International trade, dietary change, and TB control in India: Application of a fully-integrated macroeconomic-epidemiological model framework

Henning Tarp Jensen, Marcus Keogh-Brown, Susmita Chatterjee, Yoko Laurence,

Sedona Sweeney, Anna Vassal, Tom Sumner

Abstract: We apply a newly established fully integrated macroeconomic-epidemiological-demographic model framework for India to analyse the impact of eliminating import tariffs on the macroeconomic and health burdens of tuberculosis (TB) in India over 2021-2040. We find that full tariff elimination will increase Indian Net Present Value (NPV) of GDP by 103bn USD (or 2.93USD/person/year), but also that this covers tax efficiency gains of 119bn USD combined with TB health-related NPV GDP losses of 15bn USD. Substitution in household consumption away from less expensive sources of kcal's, including cereals and processed rice, and towards consumption of more expensive sources of kcal's, including meat products and edible oils, implies that Indian households will experience increasing prevalence levels of low BMI, and this will act as a risk factor for TB and lead to increases in TB incident cases (˜9.5million) and TB case fatalities (˜1.3million), as well as increasing excess deaths (˜1.1million) and declining population levels (9.2million person-years) over 2021-2040. In spite of increasing income and consumption levels, we also find that Indian low-income households will experience relatively low expansion of NPV household consumption, and excessive adverse impacts on population demographics, including population levels and excess deaths. Hence, we find that import tariff elimination trade reform in India is likely to involve both overall efficiency-related welfare gains, but also adverse overall health impacts as well as a worsening welfare inequalities and TB-related health inequities.

Introduction

The global burden of disease (GBD) of Tuberculosis (TB) has been reduced by >40% over the past two decades, but TB, nonetheless, continues to be the 12th leading cause of GBD and account for >2% of the All Cause GBD in 2019 (GBD Diseases and Injuries Collaborators 2020), and the 13th leading cause of death worldwide and the top cause of deaths from a single infectious agent in 2019 (WHO 2021). Due to the advent of Covid-19, TB was only the 2nd leading cause of deaths from a single infectious agent in 2020, but the Covid-pandemic, nonetheless, meant that TB-related deaths increased by 5.6% to >1.5 million in 2020, thereby reversing the yearlong downward global trend in TB-related deaths and increasing for the first time since the early 2000s (ibid.)

Regardless of the reversal of global trends, the global community has committed itself to end TB by 2030. Hence, Goal 3.3 of the Sustainable Development Goals (SDG) indicates "(B)y 2030, end the epidemics of AIDS, tuberculosis, malaria and neglected tropical diseases and combat hepatitis, water-borne diseases and other communicable diseases" (WB 2015), while the End TB Strategy, which was endorsed by the World Health Assembly in May 2014, spans the period 2015-2035 and sets milestones for 2020 and 2025 and targets for TB incidence rates in 2030 (80% reduction compared to 2015 levels) and 2035 (90% reduction compared to 2015 levels) (WHO 2015).

Similar to global trends, India has made progress in reducing their TB disease burden over the past two decades. However, while India has reduced their share of global All Cause GBD, from 20.7% to 18.5%, they continue to account for roughly one-third of the global TB GBD (GBD 2020). Hence, while India has generally improved health outcomes beyond the global average, this doesn't apply to India's TB health outcomes, with the Indian share of global TB cases amounting to 26% of global TB cases and 27% of global Rifampicin-resistant RR-TB cases in 2019 (and with 78% of India's RR-TB cases being Multidrug-resistant MDR-TB cases) (WHO 2020).

TB is a curable and preventable disease. It has previously been estimated that three key risk factors accounted for >11 million DALYs in 2015, including smoking (2.8 million DALYs), Alcohol use (4.7 million DALYs), and Diabetes (3.8 million DALYs) (GBD Tuberculosis Collaborators 2018). Other key TB risk factors include HIV co-morbidity, in-door air pollution, and low Body Mass Index (BMI) (Oxlade and Murray 2012). WHO guidelines on the prevention (and elimination) of TB recommends the adoption of 'programmatic management of tuberculosis preventive treatment' involving a 'cascade of care'-approach whereby people most at risk of developing TB are identified and offered Tuberculosis Preventive Treatment (TPT).

In this paper, we investigate whether a non-pharmaceutical trade policy intervention, involving full elimination of import tariffs in India, may be able to positively affect nutritional intakes, and by

implication lower the prevalence of low BMI which acts as a TB risk factor in the Indian population, and thereby support WHOs preferred TPT-focussed pharmaceutical 'cascade of care' prevention strategy. Specifically, we apply a newly established fully integrated macroeconomic-epidemiological-demographic model framework for India, developed along the lines of the framework set out by Sumner et al. (2022), to analyse the impact of eliminating import tariffs on the macroeconomic and health burdens of tuberculosis (TB) in India over 2021-2040.

In a previously published study, we constructed a fully-integrated macroeconomic-epidemiological-demographic framework to analyse how import tariffs affect the twin macroeconomic and health burdens of nutrition-related cholesterol build-up and related occurrences of stroke and IHD in the Thai population (Jensen et al. 2019). In this paper, we will similarly study the macroeconomic and health burden impacts of import tariffs but via a different nutrition-related health pathway, where nutrition acts as a risk factor for TB illness. Our macroeconomic approach follows in the nascent tradition of studying health implications of trade policy interventions. This literature includes both descriptive and statistical studies ranging from simple cross-country correlation analyses of unhealthy and imported food expenditure shares (Estimé, Lutz, & Strobel, 2014) and difference-in-difference statistical analyses of natural experiments of free trade agreements and WTO accessions (Baker et al. 2016; Barlow et al. 2017; Schram et al. 2015) to structural statistical approaches (Baker et al. 2016). The past literature is, however, marred by problems of poorly defined and specified exposures and mechanisms, and there continues to be a need for "more methodologically rigorous and consistent approaches in future quantitative studies" (Cowling, Thow, & Porter 2018).

To our knowledge, there are only a few CGE model studies analysing TB (Taylor et al. 2014; Ahmed et al. 2017), and they are focussed on the economic implications of anti-microbial resistance (AMR) across a range of illnesses, including HIV, TB, and Malaria. Moreover, they rely on simple linear projections of mortality, derived from various ad hoc scenarios of future anti-microbial resistance development, but without application of detailed epidemiological modelling. Our macroeconomic CGE model framework therefore provides a novel macroeconomic-epidemiological-demographic tool for TB-related policy analysis.

Methods

We construct a recursively-dynamic Computable General Equilibrium (CGE) model for India by calibrating the IFPRI Standard Model framework (Löfgren, Lee Harris & Robinson 2002) to a 2014 Social Accounting Matrix (SAM) for India, derived from the GTAP10 data base (Aguiar et al. 2019) and, by expanding the model to a dynamically-recursive model framework via addition of a full set of factor ownership updating-equations (see appendix A for details on the CGE model calibration). The labour factor updating equations are linked to the demographic model, which is calibrated to the Indian population projections from the World Population Prospects: 2019 Revision (UN 2019) (see appendix B for details on the demographic model calibration), while the capital factor ownership updating equations are calibrated to capital stock and depreciation data from the Penn World Tables 10.0 (Feenstra Inklaar & Timmer 2015). The CGE model includes three production factors (unskilled and skilled labour, and capital) and five representative households which are stratified according to income quintiles and derived from the 68th National Sample Survey 2011-12 for India (NSSO 2018).

The epidemiological model for TB in India (see appendix C for details on the epidemiological model calibration) is a compartmental epidemiological model (Figure 1). It has been developed, specifically, to capture treatment interventions and their ability to return infected individuals to a state of latent infection, and prevention interventions and their ability to affect risk factors and, by implication, the risk of moving from being susceptible to being latently infected or infected (Oxlade and Murray 2012). Our model framework captures six risk factors include Smoking, Indoor air pollution, Low BMI (<18.5), Alcohol intake, Diabetes, and HIV-positive (HIV+). Household-specific prevalence rates (p_{ih}) and risk-factor-specific relative risks (g_j) are presented in Table 1 along with household-specific Relative Risks of TB (RR_h) which are calculated using the following formula (1):

(1) $$RR_h = \exp\left(\sum_i g_i p_{ih}\right)$$

The compartmental epidemiological model is a discrete-time version of the standard continuous-time model framework, which is specified with weekly discrete time intervals and with parameters specified and calibrated accordingly. In order to match health outcomes with the annual time steps of the macroeconomic CGE model, annual prevalence and population aggregates are calculated as averages of weekly outcomes, while annual incidence, mortality and notification aggregates are calculated as sums of weekly outcomes across a given year. This implies that annual health-related labour force and health cost changes are calculated from respectively the sum/average of the 52 weekly epidemiological model solutions, applied to the same year demographic model, and imposed on the same year CGE model. Furthermore, the epidemiological model receives feedback from the CGE model via specification of a nutritional module which computes household-specific changes in

the distribution of kilocalorie (kcal) consumption intakes, and related changes in the distribution of population weight and BMI (see appendix C.4 for details of the nutritional module) based on consumer demand changes derived from a set of household-specific Almost Ideal Demand Systems (AIDS) (see appendix A.1 for details on our calibration of household-specific AIDS demand systems), and this allows for computing changes in household-specific prevalence rates of low BMI, a key risk factor for development of TB, and thereby to achieve full integration between the CGE, demographic, and epidemiological sub-models.

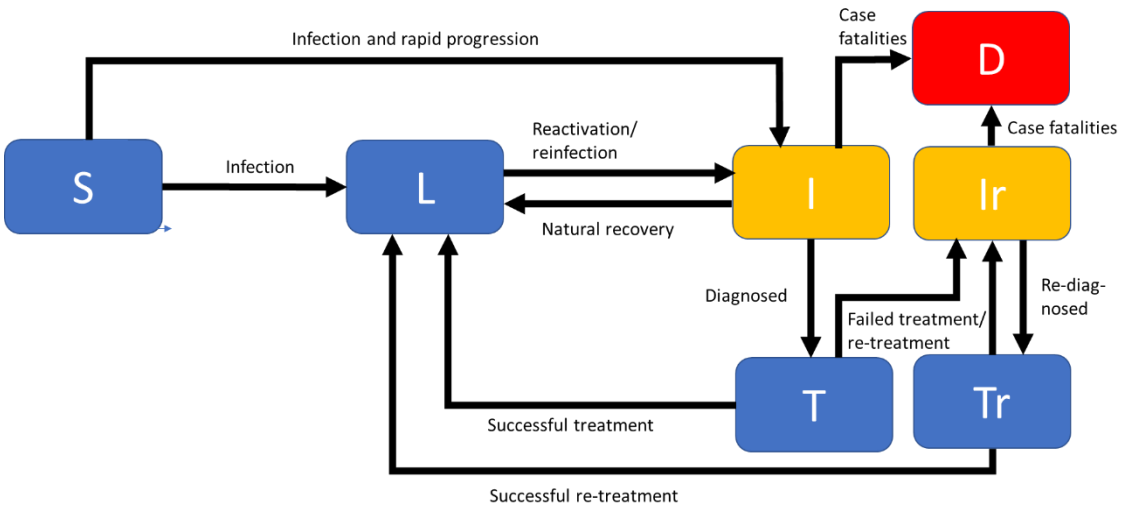

Figure 1. Epidemiological compartment-model for TB in India

Table 1. Risk Factor prevalence rates ($p_{i,h}$), Relative Risks of Risk Factors (g_i), Relative Risks of TB (RR_h)

Risk factor	**H01** ($p_{i,'H01'}$)	**H02** ($p_{i,'H02'}$)	**H03** ($p_{i,'H03'}$)	**H04** ($p_{i,'H04'}$)	**H05** ($p_{i,'H05'}$)	g_i
Smoking	0.179	0.163	0.1415	0.1235	0.0965	2.0
Indoor air pollution	0.791	0.685	0.510	0.252	0.083	1.4
Low BMI (<18.5)	0.339	0.281	0.218	0.167	0.111	2.1
Alcohol intake	0.031	0.022	0.019	0.013	0.0098	2.9
Diabetes	0.005	0.007	0.009	0.013	0.020	3.1
HIV+	0.004	0.004	0.005	0.004	0.002	26.7
RR_h	2.0008	1.8156	1.6133	1.3993	1.2421	

Note: Risk factor prevalence rates (p_{ih}) from Indian NFHS-4 (IIPS 2017) and RRs from Oxlade and Murray (2012)

Key features of the structure of the Indian economy, with particular relevance for the trade policy analyses in this paper, are presented in Table 2. Effective import tariff rates (TM) are varying across primary and secondary commodity sectors. Imported food staples such as paddy rice and wheat carry effective tariff rates of respectively 6.2% and 19.2%, while imports of other cereal grains are taxed more lightly at the border (0.9%). It can also be noted that processed food imports, including meat products, vegetable oils and fats, dairy products, sugar, and other processed foods are generally taxed more heavily at the border (>20%). That being said, most of the primary and processed food commodities have low import shares (M/Q) with imports (M) generally accounting for <4% (and mostly <1%) of domestic supplies (Q). The main exceptions are 'vegetables, fruits, and nuts' with an import share of 10.6%, and 'vegetable oils and fats' with an import share of 20.5%.

