). The difference has an economic reason. Gresham's law requires the mispricing to be visible to the person making the payment. Between gold and silver the mispricing was posted, in a tariff attached to a coin that nobody could confuse with another. Between a Severan denarius and a Caracallan one, in contrast, it was hidden inside metal that the mint had deliberately blanched to look identical (Appendix A).

B.5 The divergence measured inside the hoards

The coin-level layer of Section 6 does establish the divergence, and it escapes the objection from wastage for a structural reason. Count the coins in a hoard and, separately, sum the fine silver in the same hoard by valuing every coin group at its own striking vintage. Both quantities come from the same deposits and enter the same regression, with 25-year deposition bins, fixed effects for one-degree cells, and standard errors clustered by cell. They are therefore deflated by the same biases of survival and collection, and these biases cancel in the contrast. Of the 7,688 headline hoards, 1,194 carry coin-level rows and 873 carry observed fine silver, with 246,861 of 422,769 silver coins valued. Against a base of 125–149 CE, third-century deposits hold +0.762 log points more coins ($t = 4.21$), while their observed fine silver is flat, so the divergence is +0.956 log points (se 0.152, $t = 6.30$; wild cluster bootstrap $p = 0.000$). It survives every reweighting reported below (Figure 9).

The informative part of this result is its magnitude, not its sign. The sign is guaranteed, since debasement is a known fact, but +0.956 log points is a factor of 2.6 in fine silver per coin, and the interesting question is what it should be compared with. The natural benchmark is the same regression run on the measure that the literature has used, in which the coins of each hoard are valued at the standard in force when the hoard was buried, as though none of them had been struck earlier. This gives a divergence of +1.336 log points (se 0.068), a factor of 3.8. The whole difference between 3.8 and 2.6 is coin that was older than its burial date. This is the 36-year lag of Section 6, recovered from an entirely separate design. The two regressions share hoards, cells, and estimator, and they differ in one convention of measurement only.

The design has two limits. First, coverage is only 11.4%. Second, coverage is non-random in exactly the wrong places (4.7% in the 275 bin and 8.4% in the 250 bin, against 16.6–20.1% in the bins covering the second century). As a result, the level of the third-century fall in silver per hoard is not identified. It is −0.19 (se 0.21) with weights from the bins of the covered subset, −0.29 under inverse-probability weighting, −0.44 under population bin weights, and −0.55 when the wedge is extrapolated to the full panel. What is robust across all of these weightings is the *measurement effect*. Moving from a deflator based on the year of deposition to observed vintages shifts the third-century estimate by +0.380 (se 0.128, $t = 2.97$ on a paired test). On identical hoards this shift decomposes into +0.323 ($t = 5.95$) from the mix of vintages and +0.057 ($t = 0.49$) from the composition of the metal. The shift is a clean vintage effect.

