Supplementary material 1: Data mismatch

The employment data used in this paper is available from different official sources, such as the National Population Census (NPC), the National Economic Census (NEC), and the Shanghai Bureau of Statistics (SBS). Discrepancies in employment totals at the city level are inevitable due to the differences in data calibre. For instance, a respondent is considered as "employed" if s/he has paid work for over an hour during the survey week in the NPC. As for the NEC, legal entities self-report their number of employees. Compared with these censuses, which are conducted every 5-10 years, SBS employment data is released yearly and takes the census results into account once available. However, SBS doesn't amend its previously-released statistics except for the previous year. Adjustments have therefore to be made manually to bridge the mismatches and to fit with the overall trends. Note that this research region is considered as a closed system: the number of employed residents is assumed equivalent to the number of employed workers across the study area. Employment at the zonal level is then recalculated according to the city-level adjustments pro rata.

Table. Mismatches in employment data in Shanghai.

	2000	2005	2010	2015
Employed residents (from NPC)	8,409,778	9,087,444	12,803,286	13,003,281
Employed workers (from NEC)	9,187,055	9,081,258	10,801,916	13,055,184
Employed workers (from SBS)	7,522,600	8,633,200	10,907,600	13,615,100
Adjusted employment totals	7,903,050	8,820,040	11,564,096	13,615,100

Notes: The number of NPC employed residents is expanded from a 10% sample; the number of employed workers from NEC is extrapolated or interpolated from the years 2001, 2004, 2008 and 2013 assuming linear growth; The employment totals are adjusted based on the original and amended (if available) SBS data (2000-2017) to avoid sudden changes between years.

Supplementary material 2: Model calibration

With the average income and proportion of spending on housing at the city and regional level as known inputs, the average housing rent can be calculated. However, the initial model only assumed the same proportion of spending on housing across the city-region, while in reality the central residents should have more available housing budget. The rental gaps between the categorical zones from the online data (Figure 2(d)) were therefore used to calibrate the zonal rental patterns in the model.

In addition, we used the zone-to-zone commuting time matrix and overall statistics in the official transport survey to calibrate the commuting patterns. The 2000/2010 model calibrated the average commuting distance to 7.9/8.3km, compared with 8.6 km as reported in the Shanghai Comprehensive Travel Survey in 2014. Table 1 reports the modelled commuting matrix (in percentages) of 2010, while the MPD data set is included as a reference group. The comparison largely validates the inter-zonal commuting pattern. The discrepancies in intra-zonal commuting ratio in the city centre and the far suburbs can be explained by the potential biases in using MPD, as people living in rural areas tend to have lower mobile phone ownership rates.

Table 1. Distribution of commuting journeys from mobile phone data and the model.

	Mobile phone data (2014)				Modelled distribution (2010)			
	Centre	Near suburb	Sub-centres	Far suburb	Centre	Near suburb	Sub-centres	Far suburb
Centre	44%	4%	1%	1%	40%	2%	1%	1%
Near suburb	7%	17%	1%	1%	7%	16%	1%	1%
Sub-centres	0%	0%	7%	2%	1%	0%	9%	1%
Far suburb	1%	1%	1%	13%	1%	0%	1%	20%

Notes: The two matrices both sum up to 100%.

Supplementary material 3: Modelling key inputs and outputs

Table. Main inputs by categorical zones.

Scenarios	Jobs (million)					Housing supply (million units)				
	Centre	Near suburb	Sub-centres	Far suburb	Total	Centre	Near suburb	Sub-centres	Far suburb	Total
2010	5.7	2.1	1.2	2.6	11.6	4.4	2.1	0.8	1.9	9.2
S0 – Trend (dispersion)	5.7	2.4	1.4	3.0	12.4	4.4	2.8	1.2	2.8	11.2
S0a – Suburbanisation with job decentralisation	5.0	2.7	1.5	3.3	12.4	4.4	2.8	1.2	2.8	11.2
S0b – Monocentricity with job centralisation	6.1	2.5	1.2	2.6	12.4	4.4	2.8	1.2	2.8	11.2
S1 - Housing supply increase	5.7	2.4	1.4	3.0	12.4	4.4	2.6	2.0	2.2	11.2
S2 - Job provision	5.7	2.2	1.9	2.7	12.4	4.4	2.8	1.2	2.8	11.2
S3 - Coordinated jobs and housing growth	5.7	2.2	1.9	2.7	12.4	4.4	2.6	2.0	2.2	11.2

Table. Main outputs by categorical zone.

Scenarios	Employed residents (million)					Housing rents (1,000 ¥/unit/year)				
	Centre	Near suburb	Sub-centres	Far suburb	Total	Centre	Near suburb	Sub-centres	Far suburb	Total
2010	5.0	2.8	1.1	2.6	11.6	25.8	17.9	12.1	11.5	19.8
S0 – Trend (dispersion)	4.5	3.4	1.4	3.1	12.4	48.4	33.5	21.3	19.3	34.5
S0a – Suburbanisation with job decentralisation	4.1	3.5	1.5	3.4	12.4	45.2	34.4	23.3	21.2	34.5
S0b – Monocentricity with job centralisation	4.8	3.6	1.2	2.8	12.4	50.3	34.3	19.0	17.4	34.5
S1 - Housing supply increase	4.7	3.2	1.7	2.8	12.4	49.9	34.1	15.5	21.8	34.5
S2 - Job provision	4.5	3.2	1.9	2.9	12.4	48.4	32.1	26.9	18.1	34.5
S3 - Coordinated jobs and housing growth	4.7	3.0	2.2	2.6	12.4	49.9	32.5	18.9	20.3	34.5