Table 2. Structure of Indian Economy

Commodity Sector	TM	TS	E/X	M/Q
Paddy rice	6.2%	-10.5%	0.2%	0.0%
Wheat	19.2%	-15.1%	6.0%	0.1%
Cereal grains nec	0.9%	-4.5%	21.0%	0.3%
Vegetables, fruits, and nuts	21.7%	-1.7%	5.4%	10.6%
Crops nec	26.5%	-12.1%	9.1%	2.5%
Bovine cattle, sheep and goats, horses	15.5%	-23.7%	0.6%	0.1%
Animal products nec	6.0%	-1.1%	1.9%	0.9%
Raw milk	0.0%	0.5%	0.2%	0.0%
Wool, silk-worm cocoons	13.7%	-0.4%	5.0%	10.3%
Forestry	8.2%	0.0%	0.8%	5.8%
Fishing	11.6%	-1.9%	2.2%	0.4%
Mining Products	1.9%	0.2%	6.8%	71.3%
Bovine meat products	3.9%	-1.3%	58.3%	0.2%
Meat products nec	16.2%	-2.9%	1.7%	1.4%
Vegetable oils and fats	48.3%	1.3%	7.8%	20.5%
Dairy products	22.5%	-3.3%	1.2%	0.2%
Processed rice	5.2%	-12.5%	13.6%	0.0%
Sugar	56.3%	-0.7%	7.0%	3.9%
Food products nec	35.3%	-2.8%	14.6%	1.6%
Beverages and tobacco products	89.6%	33.8%	4.3%	1.2%
Textile and leather products	11.4%	0.6%	24.7%	4.6%
Other manufacturing	7.0%	8.1%	17.8%	16.1%
Utilities	0.0%	-8.8%	0.0%	0.3%
Construction	0.0%	2.0%	0.3%	0.2%
Trade	0.0%	-0.5%	1.0%	0.4%
Transport	0.0%	0.6%	10.3%	3.5%
public administration	0.0%	0.7%	0.4%	0.5%
education services	0.0%	1.5%	1.3%	1.6%
health services	0.0%	0.9%	0.8%	0.8%
recreation services	0.0%	32.4%	11.3%	12.0%
other services	0.0%	2.1%	18.0%	5.1%
Average	5.8%	2.1%	10.3%	10.7%

Source: 2014 India SAM (Aguiar et al. 2019)

Notes: TM – import tariff rates, TS – sales tax rates, E – Exports, M – Imports, X – Domestic production, Q – Domestic supply

Key features of the structure of TB clinical outcomes in India are presented in Table 3. In 2020, the total number of incident cases was 2.65 million, (192 per 100,000) including 53,000 incident cases among HIV+ citizens, and the total number of case fatalities was 504,000 (37 per 100,000) including 11,000 case fatalities among HIV+ citizens.

Table 3. Structure of TB disease in India (2020)

	Number	Rate per 100 000 population
Total TB incidence	2,590,000	188
HIV-positive TB incidence	53,000	3.8
HIV-negative TB mortality	493,000	36
HIV-positive TB mortality	11,000	0.78

Source: WHO (2022)

Our macroeconomic model closure (see appendix A.5 for details) specifies the GDP deflator as price numeraire. The GDP deflator was kept fixed at the counterfactual GDP price level growth path, based on a fixed annual inflation rate of 4.6% (see appendix A.2). For the purpose of calculating Net Present Value (NPV) impacts over our 2021-2040 time horizon, we furthermore assumed a nominal discount rate of 10%, implying a real discount rate of 5.4% which is consistent with average real interest rates in India over 2019-20 (WB 2021).

Our counterfactual growth path (see appendix A.4 for details) was simulated with a standard neoclassical model closure involving savings-driven investment to clear the capital account, and a balanced macro-closure to ensure that the government consumption share of domestic absorption remained fixed along the counterfactual growth path. Our trade policy scenario was, similarly, simulated with a standard neoclassical macro closure, involving savings-driven investment to clear the capital account, but it differed from the counterfactual simulation by keeping the core real government consumption 'fixed' at the counterfactual growth path (in terms of its own price deflator). Hence, our policy simulation only allowed government consumption to vary in response to changes in TB-related health costs and changes in the relative price of public services. Combined with a standard revenue-neutral government budget closure, this meant that, except for relative price changes, government savings (which adjusts to clear the government account) varied one-to-one with changes in TB-related health costs.

In the following section, we present the results of our import tariff elimination policy simulations. The discussion will focus on aggregate results over 2021-40, and we will present these results in two ways: Macroeconomic aggregates are presented in NPV terms (in 2021 prices) with future values discounted at 10% p.a., while household-specific aggregates are presented as cumulative real values (in 2021 prices) but without taking account of the time value of money.

Results

The results of our import tariff elimination policy simulations are presented in Tables 4-7, including decompositions of the 2021-2040 NPV GDP impact between efficiency gains and nutrition-related TB health externality impacts (Table 4) and across final demand components (Table 5), a decomposition of household-specific cumulative real income impacts between efficiency gains and health externality impacts (Table 6), and an overview of household-specific impacts on NPV household income and consumption, skilled and unskilled labour supplies, population and excess death demographics, and TB incident case and case fatality clinical outcomes (Table 7).

Overall, import tariff elimination will increase Indian NPV GDP by 103bn USD over 2021-2040 (equivalent to 2.93USD/person/year) (Table 4). This covers an NPV GDP efficiency gain of 119bn USD and a TB health externality loss of 15bn USD. The health externality loss can be further decomposed into NPV GDP losses due to increased TB treatment costs (2bn USD), increased TB morbidity (2bn USD), and TB mortality (11bn USD). These results indicate that the nutrition-related 'low BMI' risk factor represents an important pathway between trade policy, in this case represented by import tariff reform, TB clinical outcomes, and macroeconomic impacts. Hence, this health pathway reduces the NPV GDP efficiency gains by one-seventh, and by implication accounts for ≈-15% of the net NPV GDP gains of import tariff elimination. Our results also confirm that mortality accounts for the majority of the NPV GDP health externality losses (Table 4) due to the snowballing impact of mortality on household-specific labour supplies (Table 7).

However, the macroeconomic NPV GDP gain of 103bn USD is accompanied by a significant increase in TB health burden, including increases in TB incidence (≈9.5million cases) and TB mortality (≈1.3million case fatalities), as well as increasing excess deaths (≈1.1million deaths) and declining population levels (≈9.2million person-years) over 2021-2040 (Table 7). The import tariff elimination reform therefore represents a trade-off between macroeconomic efficiency gains and increased TB-related health burdens.

Table 4. 2021-2040 Import Tariff Elimination - ΔNPV GDP impacts (bn USD)

Disease Burden			
GDP PER CAPITA	USD	% of agg	% of total
- ΔReal GDP/capita/year (USD)	2.93	0.15%	100.0%

NPV of GDP (2021-2040) (bn USD)	bn USD	% of agg	% of total
- ΔNPV GDP	103	0.15%	100.0%
- ΔNPV GDP (Tariff elimination w/o health impacts)	119	0.17%	114.7%
- ΔNPV GDP (health impacts)	-15	-0.02%	-14.9%
- Treatment Costs	-2	0.00%	-1.6%
- ΔLabour supply - Morbidity	-2	0.00%	-2.2%
- ΔLabour supply - Mortality	-11	-0.02%	-11.0%
Memorandum Items (2021-2040):			
- Real GDP/capita/year (USD)	1,939		
- NPV GDP (bn USD)	68,198		

Note: Own calculations

The breakdown of NPV GDP gains across final demand components (Table 5) indicates that household and government consumption expands by 0.10-0.14%, while investment declines by 0.04% in NPV terms over 2021-2040. NPV exports and NPV imports expands due to a permanent real exchange rate depreciation (not shown), which creates additional incentives for exporters to export and thereby cover the foreign exchange costs of increased imports, which occur since the depreciation only partially compensates for the tariff-induced reduction in import prices.

The expansion of NPV household consumption (Table 5) stems from efficiency-related improvements in factor income generation (Table 7). Interestingly, the expansion of household consumption is accompanied by declining nutritional kcal intakes, and therefore by increasing prevalence of low BMI among the Indian population (not shown). This is caused by income-related substitution in household demand away from consumption of less expensive sources of kcal's including of cereals and processed rice, and towards consumption of more expensive sources of kcal's including meat products and edible oils (not shown).

The expansion of NPV government consumption (Table 5) stems partly from increased relative prices of domestic services caused by the real exchange rate depreciation (not shown), but mostly from increased TB treatment costs due to nutrition-related increases in TB incident cases (Table 7). In spite of the revenue-neutral government budget closure, the tariff elimination therefore leads to increasing government budget deficits due to the health externality impact on healthcare costs, and this reduces overall savings, and crowds out investment, in the Indian economy. In other words, while the tariff elimination is fully financed by increased household taxation, the health externality

leads to declining investment and capital accumulation, and thereby to reduced NPV GDP efficiency gains over time (not shown).

Table 5. 2021-2040 Import Tariff elimination - ΔNPV Final Demand (bn USD)

Disease Burden			
NPV of GDP (2021-2040)	bn USD	% of agg	% of total
- ΔNPV GDP (bn USD)	103	0.15%	100.0%
- ΔNPV Household Consumption	66	0.10%	63.7%
- ΔNPV Government Consumption	95	0.14%	91.9%
- ΔNPV Investment	-27	-0.04%	-26.0%
- ΔNPV Exports	928	1.36%	897.3%
- ΔNPV Imports	959	1.41%	926.9%
Memorandum Items (2021-2040):			
- NPV GDP (bn USD)	68,198		

Note: Own calculations

The household-specific impacts of import tariff elimination (Tables 6-7) show that, while the low-income household H01 (income quintile 1) is experiencing the largest cumulative losses of population (>3.4million person-years) and labour supplies (>1.4 million person-years), they also experience the lowest gains in terms of NPV household income (33.1bn USD) and NPV household consumption (2.5bn USD). In contrast, the high-income household H05 (income quintile 5) is experiencing the smallest cumulative losses of population (<0.9million person-years) and labour supplies (<0.4 million person-years), and the highest gains in terms of both NPV household income (≈600bn USD) and NPV household consumption (39.3bn USD).

The household-specific indicators also demonstrate that low-income households are experiencing the highest increases in TB incidence (≈3.4 million cases) and TB mortality (≈465,000 case fatalities) as well as the highest increase in excess deaths (≈401,000 deaths) (Table 7). Hence, while the high-income households pay higher direct taxes (not shown), and while the low-income households experience roughly the same relative income gains from improved factor returns as other households (1.65%), the low-income households are experiencing adverse distributional impacts in terms of NPV household consumption, population demographics including population levels and excess deaths, and TB-related clinical health outcomes including incident cases and case fatalities.