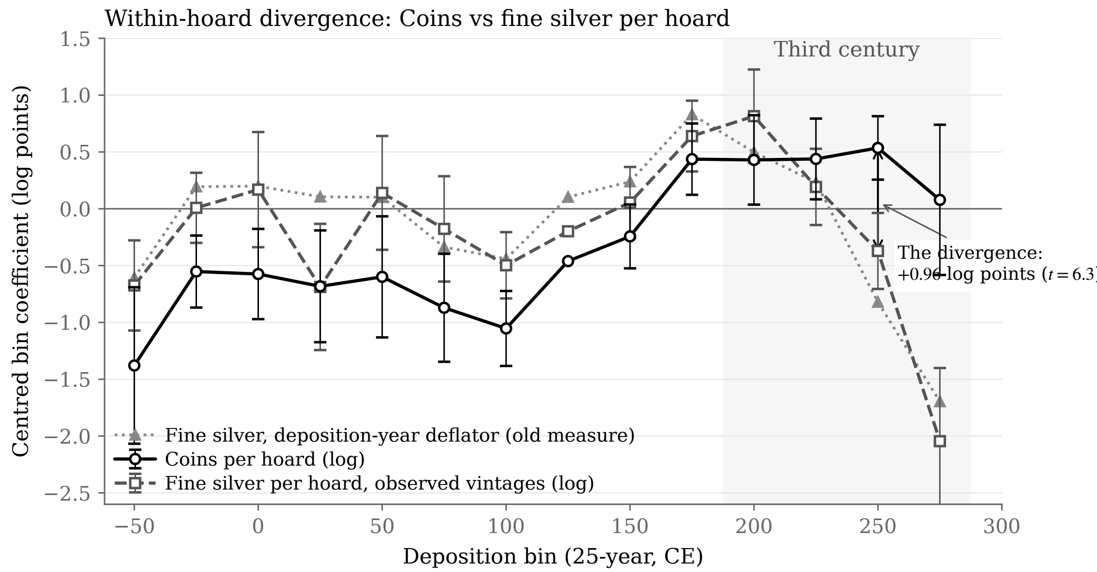


Figure 9: **Coins and silver in the same hoards.** Centered bin coefficients from a common design (25-year deposition bins, one-degree cell fixed effects, cell-clustered standard errors) on the 698 hoards for which every outcome is defined. Coins per hoard rise into the third century, while the fine silver those coins contain, valued at each coin's own striking vintage, does not. The gap is the accounting divergence $M^{\text{nom}} = M^{\text{Ag}}/c$, measured rather than constructed. The conventional deflator based on the year of deposition (triangles) assumes that every coin in a hoard was struck at the standard of the year in which the hoard was buried, and this assumption alone produces most of the apparent collapse of silver in the third century.

B.6 A test of fineness-selective survival

The natural test of fineness-selective withdrawal compares the decay of the pre-194 vintages with a counterfactual in which new coin merely dilutes old coin. The test appears to find selection, with an excess decay of 1.31 percentage points a year over 195–235 and 2.00 over 215–255, but it identifies nothing. Consider a simulated null with no preference for fineness, a single hazard applied to every vintage, and the hoard sizes and deposition dates of the corpus itself. Under this null the same estimator returns +0.7 to +2.0 points a year, which equals or exceeds every one of the estimates except that for the Caracallan window. An accumulated hoard is a lagged average of past circulation rather than a snapshot of it, and the recursion in the benchmark converts that lag into apparent excess decay. Across defensible ways of measuring the inflow of new coin, the statistic spans −2.9 to +3.9 points a year, so even its sign is a matter of convention. Tests of vintage survival based on hoards cannot identify fineness-selective withdrawal without a correction for the way each hoard accumulates, and nobody has yet built such a correction. The within-hoard cohort test of Section 6 needs no benchmark, and it is the test on which the paper rests.

B.7 Mines and mints in detail

Section 2 assumed that the two margins of the money supply move for unrelated reasons, bullion with mining geology and fineness with fiscal pressure. The circulating series allows us to check this assumption directly against lead flux at annual frequency ($N = 330$, 30 BCE–310 CE). In levels the correlation is +0.43. Both series fall across the third century, and no test can separate a common collapse into its causes. In first differences the correlation is +0.01, and its absolute value does not exceed 0.03 at any lead or lag from −20 to +20 years. The mint standard itself behaves in the same way against lead, with +0.44 in levels, −0.02 differenced, and at most 0.11 at any lead or lag, so the result is a property of the monetary series and not of the smoothing involved in $\bar{c}_t$. The test has limits in both directions. Where the assumption can be tested, it holds cleanly. Where the two series do move together, in the third-century decline that matters most, the test has no power, because $\bar{c}_t$ is a slowly moving stock and a ten-year window contains little independent variation.

B.8 The reform of 274/5 in detail

It is tempting to read Aurelian's reform in the same way as the test at 296, since 11.5–15.1× predicted against 12.5× observed looks like a success. We do not accept this reading. The event confounds the release of the premium, which the model captures, with a change in content and a re-tariffing, which it does not capture. First, the direction of the change in content is not the one that the c_t in the table assumes. Christiansen (2004), the source of the table, has Aurelian replacing the old tetradrachms with a more heavily debased standard, but the metal analyses point the other way. In a survey of these analyses, Haklai-Rotenberg (2011) records "the slight increase in their silver content in 274" for Roman-Egyptian tetradrachms of 260–280, citing Cope et al. (1997). Cope et al. measure the metal directly, by wet chemistry, and on a question of metal content this outranks reasoning from the composition of hoards. Second, if the content rose, the content term works against the release of the premium, and the prediction falls to 4.2–6.6×, which under-predicts the observed 12.5× by a factor of two to three. Third, and most importantly, Aurelian's reform re-tariffed the unit of account. The XXI/KA mark raised the face value of the coin by more than it raised its metal, so part of the price response is a fiat devaluation that the premium accounting does not model at all. We therefore rely on the 296 test alone. The direction of the disagreement is also informative in its own right. The price explosion of the 270s was larger than a pure collapse of confidence can deliver, which is evidence that the re-tariffing raised prices on top of that collapse.

