

# 1 Historical outcomes

As a way of validating the calibrated model, we use it “out of sample” to predict the time path of urbanization and informal share for a set of 18 historical countries beginning in 1800. These countries started from a similar initial level of urbanization, 12.5%, as our sample of developing countries. They also experienced declines in urban mortality, of about 22.5 per thousand, as in our baseline sample, where urban mortality fell by about 25 per thousand. The significant *difference* for the historical countries is that their UMT took place over the course of a century-plus, versus a few decades in the case of our developing countries.

In Table 4 we show the results of running our calibrated model, matching the average initial conditions of the 18 historical countries in 1800, but leaving all other parameters of the model, including the location elasticities calibrated using our baseline sample, shown in Table 2. We have not built the model to match any future data from these 18 countries.

The first row shows the observed averages from the 18 historical countries..... XXXXX REMI - either you can fill in the sources for the observed data, or let me know where it comes from, and I can fill in this paragraph describing the row of Observed data in the table. XXXXX

Row 2 shows the results of simulating the model with the historical initial conditions, and assuming an exogenous fertility process the same as in our baseline calibration. As can be seen, for 1900 urbanization in the model (40.9 percent) is reasonably close to urbanization in the data (37.9 percent). The predicted informal rate is 35.6%, which is similar to the 30% value that we set from the available data. We do not want to overstate our ability to match this, given our limited sources for actual informal shares, but the model does not deliver a prediction that seems wildly at odds with reality. We do overstate the actual urban size by a significant amount (20.5 in the model versus 7.7 in the data), which we can attribute to the exogenous fertility process as well as the fact that our model has no out-migration, unlike the historical countries. By 1950, the model again does a good job of matching the urbanization rate (58.5 percent versus 57.9 percent in the data), and the informal share (15.8 percent versus 15 percent in the data), but again we do not want to overstate that match. The simulation does a reasonable job in matching the increase in urbanization over time, while also capturing the fact that the informal share of urban areas was falling.

In row 3, we simulate the model again, this time allowing for an endogenous fertility response with the same parameters we set in our baseline sample. In the case, the model again does a reasonable job of matching the observed urbanization rates in both 1900 and 1950, and we again capture the declining informal share over time. With endogenous fertility, the implied urban sizes in our simulation become more reasonable, predicting that urban population was 10.8 times larger in 1900 (versus 7.7 in the data), and 25.7 times larger in 1950 (versus 14.7 in the data). Again, without allowing for out-migration, or a more nuanced demographic transition, we continue to overstate population growth, but the composition of that population across urban locations appears reasonable.

Aside from validating the model, the historical simulations give us another means of illustrating the importance of the UMT. In rows 4 and 5, we simulate the model again using the historical starting conditions, but impose a rapid UMT. We replicate the simulations in rows 2 and 3, changing *only* the speed of the mortality transition to match that of our baseline sample of developing countries. With exogenous fertility (row 4), by 1900 our historical sample would have been at 46.3 percent urban, and 50 percent informal, both higher than the observed rates and higher than our simulation with a slow UMT. As expected with

lower mortality, the implied urban size would have been more than twice as large (a ratio of 50 rather than 20.5), while the informal share would have been roughly fifteen percentage points larger (50 percent rather than 35.6 percent). By 1950, the urbanization rate would have been similar to that with a slow mortality transition (61.3 versus 58.5 percent), but the informal share would have been nearly twice as big (29.0 versus 15.8 percent). Urban size would have been close to three times larger, in addition (a ratio of 98.6 versus 37.9).

With endogenous fertility (row 5), a rapid UMT would have caused an even greater divergence from the observed outcomes, as the lower living standards led to higher fertility. One should not take the reported outcomes at face value; the simulated urban sizes reach improbably ratios (381.3 in 1900 and over 18 thousand in 1950). The simulations are useful, however, because they show that a fast UMT can push an economy into a vicious spiral. A rapid UMT is capable of producing rapid urbanization (64.4 percent in 1900 or 91.7 percent in 1950) that is almost entirely informal (78.1 percent in 1900 or 91.8 percent in 1950), a significant departure from the slow UMT simulations in rows 2 or 3. The significant difference in results between the slow and rapid UMT reaffirms the results we showed in Table 3 using our sample of developing countries.

Table 1: **HISTORICAL OUTCOMES USING THE CALIBRATED MODEL, 1800-1950**

Scenario	1800:			1900:			1950:		
	Urb. Rate	Inf. Rate	Urb. Size	Urb. Rate	Inf. Rate	Urb. Size	Urb. Rate	Inf. Rate	Urb. Size
1. <i>Observed data:</i>	12.6	50.0	1.0	37.9	30.0	7.7	57.9	15.0	14.7
2. Calibrated model, exog. fert.	12.6	50.0	1.0	40.9	35.6	20.5	58.5	15.8	37.9
3. Calibrated model, endog. fert.	12.6	50.0	1.0	38.1	25.0	10.8	57.9	11.6	25.7
4. Calibrated model, exog. fert., fast UMT	12.6	50.0	1.0	46.3	50.0	50.0	61.3	29.0	98.6
5. Calibrated model, endog. fert., fast UMT	12.6	50.0	1.0	64.4	78.1	381.3	91.7	91.8	18399.3

Notes: Row 1 of the table shows the average urbanization rate and informal share, as well as the relative size of the total urban population, for a set of 18 historical countries. See text for the exact countries and the sources of the data. Row 2 shows the results of simulating our model, using the parameters listed in Table 2, except for setting the initial urbanization rate to match the historical data in 1800, setting the initial demographic rates to match averages from the 18 countries, and setting the half-life of the UMT to be 50 years (as opposed to 3 in our baseline model). Row 3 replicates Row 2, but replaces the exogenous fertility process with the endogenous fertility process described in the Appendix. In Rows 4 and 5, we simulate the model again using the same conditions in Rows 2 and 3, respectively, but setting the half-life of the UMT to be only 3 years (matching the UMT in our baseline).