Table 6. 2021-2040 Import Tariff Elimination - ΔNPV household income (YH) impacts (bn USD; 2021 prices)

	Households									
	H01 (Income Q1)		H02 (Income Q2)		H03 (Income Q3)		H04 (Income Q4)		H05 (Income Q5)	
HOUSEHOLD INCOME PER CAPITA	USD	Δ%	USD	Δ%	USD	Δ%	USD	Δ%	USD	Δ%
- ΔReal YH/capita/year (USD)	4.90	1.65%	6.61	1.67%	9.31	1.66%	16.24	1.65%	58.79	1.64%
NPV of HHLD INCOME (2021-2040)	bn USD	Δ%	bn USD	Δ%	bn USD	Δ%	bn USD	Δ%	bn USD	Δ%
- ΔNPV YH	33.1	1.62%	55.7	1.64%	84.8	1.64%	155.9	1.63%	596.8	1.63%
- ΔNPV YH (excl. TB health impacts)	34.7	1.70%	56.8	1.67%	85.9	1.65%	157.5	1.63%	603.8	1.60%
- ΔNPV YH (Treatment Costs)	-0.1	1.62%	-0.1	1.65%	-0.1	1.64%	-0.2	1.63%	-0.8	1.63%
- ΔNPV YH (ΔLabour supply - Morbidity)	-0.2	1.63%	-0.2	1.65%	-0.2	1.64%	-0.2	1.63%	-1.0	1.63%
- ΔNPV YH (ΔLabour supply - Mortality)	-1.3	1.69%	-0.8	1.67%	-0.8	1.66%	-1.2	1.64%	-5.1	1.64%
Memorandum Items (2021-2040):										
- Real YH/capita/year (USD)	296		397		561		987		3,577	
- NPV YH (bn USD)	2,044		3,387		5,163		9,578		36,667	

Note: Own calculations

Table 7. 2021-2040 Import Tariff Elimination - cumulative household impacts (bn USD; 2021 prices)

	Households HH (Income quintiles Q1-Q5)									
	HH_Q1		HH_Q2		HH_Q3		HH_Q4		HH_Q5	
HOUSEHOLD INCOME & CONSUMPTION	bn USD	Δ%	bn USD	Δ%	bn USD	Δ%	bn USD	Δ%	bn USD	Δ%
- ΔNPV Household income	33.1	1.62%	55.7	1.64%	84.8	1.64%	155.9	1.63%	596.8	1.63%
- ΔNPV Labour income	31.2	1.62%	49.7	1.64%	70.6	1.63%	120.1	1.61%	326.2	1.59%
- ΔNPV Capital income	1.9	1.68%	6.0	1.68%	14.1	1.68%	35.8	1.68%	270.7	1.68%
- ΔNPV Household consumption	2.5	0.14%	5.0	0.17%	7.3	0.16%	11.7	0.14%	39.3	0.16%
LABOUR MARKET	1000s	Δ%	1000s	Δ%	1000s	Δ%	1000s	Δ%	1000s	Δ%
- ΔUnskilled labour (#person-years)	-1,392	-0.09%	-852	-0.04%	-611	-0.03%	-397	-0.02%	-233	-0.01%
- ΔSkilled labour (#person-years)	-66	-0.09%	-71	-0.04%	-74	-0.03%	-92	-0.02%	-158	-0.01%
DEMOGRAPHY	1000s	Δ%	1000s	Δ%	1000s	Δ%	1000s	Δ%	1000s	Δ%
- ΔPopulation (#person-years)	-3,470	-0.07%	-2,151	-0.04%	-1,587	-0.03%	-1,124	-0.02%	-883	-0.01%
- ΔExcess deaths (#persons)	401	1.08%	266	0.85%	197	0.56%	141	0.35%	112	0.21%
TB CLINICAL OUTCOMES	1000s	Δ%	1000s	Δ%	1000s	Δ%	1000s	Δ%	1000s	Δ%
- ΔTB incident cases (#persons)	3,410	20.1%	2,248	14.1%	1,674	12.7%	1,207	11.7%	964	11.0%
- ΔTB case fatalities (#persons)	465	18.2%	306	12.7%	227	11.4%	164	10.5%	130	9.9%

Note: Own calculations

Conclusion

Our results show that Indian import tariff elimination, if fully implemented, would bring large NPV GDP efficiency gains amounting to 119bn USD over 2021-2040. However, our results also demonstrate that there would be additional negative TB-related pecuniary externalities amounting to 15bn USD, implying that the net gain to the Indian economy would only be 103bn USD or 2.93USD/person/year. Our TB-related health externality would therefore reduce the NPV GDP efficiency gains by one-seventh. Moreover, the negative health externality would increase excess deaths (≈1.1million deaths) and reduce population levels (≈9.2million person-years). Indian policymakers will therefore face a trade-off between macroeconomic efficiency gains and increased TB-related health burdens if they decide to pursue import tariff elimination reform.

Our distributional results further demonstrate that low-income households would experience adverse distributional impacts in terms of NPV household consumption, population levels and excess deaths. Indian policymakers would therefore also have to consider the trade-offs between economic efficiency gains, on the one hand, and increasing welfare inequality and health inequity, on the other. Overall, this demonstrates the importance, for Indian policy makers, to consider potential health externalities when assessing trade policy reforms which covers primary and processed foods, and which may therefore affect nutritional risk factors for TB illness and for other health conditions.

Reference list

Aguiar A, Chepeliev M, Corong E, McDougall R, Van Der Mensbrugghe D. 2019. The GTAP Data Base: Version 10. Journal of Global Economic Analysis 4(1):1-27.

Ahmed SA, Barış E, Go DS, Lofgren H, Osorio-Rodarte I, Thierfelder K. 2017, World Bank Policy Research Working Paper 8133. World Bank: Washington DC.

Baker P, Friel S, Shram A, Labonte R. 2016. Trade and investment liberalization, food systems change and highly processed food consumption: a natural experiment contrasting the soft-drink markets of Peru and Bolivia. Globalization and Health 12:24.

Barlow P, Mckee M, Basu S, Stuckler D. 2017. The health impact of trade and investment agreements: a quantitative systematic review and network co-citation analysis. Globalization and Health 13:13.

Estimé MS, Lutz B, Strobel F. 2014. Trade as a structural driver of dietary risk factors for noncommunicable diseases in the Pacific: an analysis of household income and expenditure survey data. Globalization and Health 10:48.

Feenstra RC, Inklaar R, Timmer MP. 2015. The Next Generation of the Penn World Table. American Economic Review 105(10):3150-3182. available for download at www.ggdc.net/pwt

GBD Diseases and Injuries Collaborators. 2020. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 396(10258):1204-22.

GBD Tuberculosis Collaborators. 2018. The global burden of tuberculosis: results from the Global Burden of Disease Study 2015. Lancet Infectious Diseases 18(3):261-84.

IIPS. 2017. National Family Health Survey (NFHS-4), 2015-16. Mumbai: International Institute for Population Sciences (IIPS).

Jensen HT, Keogh-Brown MR, Shankar B, et al. 2019. International trade, dietary change, and cardiovascular disease health outcomes: Import tariff reform using an integrated macroeconomic, environmental and health modelling framework for Thailand. SSM - Population Health 9:100435.

Löfgren H, Lee Harris R, Robinson S. 2002. A Standard Computable General Equilibrium (CGE) Model in GAMS. Microcomputers and Policy Research 5. International Food Policy Research Institute, Washington DC.

NSSO. 2018. National Sample Survey (NSS) 68th round 2011-12. Electronic data. Ministry of Statistics and Programme Implementation, Government of India. URL: http://microdata.gov.in/nada43/index.php/catalog.

Oxlade O, Murray M. 2012. Tuberculosis and poverty: why are the poor at greater risk in India? PLoS One 7(11):e47533.

Schram A, Labonte R, Baker P, Friel S, Reeves A, Stuckler D. 2015. The role of trade and investment liberalization in the sugar-sweetened carbonated beverages market: a natural experiment contrasting Vietnam and the Philippines. Globalization and health 11:41.

Sumner T, Jensen HT, Keogh-Brown MR, Vassall A. 2022. Considering the interaction between TB and the economy: time to integrate epidemiological and economic models. International Journal of Tuberculosis and Lung Disease (forthcoming).

Taylor J, Hafner M, Yerushalmi E, Smith R, Bellasio J, Vardavas R, Bienkowska-Gibbs T, Rubin J. 2014. Estimating the economic costs of antimicrobial resistance. Model and Results. RAND Research Report No 911. RAND Corporation: Santa Monica, Calif., and Cambridge, UK.

UN. 2019. Electronic data. World Population Prospects, the 2019 Revision. United Nations. URL: http://esa.un.org/unpd/wpp/index.htm.

WB. 2021. Electronic data. World Development Indicators (WDI) database. Washington DC: World Bank. URL: https://databank.worldbank.org/source/world-development-indicators.

WHO. 2015. Implementing the End TB Strategy: The Essentials. World Health Organization: Geneva.

WHO. 2020. Global Tuberculosis Report 2020. World Health Organization: Geneva.

WHO. 2021. Global Tuberculosis Report 2021. World Health Organization: Geneva.

Appendix A. CGE model calibration

Our macroeconomic Computable General Equilibrium (CGE) model framework consists of a set of behavioural and accounting equations. The core static CGE model framework is structured according to the IFPRI standard model framework as documented in Löfgren, Lee Harris & Robinson (2002). In this appendix, we outline the revisions and additions which we have made to the standard static model framework, and describe the dynamic equations which we have added to turn the framework into a dynamically-recursive model framework. The revised and added equations are outlined in section A.1, while the calibration of the full CGE model framework is described in section A.2. Finally, the macro-closure of the CGE model framework is described in section A.3.

The CGE model is calibrated to the 2014 India Social Accounting Matrix (SAM) from the GTAP10 database (Aguiar et al. 2019). The original SAM contained 65 activity and commodity sectors, but was aggregated to 31 activity and commodity sectors for computational purposes. Throughout, we define subscript $a = 1, \dots, 31$ to refer to our set of 31 production activities, and subscript $c = 1, \dots, 31$ to refer to our set of 31 retail commodities.

Furthermore, while the original GTAP SAM data set only contained one representative household, we disaggregated this single household account according to income quintiles, based on data from the 68th National Sample Survey 2011-12 for India (NSSO 2018). Throughout, we define subscript $h = 1, \dots, 5$ to refer to representative households defined by Socioeconomic Status (SES) quintiles running from quintile 1 (lowest SES) to quintile 5 (highest SES).

We note that the calibration of our macroeconomic CGE model involves stratification of households according to household income quintiles. The use of income quintiles for household stratification purposes is similar to the approach taken for household stratification within our demographic model framework, but it differs somewhat from the household stratification within our epidemiological model, which were based on household wealth quintiles (due to lack of data availability). While the correlation between household wealth and household income is <1, we, however, believe that the error we make is of a second order nature.

A.1. CGE model equations and parameters

Within the macroeconomic CGE model, the IFRPI standard model allows for LES and CES demand systems. However, since nutrition is considered to be a key pathway, in our model framework, linking household demand with TB clinical health outcomes, we introduced a set of more flexible household-specific 'Almost Ideal Demand Systems' (AIDS) in equation A.1:

(A.1) $$\frac{PQ_{c,t}QH_{c,h,t}}{EH_{h,t}} = \alpha_h^{AIDS} + \beta_h^{AIDS} log(\frac{EH_{h,t}/P_t}{EH_{h,0}/P_0}) + \gamma_h^{AIDS} log(\frac{PQ_{c,t}/P_t}{PQ_{c,0}/P_0})$$

where the budget share of expenditures on commodity c ($PQ_{c,t}QH_{c,h,t}$), as share of total household expenditures ($EH_{h,t}$), for household h, is a function of the log of (normalized) real expenditures ($log(\frac{EH_{h,t}/P_t}{EH_{h,0}/P_0})$) and the log of the (normalized) relative price of commodity c ($log(\frac{PQ_{c,t}/P_t}{PQ_{c,0}/P_0})$).

The calibration of AIDS demand specifications of private demand for each of our five income quintile households, was based on a comprehensive, but fairly aggregate, set of USDA-ERS estimates of uncompensated own-price, cross-price, and income elasticities, for India, which were estimated on the basis of 2005 International Comparison Project data, and covering all food and non-food consumption items (Meade, Regmi, Seale, Muhammad 2014). These data were complemented by more recent, and more detailed, uncompensated own-price, cross-price, and income elasticities for food items for Tamil Nadu (Felix & Kumar 2020), which were estimated on the basis of data from the 68th National Sample Survey 2011-12 for India (NSSO 2018). While the latter study focussed on Tamil Nadu data, the lack of detailed food commodity elasticity estimates in the USDA-ERS study led us to adopt Felix & Kumar's food elasticity estimates under the maintained assumption that they apply consistently across all Indian states.

In order to properly parameterize the AIDS demand systems for each of our five income quintile households, we constructed a sector mapping and employed standard uncompensated price and income elasticity formulas (Green & Alston 1990, 1991) to produce prior unbalanced parameter values for our household-specific AIDS-demand systems, based on household-specific relative demand shares from our 2014 India SAM. As usual, the raw AIDS demand cross-price and income parameter matrices and vectors did not satisfy basic regularity conditions (symmetry and adding-up). A Bayesian minimum cross-entropy analysis, specified with a full set of regularity constraints, was therefore used to derive a full set of posterior parameter values including a 31x31 cross-price elasticity matrix (satisfying all regularity conditions) and a 31x1 income elasticity vector (satisfying adding-up constraints) for each of our five income quintile households.