B.9 Harper's silver conversion

Harper's silver conversion in the first three centuries is nearly constant, at ∼0.97 g of silver per drachma, the parity of the imperial denarius. It does not track the tetradrachm's own debasement, so his silver series for the first three centuries is the nominal series rescaled, and it cannot separate real from monetary shocks. An apparent decomposition within this window based on his series is an artifact. Its "monetary share" falls from 58% to 4% once two observations at breaks in the coinage are removed. A naive regression of prices on fineness is also dominated by the single observation at the onset in 282 CE, which accounts for half of its slope. The repair uses the tetradrachm's own content, as in Section 4.

B.10 Political instability and debasement

The political record allows us to check whether the debasement was driven by fiscal pressure rather than by harvests. We build an annual index of imperial political instability from the chronology of reigns, counting accessions, violent depositions, and simultaneous usurpers. The index runs at 2.6 events per year during the third-century crisis of the soldier emperors, against 0.2 under the Principate. It is negatively related to coin fineness (detrended $t = -2.9$; correlation -0.48), which is consistent with accession donatives and war finance passing into the coinage. The co-movement belongs to the crisis era and is not a first stage at annual frequency, since the year-on-year change in fineness is unrelated to instability. In addition, instability is not a valid instrument for the monetary channel on its own, because civil war moves grain prices directly, through requisitions and disrupted transport, and not only through the coinage. We read this result as narrative evidence that the great debasement was fiscal and political in origin and exogenous to Egyptian harvests, not as a two-stage estimate.

B.11 The monetization test in full

Section 4 reports that the money share of payment legs falls from 93.2% in 150–199 (2,383 legs from 654 documents) to 82.2% in 200–299 (2,123 legs from 674 documents), with $z = 11.3$, on 15,053 legs drawn from 12,370 dated documents. A well-chosen baseline can manufacture a result of this shape, so we check the fall against every available pre-200 baseline (z between 7.0 and 11.3), against a cut of 150–234 versus 235–284 ($z = 8.6$), and against collapsing the data to one observation per document ($z = 4.2$–4.8). The only specification that reverses the sign assigns the peak of 165–199 itself to the "post" window, which does not test the hypothesis. The attempted corroboration on 1,480 dated agricultural leases, in which rent is fixed either in drachmas or in wheat, fails on its own terms. In the leases the third-century fall appears only with the local peak of 150–199 as the baseline, reverses against broader pre-windows, and is insignificant when the data are collapsed by document. This is why the text makes the negative claim and no more.

B.12 The food-supply channel

Real supply shocks should move grain prices independently of the monetary channel. The Nile provides the natural test, and the agricultural calendar fixes the alignment. The flood crests in late summer, sowing follows, and the grain is harvested the next spring and sold over the following year, so a flood prices the harvest of the *following* year. Matched in this way, a stronger flood goes with a *cheaper* real (silver) price of Egyptian wheat (correlation -0.28, $n = 18$ overlapping quotations), and the independent graded flood scale of Huebner (2025) agrees (-0.24, $n = 15$). The effect decays at a lag of two years, as a harvest shock should. The two reconstructions of the flood agree on the sign, but not on the magnitude. Three further limits apply. The low-flood flag is set in only three of the price-years, too few to estimate the effect of a drought separately. The index for the Roman window is a documentary code with three values, not a measured crest height. The overlap with dated prices is also thin. Famines, geolocated to the regions that supplied grain, act on prices rather than on mining output, since the same famines leave the lead series unmoved, and this is what a shock to the food supply should do. The Egyptian population series ($n = 28$, 14–575 CE) records the three pandemics as discrete drops (Antonine -15%, Cyprianic -20%, Justinianic -51%), and it supplies Y_t for the quantity identity.