In order to model the healthcare cost feedback pathway from the calculation of TB morbidity impacts within the epidemiological and demographic models (see appendix D.1) to the macroeconomic CGE model, we made the simplifying assumption that all healthcare costs will be borne by the Indian government. Furthermore, as also noted in appendix D.1, we take a conservative approach to the measurement of TB healthcare costs and focus on measuring the treatment costs of MDR-TB and DS-TB treatment regimens.

With the above in mind, and noting our calculated average treatment cost of 10,310.6 INR/case in 2014 fixed prices ($TBtreat\_unitcost$) and that total healthcare treatment costs in fixed 2014 prices

($TBtreat\_cost$) are calculated in equation D.5 (see appendix D.2), we are able to complement our standard equation A.2, which specifies real government consumption of healthcare commodity 'c29' ($QG^{scen}_{'c29',t}$) as being equal to fixed core real government consumption of the healthcare commodity 'c29' ($\overline{qg}^{count}_{'c29',t}$) (for use in counterfactual simulations), with equation A.2' which adds endogenous fixed price healthcare costs of TB morbidity outcomes to the equation (for use in policy simulations):

(A.2) $$QG^{scen}_{'c29',t} = \overline{qg}^{count}_{'c29',t}$$

(A.2') $$QG^{scen}_{'c29',t} = \overline{qg}^{count}_{'c29',t} + (TBtreat\_cost^{scen} - TBtreat\_cost^{count})$$

Specifically, the additional term in equation A.2', involving the difference between the policy scenario TB treatment costs ($TBtreat\_cost^{scen}$) and the counterfactual TB treatment costs ($TBtreat\_cost^{count}$) (see equation D.5 in appendix D.1), accounts for TB morbidity impacts on government consumption of the healthcare services commodity 'c29'.

A.2. CGE model calibration

Production functions are specified as Constant Elasticity of Substitution (CES) functions of aggregate intermediate input demands (disaggregate commodity input demands are determined by Leontief specifications) and aggregate factor input demands (disaggregate factor input demands are also determined by CES specifications) with standard elasticity values for the top-level production specifications (0.8) and the bottom-level factor input demand specifications (0.6); Trade between domestic and foreign agents is specified as a function of relative prices (determined by the real exchange rate), based on Armington CES specifications on the import side and Constant Elasticity of Transformation (CET) specifications on the export side. Standard trade elasticity values were applied on the import side (1.2) and on the export side (1.5).

Based on the above-mentioned CGE model parameters, we calibrated our core static CGE model to the 2014 India SAM from the GTAP10 database (Aguiar et al. 2019).

A.3. CGE model dynamically-recursive factor updating equations

The core static macroeconomic CGE model framework, which has been outlined in the sub-sections A.1-A.2, is turned into a dynamically-recursive CGE model framework through the addition of a set of household-specific labour and capital factor supply equations, which tracks the household-specific accumulation and supply of labour and capital factors over time. The core static IFPRI standard model framework does not include capital factor ownership and supply equations. While the lack of capital factor ownership and supply equations is perfectly fine for a single-household model, it is not appropriate for a model with multiple households, where the tracking of household-specific factor

ownership and supply is critical for capturing distributional impacts of policy scenarios. We therefore added a full set of labour and capital factor supply equations to our model framework.

The household-specific labour factor ownership and supply equations are linked to the household-specific population projections from our demographic model, which are calibrated to the Indian population projections from the World Population Prospects: 2019 Revision (UN 2019) (see appendix B for details); while the household- and enterprise-specific capital factor ownership equations are calibrated to Indian capital stock and depreciation data from the Penn World Tables 10.0 (Feenstra Inklaar & Timmer 2015).

Specifically, we model the dynamic updating of household- and enterprise-specific factor supplies through the specification of the following household- and enterprise-specific labour and capital factor supply equations A.3-A.4:

(A.3) $$QFH_{h,flab,t} = skl\_share_{h,flab} part\_rate_{h,flab} \sum_{gen,age} POP_{h,gen,age,t}$$

(A.4a) $$QFH_{h,fcap,t+1} = (1 - dep\_rate) QFH_{h,fcap,t} + qfh\_share_{h,fcap} \frac{\sum_c PQ_{c,t} QINV_{c,t}}{IPI_t}$$

(A.4b) $$QFH_{,ent',fcap,t+1} = (1 - dep\_rate) QFH_{,ent',fcap,t} + qfh\_share_{,ent',fcap} \frac{\sum_c PQ_{c,t} QINV_{c,t}}{IPI_t}$$

where the labour factor ownership equation A.2 calculates household-specific factor ownerships of unskilled and skilled labour factor types ($QFH_{h,flab,t}$) as functions of household-specific skill-shares of skilled and unskilled labour ($skl\_share_{h,flab}$), household- and labour factor-specific labour force participation rates ($part\_rate_{h,flab}$), and age-, gender- and household-specific population numbers projected by the demographic model ($POP_{h,gen,age5,t}$); and the capital factor ownership equations A.4a and A.4b calculate household- and enterprise-specific capital factor ownerships ($QFH_{h,fcap,t}$ and $QFH_{,ent',fcap,t}$) based on a constant capital depreciation rate (dep_rate), and constant household- and enterprise-specific capital factor ownership-shares ($qfh\_share_{i,fcap}$ and $qfh\_share_{,ent',fcap}$).

When simulating policy scenarios, where TB morbidity and mortality impacts (see appendix D.1) differ from baseline outcomes, we apply the following adapted labour factor ownership and supply equation A.2':

(a.3') $$QFH_{h,flab,t} = skl\_share_{h,flab} part\_rate_{h,flab}$$
$$+ \sum_{age} ((\sum_{gen} POP_{h,gen,age,t})$$
$$+ (TBcases\_mort_{h,age,t}^{annual,scen} - TBcases\_mort_{h,age,t}^{annual,count})$$

$$+TBisolat\_durat * (TBcases\_incid_{h,age,t}^{annual,scen} - TBcases\_incid_{h,age,t}^{annual,count}))$$

where the two added terms accounts for net changes to labour supplies among working age adults due to policy-induced changes in respectively TB mortality and TB morbidity. Specifically, the first additional term, involving the difference between the policy scenario excess TB case fatalities ($TBcases\_mort_{h,age,t}^{annual,scen}$) and the counterfactual excess TB case fatalities ($TBcases\_mort_{h,age,t}^{annual,count}$) (see equation D.1 in appendix D.1), accounts for TB mortality impacts on labour supplies, while the second additional term, involving the difference between policy scenario excess TB incident cases ($TBcases\_incid_{h,age,t}^{annual,scen}$) and counterfactual excess TB incident cases ($TBcases\_incid_{h,age,t}^{annual,count}$) (see equation D.2 in appendix D.1), accounts for TB morbidity impacts on labour supplies, where morbidity impacts involves average isolation requirements of $0.158\ years/case$ ($TBisolat\_durat = 0.158\ years/case$) (see details of the calculation of $TBisolat\_durat$ in appendix D.2).

The household-specific factor supplies, which are determined by equations A.3-A.4, are subsequently cumulated to form total factor supplies of labour ($QFS_{flab,t}$) and capital ($QFS_{fcap,t}$) in equations A.5a-A.5b:

(A.5a) $QFS_{flab,t} = \sum_{i=1}^{5} QFH_{i,flab,t}$

(A.5b) $QFS_{fcap,t} = \sum_{i=1}^{5} QFH_{i,fcap,t} + QFH_{,ent,fcap,t}$

and these factor supplies are then allocated across production sectors, as a function of relative returns, and subsequently determines the growth path of the Indian economy (see Löfgren, Lee Harris & Robinson (2002) for details on the efficient allocation of production factors across production activities according to relative returns specified by sector-specific first order conditions for profit-maximization).

We note that the link between the demographic and epidemiological model predictions of TB morbidity and mortality clinical health outcomes (se appendix D.1), on the one hand, and labour factor supplies, on the other, (partially) determine the dynamic growth path of the macroeconomic CGE model. Hence, this provides a key pathway for the full integration of our three epidemiological, demographic, and macroeconomic models into one fully integrated model framework.

A.4. Establishing 2021 as base year, and constructing counterfactual growth path

Based on the above-mentioned calibration of our core static CGE model (section A.2) and the specification of our dynamic factor ownership equations (section A.3), as well the macro closure rules discussed in the next section, we were finally able to run the dynamically-recursive model forward to 2020 to match 2015-2020 nominal and real GDP aggregates from the World

Development Indicators database (WB 2021) and to establish 2021 as our base year. Again referring to the closure rules discussed in the next section, we then simulated a counterfactual 2021-40 growth path to match the average 2011-2018 Indian growth rates of real GDP (7.0% p.a.) and nominal GDP (11.6% p.a.) with an underlying implied GDP deflator inflation rate of 4.6% p.a., thereby completing the calibration of the dynamically-recursively model, including establishing 2021 as our base year, and the construction of our 2021-2040 counterfactual growth path.

A.5. CGE model macro closure

Our macroeconomic model closure specifies the GDP deflator as price numeraire. Hence, the GDP deflator was kept fixed at the counterfactual GDP price level growth path, based on a fixed annual inflation rate of 4.6% (see section A.2), in both our counterfactual simulation and our trade policy simulation.

Our counterfactual growth path (see section A.4 above) was simulated with a standard neoclassical model closure involving (1) price clearing of all goods and factor markets, (2) real exchange rate clearing of the Balance of Payments, and (3) savings-driven investment to clear the capital account. In addition, the counterfactual growth path was simulated with a balanced macro-closure which ensures that the government consumption share of domestic absorption is fixed along the counterfactual growth path.

Our import tariff elimination trade policy scenario was, similarly, simulated with a standard neoclassical macro closure, involving (1) price clearing of all goods and factor markets, (2) real exchange rate clearing of the Balance of Payments, and (3) savings-driven investment to clear the capital account. However, in contrast to the counterfactual simulation, our policy simulation kept the core government consumption 'fixed' at the counterfactual growth path (in terms of its own price deflator). Hence, our policy simulation only allowed government consumption to vary in response to changes in TB-related health costs. Combined with a standard revenue-neutral government budget closure, where non-distortionary household taxes adjust endogenously to compensate for tariff-related shocks and keep government revenues fixed at the counterfactual growth path, this means that, except for minor relative price changes, government savings (which adjusts to clear the government account) will vary almost one-to-one with changes in TB-related health costs.

References

Aguiar A, Chepeliev M, Corong E, McDougall R, Van Der Mensbrugghe D. 2019. The GTAP Data Base: Version 10. Journal of Global Economic Analysis 4(1):1-27.

Feenstra RC, Inklaar R, Timmer MP. 2015. The Next Generation of the Penn World Table. American Economic Review 105(10):3150-3182. available for download at www.ggdc.net/pwt

Felix KT, Kumar PN. 2020. An Economic Analysis of Food Demand in Rural and Urban Households of Tamil Nadu. Indian Journal of Economics & Development 16(2):173-79.

Löfgren H, Lee Harris R, Robinson S. 2002. A Standard Computable General Equilibrium (CGE) Model in GAMS. Microcomputers and Policy Research 5. International Food Policy Research Institute, Washington DC.

Meade B, Regmi A, Seale Jr JL, Muhammad A. 2014. New International Evidence on Food Consumption Patterns. A Focus on Cross-Price Effects Based on 2005 International Comparison Program Data. USDA-ERS Technical Bulletin Number 1937. Washington DC: United States Department of Agriculture. URL: https://www.ers.usda.gov/webdocs/publications/47609/42021_tb1937.pdf?v=41680

Muhammad A, Seale Jr JL, Meade B, Regmi A. 2011. International Evidence on Food Consumption Patterns. An Update Using 2005 International Comparison Program Data. USDA-ERS Technical Bulletin Number 1929. Washington DC: United States Department of Agriculture. URL: https://www.ers.usda.gov/webdocs/publications/47579/7637_tb1929.pdf?v=41680

NSSO. 2018. National Sample Survey (NSS) 68th round 2011-12. Electronic data. Ministry of Statistics and Programme Implementation, Government of India. URL: http://microdata.gov.in/nada43/index.php/catalog.

WB. 2021. Electronic data. World Development Indicators (WDI) database. Washington DC: World Bank. URL: https://databank.worldbank.org/source/world-development-indicators.

UN. 2019. Electronic data. World Population Prospects, the 2019 Revision. United Nations. URL: http://esa.un.org/unpd/wpp/index.htm.

Appendix B. Demographic model calibration

The demographic module is designed to mirror the population stratification of the macroeconomic CGE model presented in appendix A, and the epidemiological model presented in appendix C. A set of aggregate 2014-40 Indian population projections, along with age-, gender-, and time-specific parameters for sex ratios at birth (sex_ratio_{gen}), fertility rates ($asfr_{age,t}$), mortality rates ($\mu_{gen,age,t}$), was derived from the 2019 revision of the United Nations World Population Prospects (UN 2019), and used as the basis for calibration of our demographic model. However, in order to match the disaggregated income quintile disaggregation of representative households within the CGE model (see appendix A), we further disaggregated the Indian population across household income quintiles based on weighted age- and gender-specific population shares across income quintiles from the 2011-12 Indian National Sample Survey NSS68 (NSSO 2013). By employing the NSS68 household survey dataset to derive population composition shares, we ensured consistency between our macroeconomic and demographic household disaggregation across income quintiles by consistently employing data from the same household survey to disaggregate household-specific population composition within the demographic model, and household-specific income and expenditure composition within the macroeconomic SAM database (see appendix A). However, the focus on income quintiles, as measures of SES quintiles, differs somewhat from our epidemiological model, where households are stratified according to household wealth quintiles (due to data issues). While the correlation between household wealth and household income is <1, we believe, as previously noted, that the error we make is of a second order nature.

The core demographic model parameters, including sex ratios at birth (sex_ratio_{gen}), fertility rates ($asfr_{age,t}$), and mortality rates ($\mu_{gen,age,t}$), were assumed to be constant across household types, while two additional sets of household-specific parameters including net immigration rates ($\alpha^{imigr}_{h,gen,age,t}$), and annual transition probabilities between 5 year age groups ($p^{trans}_{h,gen,age,t}$) were calibrated to target our set of 2014-40 household-specific population projections. Hence, the data set, described above, was employed to calibrate our demographic model, encompassing a compact set of equations for calculating Births ($Births_{h,gen,t}$), Deaths ($Deaths_{h,gen,age,t}$), International Net Migration ($Migr_{h,gen,age,t}$), and Population ($POP_{h,gen,age,t}$) levels for our five household type (H), 26 annual time periods (T), 15 quinquennial age categories (AGE) and two gender types (GEN):

(B.1) $$Births_{h,gen,t} = sex\_ratio_{gen} \sum_{(gen|gen='female',age\in[15;49]} asfr_{age,t} POP_{h,gen,age,t-1}\,, \quad h \in H, gen \in GEN, t \in T$$

(B.2) $$Deaths_{h,gen,age,t} = \mu_{h,gen,age,t} POP_{h,gen,age,t-1}\,, \quad h \in H, gen \in GEN, age \in AGE, t \in T$$

(B.3) $$Migr_{h,gen,age,t} = \alpha^{imigr}_{h,gen,age,t}\left(1-\mu_{h,gen,age,t}\right)POP_{h,gen,age,t-1}\,,$$

$$h \in H, gen \in GEN, age \in AGE, t \in T$$

(B.4a) $$POP_{h,gen,age,t|age=[0;4]} = \left(1-p^{trans}_{h,gen,age,t}\right)\left(1-\alpha^{imigr}_{h,gen,age,t}\right)\left(1-\mu_{h,gen,age,t}\right)POP_{h,gen,age,t-1}$$

$$+Births_{h,gen,t}, \qquad h \in H, gen \in GEN, t \in T$$

(B.4b) $$POP_{h,gen,age,t|age>[0;4]\,\&\,age<[80;]} = p^{trans}_{h,gen,age-1,t}\left(1-\alpha^{imigr}_{h,gen,age-1,t}\right)\left(1-\mu_{h,gen,age-1,t}\right)POP_{h,gen,age-1,t-1}$$

$$+\left(1-p^{trans}_{h,gen,age,t}\right)\left(1-\alpha^{imigr}_{h,gen,age,t}\right)\left(1-\mu_{h,gen,age,t}\right)POP_{h,gen,age,t-1},$$

$$h \in H, gen \in GEN, age \in AGE\backslash\{[0;4],[80;]\}, t \in T$$

(B.4c) $$POP_{h,gen,age,t|age=[80;]} = p^{trans}_{h,gen,age-1,t}\left(1-\alpha^{imigr}_{h,gen,age-1,t}\right)\left(1-\mu_{h,gen,age-1,t}\right)POP_{h,gen,age-1,t-1}$$

$$+\left(1-\alpha^{imigr}_{h,gen,age,t}\right)\left(1-\mu_{h,gen,age,t}\right)POP_{h,gen,age,t-1}\,, h \in H, gen \in GEN, t \in T$$

Equations B.1-B.4 were all included within the integrated macroeconomic-demographic-epidemiological model framework, and the four sets of variables including Births ($Births_{h,gen,t}$), Deaths ($Deaths_{h,gen,age,t}$), International Net Migration ($Migr_{h,gen,age,t}$), and Population ($POP_{h,gen,age,t}$) are all endogenous variables within the model framework.

In terms of how the demographic models link with the epidemiological models, the endogenous household-specific all-cause mortality rates ($\mu_{h,gen,age,t}$) are aggregates of exogenous baseline all-cause mortality rates ($\mu^{0}_{h,gen,age,t}$) and net changes in TB excess mortality rates ($\mu^{TB}_{h,gen,age,t} - \mu^{TB,0}_{h,gen,age,t}$):

(B.5) $$\mu_{h,gen,age,t} = \mu^{0}_{h,gen,age,t} + \left(\mu^{TB}_{h,gen,age,t} - \mu^{TB,0}_{h,gen,age,t}\right)$$

where the impacts of changes in endogenous TB excess mortality rates ($\mu^{TB}_{h,gen,age,t}$) on endogenous all-cause mortality rates ($\mu_{h,gen,age,t}$) are calculated net of baseline all-cause mortality rates ($\mu^{0}_{h,gen,age,t}$) and baseline TB excess mortality rates ($\mu^{TB,0}_{h,gen,age,t}$) in order to avoid double-counting, and where the endogenous excess TB mortality rates are determined within the epidemiological model (see appendix C).

In terms of how the demographic models link with the macroeconomic CGE model, the household-specific population variables ($POP_{h,gen,age,t}$) are used to calculate household-specific labour factor ownerships and supplies, which feed into the macroeconomic CGE model (see equations A.3 and A.3' in appendix A).

References

NSSO. 2018. National Sample Survey (NSS) 68th round 2011-12. Electronic data. Ministry of Statistics and Programme Implementation, Government of India. URL: http://microdata.gov.in/nada43/index.php/catalog.

UN. 2019. Electronic data. World Population Prospects, the 2019 Revision. United Nations. URL: http://esa.un.org/unpd/wpp/index.htm.

Appendix C. Epidemiological model specification and model fitting

Our epidemiological model for household-specific TB transmission, within and between our five household types, is specified as a set of difference equations with bi-weekly time periods. We made this choice as a trade-off between our macroeconomic and demographic models being specified with discrete annual time steps, and the need for our discrete time epidemiological model framework to have sufficiently high frequency time steps to mirror the standard continuous time epidemiological model specifications, but also with sufficiently few time steps, within a given year, to make model solutions over >25 annual time periods tractable.

It is important to understand that the nature of the epidemiological model (similar to the demographic model) is a simulation model which traces out a non-equilibrium non-steady state growth path, and that this is fundamentally different from the macroeconomic CGE model, which in its essence is producing annual equilibrium solutions of the Indian economy. In essence, the epidemiological and demographic models are (partly) the growth engines which drive the CGE model forward from one annual equilibrium solution to the next (another growth engine is the endogenous capital accumulation which occurs as part of the dynamically-recursive CGE model specification).

The difference in time steps between our epidemiological vs. demographic and macroeconomic CGE models have implications for how we specify the interaction between the three models within our integrated model framework. Specifically, the annual values of TB incidence ($Incid_{h,t}^{annual}$) and TB mortality ($Mort_{h,t}^{annual}$), which are used to calculate inputs for the CGE and demographic models, are calculated as the sum of the 26 bi-weekly incidence and mortality numbers ($Incid_{h,t\_2w}^{bi-week}$ and $Mort_{h,t\_2w}^{bi-week}$) across the given year (see equations C.10 and C.16), while the annual TB prevalence ($Prev_{h,t}^{annual}$) is calculated as the average of the 26 bi-weekly prevalence numbers ($Prev_{h,t\_2w}^{bi-week}$) across the given year (see equation C.12).

IT should be noted that 26 bi-weekly model solutions for >25 annual time periods results in >650 bi-weekly time steps, each involving separate epidemiological model solutions for each of our five household types. Due to the potential for numerical problems caused by the dimensionality of the illness at hand, we chose to specify a focussed model of TB transmission, which we believe exhibits the core features of Indian TB transmission within and between Indian households stratified by SES status.

We note again (also noted in appendices A and B) that the calibration of our epidemiological model involves stratification of households according to household wealth quintiles, and that this differs somewhat from our demographic and macroeconomic CGE models, discussed in appendices A and

B, where households are stratified according to household income quintiles based on the NSS68 household survey data (NSSO 2018). The alternative stratification, within the epidemiological model, according to household wealth quintiles, is due to data access issues; it was only possible to obtain epidemiological model parametric data stratified according to wealth quintiles, but not according to income quintiles. While the correlation between household wealth and household income is <1, we believe, as previously noted, that the error we make is of a second order nature.

C.1. Epidemiological model: State/compartment definitions

The states of our epidemiological model, used to define the various compartments of our epidemiological model, include a 'Susceptible' state (S), a 'Latently Infected' state (L), and 'Infectious' state (I), an 'Infectious with previous treatment' state (Ir), an 'On Treatment' state (T), and an 'On treatment with previous treatment" state (Tr) (see Table C.1.)

Table C.1. Epidemiological model – state/compartment definitions

S	Susceptible (never exposed to M.tb.)
L	Latently Infected (infected with M.tb, not infectious)
I	Infectious (TB disease)
Ir	Infectious with previous treatment (TB disease, previous treatment for current episode)
T	Treatment (on treatment for TB disease)
Tr	Treatment with previous treatment (on treatment for TB disease, previous treatment for current episode)
D*	Deaths (excess mortality among persons with TB disease)

Notes: *Since we're modelling living population shares, the compartment for Deaths (D) is not strictly speaking a compartment within our model. However, we keep track of the population share which exits to the D compartment every time period.

C.2. Epidemiological model: Parameter definitions

Our epidemiological model is specified in terms of a set of difference-equations, which are outlined in the following section. The parameters of our epidemiological framework are listed in Table C.2., and the non-calibrated parameter values are listed in Table C.3.

Table C.2. Epidemiological model – parameter definitions

u	background mortality rate
m	excess mortality rate due to active TB
a_h*	proportion progressing directly to disease on infection
c_h*	risk of developing active TB from latent infection
Q	relative risk of active TB following re-infection
E	risk of self-cure of active TB
D	risk of diagnosis of active TB*
T	proportion successfully treated*
V	1/average duration of treatment
β_{in}	transmission risk between those in same SES quintile
β_{rel}	relative transmission risk between SES quintiles
N	Total population

Notes: *Allowed to vary by household SES quintile

C.3. Epidemiological model: Equation definitions

Our epidemiological model framework consists of a set of behavioural difference equations and a set of accounting equations. The difference equations are outline in section C.3.1, while the accounting equations are outlined in section C.3.2. Throughout, we define subscript h = 1, … ,5 to refer to representative households defined by Socioeconomic Status (SES) quintiles running from quintile 1 (lowest SES) to quintile 5 (highest SES). Furthermore, we distinguish between annual time steps $t \in [2014; 2040]$ and bi-weekly time steps within a given year $t\_2w \in [1; 26]$.

C.3.1. Epidemiological model: Difference equation definitions

Our approach to epidemiological modelling is focussed on modelling the share of the Indian population which, at any given point in time, is in a given TB state, and to, subsequently, combine our epidemiological model outcomes with demographic model predictions, in order to derive population level TB health indicators. As a consequence, our epidemiological TB model is a closed system, where the population (N) is normalized to 1, implying that sum of our five household-specific populations sums to 1 at any given time ($\sum_{h=1}^{5} N_{h,t\_2w} = 1$).