Declarations

Competing interests. The authors declare no competing financial or personal interests.

Funding. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability. The derived datasets and the code that reproduce every figure, table, and number in this paper will be deposited in a public repository with a DOI on acceptance, and are available from the authors on request before then. The Coin Hoards of the Roman Empire data are used under CC BY-NC-SA 4.0 and the files derived from them inherit that license; the Duke Databank of Documentary Papyri corpus, the Heidelberger Gesamtverzeichnis metadata, and Harper's DARMC geodatabase are public.

References

Allen, R.C. (2009). How prosperous were the Romans? Evidence from Diocletian's Price Edict (AD 301). In A. Bowman and A. Wilson (eds.), *Quantifying the Roman Economy: Methods and Problems*, Oxford University Press, 327–345.

Alonso Ortiz, J. and Da-Rocha, J.-M. (2026). Compound Stress and the Roman Pandemic Depression. Working paper.

Bagnall, R.S. and Bransbourg, G. (2019). The Constantian monetary revolution. *ISAW Papers* 14.

Bowman, A. and Wilson, A. (eds.) (2009). *Quantifying the Roman Economy: Methods and Problems*. Oxford University Press.

Butcher, K. and Ponting, M. (2015). *The Metallurgy of Roman Silver Coinage: From the Reform of Nero to the Reform of Trajan*. Cambridge University Press.

Cagan, P. (1956). The monetary dynamics of hyperinflation. In M. Friedman (ed.), *Studies in the Quantity Theory of Money*, University of Chicago Press, 25–117.

Coin Hoards of the Roman Empire (2026). *Coin Hoards of the Roman Empire dataset*. Ashmolean Museum, University of Oxford, `https://chre.ashmus.ox.ac.uk` [accessed August 12, 2026].

Christiansen, E. (2004). *Coinage in Roman Egypt: The Hoard Evidence*. Aarhus University Press.

Coin Hoards of the Roman Republic Online (2026). Ed. K. Lockyear. American Numismatic Society, `https://numismatics.org/chrr/` [accessed July 2026].

Cope, L.H., King, C.E., Northover, J.P. and Clay, T. (1997). *Metal Analyses of Roman Coins Minted under the Empire*. British Museum Occasional Paper 120. London: British Museum.

Crawford, M.H. (1974). *Roman Republican Coinage*. Cambridge University Press. (Die counts: Roman Republican Die Project, American Numismatic Society.)

Digital Atlas of Roman and Medieval Civilizations (2016). *Prices, Wages, and Rents for Roman Egypt, AD 1–700* (K. Harper). DARMC Scholarly Data Series, Harvard University.

Duke Databank of Documentary Papyri and Heidelberger Gesamtverzeichnis der griechischen Papyrusurkunden Ägyptens (2026). EpiDoc TEI corpus, `https://papyri.info` (`github.com/papyri/idp.data`) [accessed July 2026].

de Callataÿ, F. (2011). Quantifying monetary production in Greco-Roman times: a general frame. In F. de Callataÿ (ed.), *Quantifying Monetary Supplies in Greco-Roman Times*, Bari, 7–29.

Drexhage, H.-J. (1991). *Preise, Mieten/Pachten, Kosten und Löhne im römischen Ägypten bis zum Regierungsantritt Diokletians*. St. Katharinen: Scripta Mercaturae.

Duncan-Jones, R.P. (1976). The price of wheat in Roman Egypt under the Principate. *Chiron* 6, 241–262.

Duncan-Jones, R. (1994). *Money and Government in the Roman Empire*. Cambridge University Press.

Eichler, A., Gramlich, G., Kellerhals, T., Tobler, L. and Schwikowski, M. (2015). Pb pollution from leaded gasoline in South America in the context of a 2000-year metallurgical history. *Science Advances* 1(2), e1400196.

Fisher, I. (1911). *The Purchasing Power of Money*. Macmillan.

Friedman, M. and Schwartz, A.J. (1963). *A Monetary History of the United States, 1867–1960*. Princeton University Press.

Haklai-Rotenberg, M. (2011). Aurelian's monetary reform: between debasement and public trust. *Chiron* 41, 1–39.

Harl, K.W. (1996). *Coinage in the Roman Economy, 300 B.C. to A.D. 700*. Johns Hopkins University Press.