The difference equation for modelling changes to the population share in the 'Susceptible' state (S) is described in equation C.1, and it is a function of the Force of Infection ($\lambda_{h,t}$), defined below in equation X and describing the transition rate from the S-state to the L-state (see Figure 1). The equation also includes two accounting terms which are related to background and excess TB mortality, and which ensures that the system remains closed:

(C.1) $$S_{h,t\_2w+1} = S_{h,t\_2w} - \lambda_{h,t\_2w} S_{h,t\_2w} + u_{h,t\_2w}\left(N_{h,t\_2w} - S_{h,t\_2w}\right) + m_{h,t\_2w}\left(I_{h,t\_2w} + Ir_{h,t\_2w}\right)$$

The difference equation for modelling changes to the population share in the 'Latently infected' state (L) is described in equation C.2, and it reflects the rates of transition from the L-state to the I-state (negative term) and transitions in the opposite direction from the S-, I-, Ir-, T- and Tr-states to the L-state (positive terms) (see Figure 1):

(C.2) $$L_{h,t_{2w}+1} = L_{h,t_{2w}} + \left(1 - a_{h,t_{2w}}\right)\lambda_{h,t_{2w}} S_{h,t_{2w}} - \left(q a_{h,t_{2w}} \lambda_{h,t_{2w}} + c_{h,t_{2w}} + u_h\right) L_{h,t_{2w}}$$
$$+\varepsilon\left(I_{h,t\_2w} + Ir_{h,t\_2w}\right) + v\tau_{h,t\_2w}\left(T_{h,t\_2w} + Tr_{h,t\_2w}\right)$$

The difference equation for modelling changes to the population share in the 'Infectious' state (I) is described in equation C.3, and it reflects the rates of transition from the I-state to the L-, T-, and D-states (negative terms) and transitions in the opposite direction from the S- and L-states to the I-state (positive terms) (see Figure 1):

(C.3) $$I_{h,t_{2w}+1} = I_{h,t_{2w}} + a_{h,t_{2w}} \lambda_{h,t_{2w}} S_{h,t_{2w}} + \left(q a_{h,t_{2w}} \lambda_{h,t_{2w}} + c_{h,t_{2w}}\right) L_{h,t_{2w}}$$

$$-\left(d_{h,t\_2w} + \varepsilon + u_h + m_{h,t\_2w}\right) I_{h,t\_2w}$$

The difference equation for modelling changes to the population share in the 'Infectious with previous treatment' state (Ir) is described in equation C.4. This compartment was included to avoid double-counting notifications, and the equation reflects the rates of transition from the Ir-state to the L-, Tr-, and D-states (negative terms) and transitions in the opposite direction from the T- and Tr-states to the Ir-state (positive terms) (see Figure 1):

(C.4) $$Ir_{h,t_{2w}+1} = Ir_{h,t_{2w}} + v\left(1 - \tau_{h,t_{2w}}\right)\left(T_{h,t_{2w}} + Tr_{h,t_{2w}}\right)$$

$$-\left(d_{h,t\_2w} + \varepsilon + u_h + m_{h,t\_2w}\right) Ir_{h,t\_2w}$$

The difference equation for modelling changes to the population share in the 'Treatment' state (T) is described in equation C.5, which reflects the rates of transition from the T-state to the L-, Ir-, and D-states (negative terms) and transitions in the opposite direction from the I-state to the T-state (positive term) (see Figure 1):

(C.5) $$T_{h,t\_2w+1} = T_{h,t\_2w} + d_{h,t\_2w} I_{h,t\_2w} - (v + u_h) T_{h,t\_2w}$$

The difference equation for modelling changes to the population share in the 'Retreatment' state (Tr) is described in equation C.6. Similar to the Ir-compartment, described in equation C.4, the Tr-compartment was included to avoid double-counting notifications, and the equation reflects the rates of transition from the Tr-state to the L-, Ir-, and D-states (negative terms) and transitions in the opposite direction from the Ir-state to the Tr-state (positive terms) (see Figure 1):

(6) $$Tr_{h,t\_2w+1} = Tr_{h,t\_2w} + d_{h,t\_2w} Ir_{h,t\_2w} - (v + u_h) Tr_{h,t\_2w}$$

C.3.2. Epidemiological model: Accounting equation definitions

As mentioned above in relation to equation C.1, our epidemiological TB model is a closed system, where the population (N) is normalized to 1, implying that sum of our five household-specific populations sums to 1 at any given time ($\sum_{h=1}^{5} N_{h,t} = 1$). Due to the definition of representative households, based on income quintiles among population-weighted households, the household-specific population shares (N_{h,t_2w}) are fixed at 20% each, i.e. $\overline{N_{h,t\_2w}} = 0.2$. With that in mind, the population shares of the household-specific compartments have to add up $\overline{N_{\iota,t\_2w}}$ as specified in the following equation C.7:

(C.7) $$\overline{N_{h,t\_2w}} = S_{h,t\_2w} + L_{h,t\_2w} + I_{h,t\_2w} + Ir_{h,t\_2w} + T_{h,t\_2w} + Tr_{h,t\_2w}$$

The household-specific Force of Infection ($\lambda_{i,t}$) is an endogenous parameter which measures the rate at which individuals within a given household *i* transition from the S-state to the L-state. Specifically, the Force of Infection for household *h* is calculated as the 'contact rate'-weighted average of infectious population-shares among all five households, including weighting of the within-household infectious population share by a within-household contact rate parameter (β_{in}), and weighting of other household infectious population shares by a between-household contact rate parameter (β_{rel}) as outlined in the following equation C.8:

(C.8) $$\lambda_{h,t\_2w} = \beta_{in}\left(\left(I_{h,t\_2w} + Ir_{h,t\_2w}\right) + \beta_{rel}\sum_{l\neq h}\left(I_{l,t\_2w} + Ir_{l,t\_2w}\right)\right)$$

Instantaneous TB incidence, i.e. the instantaneous intensity of TB infections which occur during the two-week time intervals which constitutes our discrete time periods, is defined as a function of the rates of transition from the S- and L-states to the I-state as outlined in the following equation C.9:

(C.9) $$Incid_{h,t\_2w}^{instantaneous} = a_{h,t\_2w}\lambda_{h,t\_2w}S_{h,t\_2w} + \left(qa_{h,t\_2w}\lambda_{h,t\_2w} + c_{h,t\_2w}\right)L_{h,t\_2w}$$

Annual TB incidence, i.e. the number of TB infections which occur during a discrete annual time period *t_ann*, is defined as the sum of instantaneous TB incidence across the 26 bi-weekly time intervals within that given annual period *t_ann*, as outlined in the following equation C.10:

(C.10) $$Incid_{h,t}^{annual} = \sum_{t\_2w=1}^{26} Incid_{h,t\_2w}^{instantaneous}$$

Instantaneous TB prevalence, i.e. the instantaneous prevalence of TB infections which exists during a specific two-week time interval *t*, is defined as the sum of the population shares of the I and Ir-compartments during that specific two-week time interval *t*, as outlined in the following equation C.11:

(C.11) $$Prev_{h,t\_2w}^{instantaneous} = I_{h,t\_2w} + Ir_{h,t\_2w}$$

Annual TB prevalence, i.e. the prevalence of TB infections during a discrete annual time period *t_ann*, is defined as the average instantaneous prevalence rate across the 26 bi-weekly time intervals within that given annual period *t_ann*, as outlined in the following equation C.12:

(C.12) $$Prev_{h,t}^{annual} = \frac{1}{26}\sum_{t\_2w=1}^{26} Prev_{h,t\_2w}^{instantaneous}$$

Instantaneous TB notifications, i.e. the instantaneous number of TB infections which are diagnosed and started on treatment during a specific two-week time interval *t*, is defined as a function of the rate of transition from the I-state to the T-state as outlined in the following equation C.13:

(C.13) $$Notif_{h,t\_2w} = d_{h,t\_2w}I_{h,t\_2w}$$

Annual TB notifications, i.e. the notifications of TB infections which occur during a discrete annual time period *t*, is defined as the sum of instantaneous TB notifications across the 26 bi-weekly time intervals within that given annual time period *t*, as outlined in the following equation C.14:

(C.14) $Notif_{h,t}^{annual} = \sum_{t\_2w=1}^{26} Notif_{h,t\_2w}^{instantaneous}$

Instantaneous TB mortality, i.e. the number of instantaneous TB case fatalities which occur during a specific two-week time interval *t*, is defined as the sum of the rates of transition from the I- and Ir-states to the D-state as outlined in the following equation C.15:

(C.15) $Mort_{h,t\_2w} = m_{h,t\_2w}(I_{h,t\_2w} + Ir_{h,t\_2w})$

Annual TB mortality, i.e. the number of TB case fatalities which occur during a discrete annual time period *t*, is defined as the sum of instantaneous TB mortality measures across the 26 bi-weekly time intervals within that given annual time period *t*, as outlined in the following equation C.16:

(C.16) $Mort_{h,t}^{annual} = \sum_{t\_2w=1}^{26} Mort_{h,t\_2w}^{instantaneous}$

C.3.3. Epidemiological model: Risk Factor equations

We model the impact of risk factors on endogenous model parameters via a set of equations, which are outlined in the sub-section. Specifically, we assume that TB risk is mainly mediated by six risk factors including tobacco smoking, indoor air pollution, low BMI, daily alcohol use, HIV infection, and diabetes (Oxlade and Murray 2012, Andrews et al 2015). Moreover, following the latter study by Andrews et al., we model the prevalence of the six main risk factors, and the relative risks for each of these factors, to as determinants of the relative progression from the susceptible and latent states to the TB disease state (parameters *a* and *c* are defined in Table C.2); and we also follow Andrews et al. in modelling relative risks of the six main risk factors across Indian Socioeconomic Status (SES) quintiles.

In terms of the two key disease progression parameters *a* and *c*, we adopt base line values from Sumner and White (2020) (see Table C.3). Our approach assumes that these values represent those in the absence of risk factors. Following Andrews et al. (2015), we represent the relationship between the risk of TB (if exposed) on the one hand, and the six risk factors on the other, using a poisson model (see p. ii in the appendix to Andrews et al.). The log risk of TB disease in SES strata *i* is then given by:

$$\ln(z_h) = \alpha + \sum_j g_j p_{hj}$$

where g_j is the **log** of the relative risk for risk factor *j* to lead to development of TB disease, and where p_{hj} is the prevalence of risk factor *j* in household SES strata *h*.

The risk rate ratio for TB disease in household SES strata *h* (RR_h), is defined as the ratio of risk of TB disease in household SES strata *h* (z_h) relative to the risk of TB disease in a population with no risk factors (z_0):

$$RR_h = \exp\left(\ln(z_h) - \ln(z_0)\right)$$

Because $p_{hj} = 0$ in the absence of risk factors, $\ln(z_0) = \alpha$ and therefore:

(C.17) $RR_h = \exp\left(\sum_j g_j p_{hj}\right)$

The calibration of the risk rate ratios, based on equation C.17 and stratified across household SES strata *h*, is presented in Table C.4 below. Based on the calibrated risk rate ratios, we were also able to calibrate the endogenous parameters a_h and c_h, stratified across household SES strata *h*, based on the following equations:

(C.18a) $a_h = \overline{RR_i} * a_0\,, i \neq \text{'low BMI'}$

(C.18b) $a_h = RR_i * a_0\,, i = \text{'low BMI'}$

(C.19a) $c_h = \overline{RR_i} * c_0\,, i \neq \text{'low BMI'}$

(C.19b) $c_h = RR_i * c_0\,, i = \text{'low BMI'}$

where equations C.18a and C.19a were used to calibrate parameters *a* and *c* for the five exogenous risk factors, i.e. tobacco smoking, indoor air pollution, daily alcohol use, HIV infection, and diabetes, while equations C.18b and C.19b were used to calibrate parameters *a* and *c* for the endogenous risk factor, i.e. low BMI.

Equations C.18b and C.19b form part of the endogenous pathway linking nutritional intakes, determined by household-specific AIDS demand systems in the CGE model, to TB risk factor ratios (equation C.17), disease progression parameters (equations C.18b and C.19b) and TB health outcomes determined by the epidemiological model. Equations C.18b and C.19b therefore form an integral part of the fully integrated macroeconomic-epidemiological model framework. In contrast, equations C.18a and C.19a specifies TB risk factor ratios for the five exogenous risk factors, and they are therefore not adjusting, endogenously, within the model framework. However, these equations allow for analysing exogenous prevention-focussed changes to TB risk factors, and how they may affect economic and clinical health outcomes. Equations C.18a and C.19a are therefore also retained as

integral parts of the integrated model framework (even if the pathways, specified by these five risk factors, are not endogenously modelled).