Harper, K. (2016). People, plagues, and prices in the Roman world: the evidence from Egypt. *Journal of Economic History* 76(3), 803–839.

Harper, K. (2017). *The Fate of Rome: Climate, Disease, and the End of an Empire*. Princeton University Press.

Howgego, C., Butcher, K., Ponting, M. and Heuchert, V. (2010). Coinage and the Roman economy in the Antonine period: the view from Egypt. *Oxford Roman Economy Project working paper*.

Huebner, S.R. (2025). Reconstructing the Nile: climate, agriculture, and socio-economic change in Roman Egypt. *Journal of Ancient History* 13(2), 307–362.

Hume, D. (1752). Of money. In *Political Discourses*. Edinburgh.

Katsari, C. (2011). *The Roman Monetary System: The Eastern Provinces from the First to the Third Century AD*. Cambridge University Press.

Kiyotaki, N. and Wright, R. (1989). On money as a medium of exchange. *Journal of Political Economy* 97(4), 927–954.

Klein Goldewijk, K., Beusen, A., Doelman, J. and Stehfest, E. (2017). Anthropogenic land use estimates for the Holocene—HYDE 3.2. *Earth System Science Data* 9, 927–953. Version 3.5 (2025), Utrecht University and PBL Netherlands Environmental Assessment Agency.

Lucas, R.E. (1980). Two illustrations of the quantity theory of money. *American Economic Review* 70(5), 1005–1014.

McCandless, G.T. and Weber, W.E. (1995). Some monetary facts. *Federal Reserve Bank of Minneapolis Quarterly Review* 19(3), 2–11.

McConnell, J.R. et al. (2018). Lead pollution recorded in Greenland ice indicates European emissions tracked plagues, wars, and imperial expansion during antiquity. *Proceedings of the National Academy of Sciences* 115(22), 5726–5731.

McConnell, J.R. et al. (2019). Pervasive Arctic lead pollution suggests substantial growth in medieval silver production modulated by plague, climate, and conflict. *Proceedings of the National Academy of Sciences* 116(30), 14910–14915.

Obstfeld, M. (1986). Rational and self-fulfilling balance-of-payments crises. *American Economic Review* 76(1), 72–81.

Online Coins of the Roman Empire and Nomisma.org (2026). American Numismatic Society, `https://numismatics.org/ocre/` [accessed July 2026].

Rathbone, D.W. (1996). Monetisation, not price-inflation, in third-century AD Egypt? In C.E. King and D.G. Wigg (eds.), *Coin Finds and Coin Use in the Roman World*, Berlin, 321–339.

Redish, A. (2000). *Bimetallism: An Economic and Historical Analysis*. Cambridge University Press.

Rolnick, A.J. and Weber, W.E. (1986). Gresham's law or Gresham's fallacy? *Journal of Political Economy* 94(1), 185–199.

Rolnick, A.J., Velde, F.R. and Weber, W.E. (1996). The debasement puzzle: an essay on medieval monetary history. *Journal of Economic History* 56(4), 789–808.

Sargent, T.J. (1982). The ends of four big inflations. In R.E. Hall (ed.), *Inflation: Causes and Effects*, University of Chicago Press, 41–97.

Sargent, T.J. and Velde, F.R. (2002). *The Big Problem of Small Change*. Princeton University Press.

Scheidel, W. (2010). Real wages in early economies: evidence for living standards from 1800 BCE to 1300 CE. *Journal of the Economic and Social History of the Orient* 53(3), 425–462.

Spoerri Butcher, M. (2019). Reka Devnia (Bulgaria): the challenges of creating a digital dataset of 80,000 coins. In B. Callegher (ed.), *Too Big to Study?*, Trieste, 161–170.

Strauss, J., Wilson, A.I. and Flohr, M. (2017). *OxREP Shipwrecks Database*, Version 2.1. Oxford Roman Economy Project, `https://oxrep.classics.ox.ac.uk/databases/shipwrecks_database/`.

Temin, P. (2013). *The Roman Market Economy*. Princeton University Press.

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

Velde, F.R., Weber, W.E. and Wright, R. (1999). A model of commodity money, with applications to Gresham's law and the debasement puzzle. *Review of Economic Dynamics* 2(1), 291–323.