C.4. Endogenous modelling of low BMI risk factor

As noted at the end of the previous section C.3, our integrated model framework specifies feedback effects, from the macroeconomic CGE model to the epidemiological TB model, via endogenous modelling of the 'low BMI' risk factor. Similar to the methodology used in Jensen et al. (2019), we specify intra-household distributions of the biomarker/risk factor in focus, and shift the intra-household distributions by changes in the mean, derived from the macroeconomic CGE Model.

In the current case, BMI is mediated by household-specific consumption of macro-nutrients in the form of kcal intakes. In order to measure mean changes in household-specific kcal intake, we specify a set of kcal nutrition-coefficients across the full range of (primary and processed food) commodities in our CGE model ($kcal\_coeff_c$) and apply them to household-specific consumption ($QH_{h,c,t}$) in order to measure changes in kcal intakes:

(C.20) $KCAL_{h,t} = \sum_{c=1}^{31} kcal\_coeff_c PQ0_{c,t} QH_{h,c,t}$

where household-specific consumption components ($QH_{h,c,t}$) are determined by the household-specific AIDS demand systems, specified in equation A.1 (see appendix A), and valued at constant prices ($PQ0_{c,t}$) allowing for application of our value-based nutritional coefficients ($kcal\_coeff_c$) to derive household-specific kcal intakes ($KCAL_{h,t}$).

The nutritional coefficients for kcal-contents of individual (primary and processed food) commodities ($KCAL_{h,t}$) were derived from the Indian food composition table of the SMILING database (SMILING 2016) and the USDA food composition database (USDA 2014). Food categories were mapped to the 31 commodities of the India SAM database.

In order to link average changes in kcal nutritional intakes to average changes in weight, and, by implication, to average changes in BMI, we relied on evidence indicating that a reduction in nutritional intakes of 7715 kcal is required to lose 1kg in weight (Sheelbeek et al. 2019). In order to measure specific policy scenario impacts on household-specific changes in average weight ($\Delta WGHT_{h,t}^{scen}$), we specified the change in weight as the difference in kcal nutritional intakes between the counterfactual simulation ($KCAL_{h,t}^{count}$) and scenario simulation ($KCAL_{h,t}^{scen}$) scaled by the 7715 kcal/kg conversion factor:

(C.21) $\Delta WGHT_{h,t}^{scen} = \frac{KCAL_{h,t}^{scen} - KCAL_{h,t}^{count}}{7215}$

While our simulation framework is set up to model average weight changes at the representative household level, the availability of individual-specific data on weight, height, and BMI, derived from the India 2015-16 NFHS-4 National Family Health Survey (IIPS 2017), allowed us to compute intra-household weight and height distributions, and thereby to allow for calculation of (changes to) household-specific distributions of weight and BMI. After data-cleaning, we managed to extract 787,840 individual records on coherent weight and height measurements, and to categorize them according to household SES quintiles (based on a wealth index provided as part of the NFHS-4 data base) (ibid.) Throughout, we define subscript $i\_strata = 1, \dots, 787{,}840$ to refer to our set of 787,840 individuals with associated coherent records of weight and height.

In terms of the calculation of intra-household distributions of weight and BMI, we shift intra-household distributions of weight levels ($WGHTstrata^{scen}_{i\_strata,t}$) across micro-individuals (i_strata), by changes in the mean weight ($\Delta WGHT^{scen}_{h,t}$) of the representative household to which these micro-individuals belong ($i\_strata \in h$), according to equation (C.22):

(C.22) $$WGHTstrata^{scen}_{i\_strata,t} = WGHTstrata^{scen}_{i\_strata,,t-1} + \Delta WGHT^{scen}_{h,t|i\_strata \in h}$$

Based on the new scenario-specific intra-household weight distributions ($WGHTstrata^{scen}_{i\_strata,t}$), we are then able to calculate the new scenario-specific intra-household BMI distributions ($WGHTstrata^{scen}_{i\_strata,t}$) according to equation (C.23):

(C.23) $$BMIstrata^{scen}_{i\_strata,t} = \frac{WGHTstrata^{scen}_{i\_strata,t}}{HGHTstrata_{i\_strata,t}{}^{2}}$$

and to calculate household-specific changes to the prevalence of the 'low BMI' risk factor ($p_{h,'lowBMI'}$) according to equation (C.24):

(C.24) $$p_{h,'lowBMI'} = \frac{1}{\sum_{i\_strata=1}^{787,840} 1[i\_strata \in h]} \sum_{i\_strata=1}^{787,840} 1[BMIstrata^{scen}_{i\_strata,t|i\_strata \in h} < 18.5]$$

which feeds into the calculation of household-specific relative risk ratios (RR_h) in equation C.17 (see section C.3.3 above), thereby completing the feedback mechanism pathway running from endogenous (food) demand determination in the macroeconomic CGE model to endogenous adjustment in the 'low BMI' risk factor within the epidemiological mode.

C.5. Epidemiological model: Parametrization and dynamic model fitting

We parameterized the household-specific epidemiological models using a set of exogenous parameter values, from the literature, which are constant across household types and not dependent on time step: (1) Proportion progressing directly to disease on infection (a_0), (2) Relative risk of active TB following re-infection (q), (3) Proportion successfully treated (τ), (4) Relative effective contact between SES quintiles (β_{out}); a set of exogenous parameter values, from the literature, which are constant across household types but dependent on time step (and therefore scaled to our bi-weekly time frequency): (5) Background mortality (u), (6) Excess mortality due to active TB (m), (7) Self-cure of active TB (ε), (8) Developing active TB from latent infection (c_0), (9) Completion of treatment (v); and two exogenous household-specific parameters where values were calibrated to fit steady-state household-specific TB prevalence rate outputs from the model to existing household-specific TB prevalence rates: (10) Effective contact between those in same SES quintile (β_{in}), (11) Diagnosis of active TB (d). All non-household-specific parameter values, both exogenous and calibrated, are presented in Table C.3. Data sources for the various data and parameter values used for calibrating the non-household-specific parameters are listed in Table C.3.

The calibration of initial values for the household-specific parameters, a and c, presented in Table C.4., were based on equations C.18-C.19 which defined a- and c-parameters in terms of baseline a0 and c0 values (Table C.3) and household-specific relative risk ratios (RR_i) (Table C.4). The household-specific RR_i were, in turn, derived from equation C. 17 which defines RR_i as a function of baseline risk factor prevalence rates (p_{ij}) (Table C.4) and baseline relative risks of risk factors (g_j) (Table C. 4). Data sources for the various data used for calibrating the household-specific parameters, a and c, are listed in Table C.4.

Table C. 3. Epidemiological model core parameter values

Symbol	Definition	Value	Source
a0	Proportion progressing directly to disease on infection*	0.084	Sumner and White (2020)
q	Relative risk of active TB following re-infection**	0.6	Vynnycky and Fine (1997)
τ	Proportion successfully treated***	0.8	Own calculation
β_{rel}	Relative effective contact between SES quintiles	7.22595E-2	Calibrated from simple fitting of model to data
Symbol	**Definition**	**Risk (weekly)******	**Source**
β_{in}	Effective contact between those in same SES quintile	3.44266E-01	Calibrated from simple fitting of model to data
U	Background mortality	3.84541E-04	Own calculation – assuming average life-expectancy of 50 years
m	Excess mortality due to active TB	1.62710E-03	Calibrated from simple fitting of model to data
ε	Self-cure of active TB	2.35629E-03	Calibrated from simple fitting of model to data
c_0	Developing active TB from latent infection*	1.14230E-05	Sumner and White (2020)
d	Diagnosis of active TB	1.01282E-2	Calibrated from simple fitting of model to data
v	Completion of treatment	3.77313E-02	Own calculation – assuming average treatment duration of 6 months

Notes: *Values of a0 and c0 are assumed to apply in absence of risk factors (see Table C.4 and equations C.18-C.19 for calculation of household-specific and risk-factor adjusted values a_h and c_h values); **Value for q assumes 40% protection against disease from re-infection due to prior exposure (point estimate for age>20 in Vynnycky and Fine (1997)) ***Value for τ is a weighted average of public and private sector treatment outcomes and DS and MDR (see section C.7 for details) ****Annual rates are converted to bi-weekly risks for the difference equation model using Risk = 1-exp(-rate/26).

Table C.4. Household-specific TB risk factor prevalence and TB relative risks, and household-specific epidemiological model parameters (a and c)

Risk factor prevalence data*						
	Household SES quintile (wealth)					
Risk factor	1	2	3	4	5	RR for risk factor**
Smoking	17.9	16.3	14.15	12.35	9.65	2.0
Indoor air pollution	79.1	68.5	51	25.2	8.3	1.4
Low BMI	33.9	28.1	21.8	16.7	11.1	2.1
Alcohol use	3.1	2.2	1.9	1.3	0.98	2.9
Diabetes illness	0.5	0.7	0.9	1.3	2.0	3.1
HIV infection	0.4	0.4	0.5	0.4	0.2	26.7
RR for risk of TB illness***	2.0008	1.8156	1.6133	1.3993	1.2421	
Household-specific parameters						
a_h****	0.16807	0.15251	0.13551	0.11754	0.10434	
c_h****	0.0000229	0.0000207	0.0000184	0.0000160	0.0000142	

Notes: *Risk factor prevalence data were obtained from the Indian 2015-16 NFHS-4 National Family Health Survey (IIPS 2017), including Smoking (NFHS-4 definition: Any use), Indoor air pollution (NFHS-4 definition: Use of biomass cooking fuel inside the home), Low BMI (NFHS-4 Definition: <18.5), Alcohol (NFHS-4 Definition: Daily use), Diabetes (NFHS-4 definition: Self-reported), HIV (NFHS-4 Definition: Testing in DHS); **Parameter values for 'RR for risk factor' were obtained from Oxlade and Murray (2012); ***Household-specific values for 'RR for risk of TB illness' were derived from risk factor prevalence data and 'RR for risk factor' parameter values via equation C.17 ****Household-specific and relative risk-adjusted values for parameters a_h and c_h were derived from values of a0 and c0 (Table C.3) and values for 'RR for risk of TB illness' (presented here) via equations C.18-C.19.

C. 6. Epidemiological model: Dynamic model fitting

We did a simple fitting of our epidemiological model to a set of aggregate TB incidence, mortality, and notification point estimates, and SES-specific prevalence data (see Table C5) by minimizing a mean-squared-error metric. In practice, we fitted the stand-alone TB model framework and derived fitted values for the parameters β_{in}, β_{out}, and d, and subsequently integrated the TB model framework with the demographic and macroeconomic CGE model frameworks using the full set of exogenous, endogenous (initial), and fitted parameter values listed in Tables C.3-C.4 and discussed in section C.5. For the purpose of model fitting, we assumed constant initial values for endogenous variables across household SES quintiles: $S_{h,t0} = \frac{0.98}{5}, L_{h,t0} = I_{h,t0} = \frac{0.02}{5}, Ir_{h,t0} = T_{h,t0} = Tr_{h,t0} = 0.$

Table C.5. Observed target values and fitted values for epidemiological model fitting, (per 100,000)

Variable		Observed values	Fitted values
TB incidence (2019)*		193 (132-266)	233
TB mortality (2019)*		32 (30-34)	36
TB notifications (2019)**		186	149
TB prevalence (2006-12)***		301 (224-378)	292
- by household SES quintile	1	497	468
(2015-16)****	2	365	362
	3	301	272
	4	234	199
	5	148	157

Notes: *WHO (2020, 2021b); **Observed nation-wide values for TB notifications obtained from WHO TB database (WHO 2021a) and adjusted for private sector´(NB: the model consistent notifications number differs slightly from WHO (2021b), where India notifications for 2019 was reported as 127 (public) + 50 (private) = 177 per 100,000) ***Observed nation-wide value for TB prevalence obtained from Chadha et al. (2019); ****Observed household-specific values for TB prevalence by SES quintile was derived from the India 2015-15 India NFHS-4 National Family Health Survey (IIPS 2017).

Table C.6. Approximate steady state values for endogenous variables within our stand-alone epidemiological TB model

	Initial values	Approximate steady-state values				
	(constant	by household SES quintile				
Variable	across hhlds)	1	2	3	4	5
$S_{h,t}$	0.98/5	9.4e-2	1.0e-1	1.1e-1	1.2e-1	1.3e-1
$L_{h,t}$	0.01/5	1.0e-1	9.5e-2	8.6e-2	7.6e-2	7.0e-2
$I_{h,t}$	0.01/5	8.2e-4	6.3e-4	4.8e-4	3.5e-4	2.8e-4
$Ir_{h,t}$	0	1.2e-4	9.2e-5	6.9e-5	5.1e-5	4.0e-5
$T_{h,t}$	0	2.4e-4	1.8e-4	1.4e-4	1.0e-4	8.0e-5
$Tr_{h,t}$	0	3.4e-5	2.7e-5	2.0e-5	1.5e-5	1.2e-5

Note: Own calculations with approximate steady-state values derived from stand-alone TB model simulations using single-week time steps.

C.7. Epidemiological model: Derivation of notification target for model fitting

This sub-section presents technical details on our calculation of the parameter-value for the "proportion of successfully treated"-parameter (τ) presented in Table C.3.

Notification data represents the reported number of people started on TB treatment. However, notification data reported to WHO by the India national TB programme does not capture all cases treated in the private sector so fitting to this would underestimate the diagnostic rate. If we know the proportion of cases started on treatment that are included in the notification data, P, we can use this to adjust the notification data to give the "true" number started on treatment (the number entering the T state in our model).

If we assume that all public sector treatments are notified, then:

$$P = P_{T,PUB} + xP_{T,PRI}$$

where $P_{T,PUB}$ is the proportion of all TB patients treated in the public sector, $P_{T,PRI}$ is the proportion treated in the private sector, and x is the proportion of private sector treatment that is notified.

Data from the National Family Health Survey (Pardeshi et al. 2018) suggests that the proportion of all TB patients that are treated in the private sector $P_{T,PRI} = 0.4$ (implying that $P_{T,PUB} = 0.6$), but we don't know what proportion of the privately treated who gets notified. We do have an idea of the proportion of all TB notifications that come from the private sector $P_{N,PRI} = 0.3$ in 2020 (WHO 2021c).

The proportion of all TB notifications that come from the private sector is given by:

$$P_{N,PRI} = \frac{x.P_{T,PRI}}{x.P_{T,PRI} + P_{T,PUB}}$$

Rearranging the above equation, we get the following expression for x:

$$x = \frac{P_{N,PRI}P_{T,PUB}}{P_{T,PRI}(1 - P_{N,PRI})} = \frac{0.3 \times 0{,}6}{0.4\ (1 - 0.3)} = 0.64$$

Plugging this value into the equation for P gives $P = 0.85$. This proportion can be used to scale up the notifications, reported by WHO, to reflect the true number of people started on TB treatment.

There is limited data on treatment outcomes among those treated in the private sector. Arinaminpathy et al. (2019) use data from Uplekar et al. (1998) to inform the proportion of individuals completing first line treatment in the private sector ($\tau_{PRI} = 0.6\ (0.4 - 0.8)$). Overall treatment success reported by WHO is currently $\tau_{WHO} = 0.84$ (WHO 2021a). Assuming that this

covers both public and private sector treatment success, and using the proportion of notifications stemming from the private sector, we can write

$$\tau_{WHO} = \left(1 - P_{N,PRI}\right)\tau_{PUB} + P_{N,PRI}\tau_{PRI}$$

which can be rearranged to give the treatment success in the public sector

$$\tau_{PUB} = \frac{\tau_{WHO} - P_{N,PRI}\tau_{PRI}}{1 - P_{N,PRI}} = \frac{0.84 - 0.3 \times 0.6}{1 - 0.3} = 0.94$$

The weighted average treatment success for DS-TB, τ_{DS}, can then be calculated as

$$\tau_{DS} = \tau_{PRIV}P_{T,PRI} + \tau_{PUB}P_{T,PUB} = (0.6 \times 0.4) + (0.94 \times 0.6) = 0.81$$

Finally, we note that approximately 3% of treated TB cases are being treated for multi-drug resistant (MDR) TB. Assuming that no MDR treatment is given in the private sector, and using an MDR treatment success rate of 56% (WHO 2021a), we can calculate the final weighted treatment success rate (τ) as:

$$\tau = \tau_{MDR}P_T^{MDR} + \tau_{DS}(1 - P_T^{MDR}) = (0.56 \times 0.03) + (0.81 \times (1 - 0.03)) = 0.80$$

Hence, we proceed to employ the parameter value of $\tau = 0.80$ for the purposes of our epidemiological simulations (see Table C.3).

References

Arinaminpathy N, Deo S, Singh S, et al. 2019. Modelling the impact of effective private provider engagement on tuberculosis control in urban India. Scientific Reports 9:3810. DOI: https://doi.org/10.1038/s41598-019-39799-7

Chadha VK, Anjinappa SM, Dave P, et al. 2019. Sub-national TB prevalence surveys in India, 2006-2012: Results of uniformly conducted data analysis. PLoS One 19(2):e0212264.

IIPS. 2017. Electronic data. National Family Health Survey (NFHS-4), India: 2015-16. Mumbai: International Institute for Population Sciences.

Menzies NA, Wolf E, Connors D, et al. 2018. Progression from latent infection to active disease in dynamic tuberculosis transmission models: a systematic review of the validity of modelling assumptions. Lancet Infectious Diseases 18(8):e228-38.

Pardeshi G, Deluca A, Agarwal S, Kishore J. 2018. Tuberculosis patients not covered by treatment in public health services: Findings from India's National Family Health Survey 2015-16. Tropical Medicine and International Health 23(8):886-95. URL: https://doi.org/10.1111/tmi.13086.

Scheelbeek PFD, Cornelsen L, Marteau TM, Jebb SA, Smith RD. 2019. Potential impact on prevalence of obesity in the UK of a 20% price increase in high sugar snacks: modelling study. BMJ 366:14786.

SMILING. 2016. Electronic data. Sustaining Micronutrient Interventions to Control Deficiencies and Improve Nutritional Status and General Health in Asia project. Institut de Recherche pour le Developpement. France. URL: http://www.nutrition-smiling.eu/.

Sumner T, White RG. 2020. The predicted impact of tuberculosis preventive therapy: the importance of disease progression assumptions. BMC Infectious Diseases 20:880. URL: https://doi.org/10.1186/s12879-020-05592-5.

Uplekar M, Juvekar S, Morankar S, Rangan S, Nunn P. 1998. Tuberculosis patients and practitioners in private clinics in India. International Journal of Tuberculosis and Lung Disease 2(4): p. 324-9.

USDA. 2014. Electronic data. USDA food composition database. United States Department of Agriculture Agricultural Research Service. United States. URL: https://ndb.nal.usda.gov/ndb/.

Vynnycky E, Fine PEM. 1997. The natural history of tuberculosis : the implications of age-dependent risks of disease and the role of reinfection. Epidemiology & Infection 119(2):183-201.

WHO. 2020. Global tuberculosis report 2020. Geneva: World Health Organization. URL: https://www.who.int/publications/i/item/9789240013131.

WHO. 2021a. Electronic data. WHO TB database for Global Tuberculosis Report 2021. World Health Organization: Geneva. URL: https://www.who.int/teams/global-tuberculosis-programme/data

WHO. 2021b. The COVID-19 pandemic and TB in India: impact and response. Online annex to Global Tuberculosis Report 2021. World Health Organization: Geneva.

WHO. 2021c. Global tuberculosis report 2021. Geneva: World Health Organization. URL: https://www.who.int/publications/i/item/9789240037021.

Appendix D. Calculation of TB clinical outcomes and health care costs

This appendix presents the equations which are used to calculate changes in clinical outcomes and health care costs. The calculations are based on the outputs from the epidemiological model (population shares in different states) and demographic model (population projections), and the changes in clinical outcomes and health care costs are subsequently used as inputs to the macroeconomic CGE model, thereby providing a feedback mechanism, and core pathway, from the demographic and epidemiological models to the macroeconomic CGE model.

D.1. Calculation of TB clinical outcomes

The calculation of clinical outcomes, including population-inflated numbers of incident cases, prevalent case, notifications, and deaths, are accomplished by equations D.1-D.4:

(D.1) $$TBcases\_incid_{h,age,t}^{annual} = Incid_{h,t}^{annual} TBincid\_share_{age} \sum_{gen} POP_{h,gen,age,t}$$

(D.2) $$TBcases\_mort_{h,age,t}^{annual} = Mort_{h,t}^{annual} TBmort\_share_{age} \sum_{gen} POP_{h,gen,age,t}$$

(D.3) $$TBcases\_prev_{h,age,t}^{annual} = Prev_{h,t}^{annual} \sum_{gen} POP_{h,gen,age,t}$$

(D.4) $$TBcases\_notif_{h,age,t}^{annual} = Notif_{h,t}^{annual} \sum_{gen} POP_{h,gen,age,t}$$

where the annual rate-variables of incidence ($Incid_{h,t}^{annual}$), mortality ($Mort_{h,t}^{annual}$), prevalence ($Prev_{h,t}^{annual}$), and notifications ($Notif_{h,t}^{annual}$), derived from equations C.10, C.12, C.14, and C.16 in appendix C.3.2, are multiplied by population aggregates ($\sum_{gen} POP_{h,gen,age,t}$) to calculate population level-variables for incidence ($TBcases\_incid_{h,age,t}^{annual}$), mortality ($TBcases\_mort_{h,age,t}^{annual}$), prevalence ($TBcases\_prev_{h,age,t}^{annual}$), and notifications ($TBcases\_notif_{h,age,t}^{annual}$).

We further note that the calculations of incident cases ($TBcases\_incid_{h,age,t}^{annual}$) and case fatalities ($TBcases\_mort_{h,age,t}^{annual}$) are also corrected for the age distributions of incident cases ($TBincid\_share_{age}$) and case fatalities ($TBmort\_share_{age}$). The age distribution parameters were established from WHO data on age-specific TB incidence and mortality rates in India, derived from the database underlying the Global Tuberculosis Report 2021 (WHO 2021).

D.2. Calculation of TB health care costs and isolation requirements

In calculating TB health care costs, we take a conservative approach and focus narrowly on the treatment costs associated with MDR- and DS-cases. In so doing, we are conscious of the fact that we disregard costs associated with e.g. testing and hospitalization of severe cases. However, we believe that the costs of treatment regimens constitute the main part of TB health care costs, and

that we therefore capture the main impacts on macroeconomic outcomes, but in a conservative way.

The cost of treatment regimens in India, for new and relapse adult pulmonary TB cases, range from a mean 136.15USD for DS-TB cases ($TBtreat\_cost^{DS} = 136.15USD/case$) to a mean 1229.29USD for MDR-TB cases ($TBtreat\_cost^{MDR} = 1220.29USD/case$) according to the recently published Value-data set (Sweeney et al. 2021). Furthermore, we assume that treatment involves isolation requirements of 8 weeks for DS-TB treatment ($TBisolat\_durat^{DS}$) and 16 weeks for MDR-TB treatment ($TBisolat\_durat^{MDR}$). We also note that MDR-TB cases were estimated to account for 2.84% (2.27–3.50%) of new TB cases in the First National Anti-Tuberculosis Drug Resistance Survey India 2014-16 (MoHFW 2018) prompting us to assume that the share MDR-TB cases, going forward, will be 3% ($MDRcase\_share = 0.03$).

Based on the above the above parameter values, we're able to calculate an estimate of the average TB treatment regimen unit cost:

$$TBtreat\_unitcost = MDRcase\_share * TBtreat\_cost^{MDR} + (1 - MDRcase\_share) * TBtreat\_cost^{DS}$$

$$= 0.03 * 1229.29 + (1 - 0.03) * 136.15$$

$$= 168.94\ USD/case$$

$$= 10{,}310.6\ INR/case$$

where the last calculation is based on a 2014 exchange rate of 61.03 INR/USD (WB 2021).

Based on the above average TB treatment regimen unit cost estimate ($TBtreat\_unitcost$), and the calculation of annual incident case numbers ($TBcases\_incid_{h,age,t}^{annual}$) in equation D.1, we're then able to specify the equation for calculation of total healthcare treatment regimen costs ($TBtreat\_cost$) as equation D.5:

(D.5) $$TBtreat\_cost = TBtreat\_unitcost * TBcases\_incid_{h,age,t}^{annual}$$

Based on the isolation requirements for the treatment of respectively MDR-TB and DS-TB, noted above, and for the purposes of calculating the TB morbidity impacts on labour supplies (see equation A.2' in appendix A.3), we also calculate average isolation requirement for TB treatment:

$$TBisolat\_durat = MDRcase\_share * TBisolat\_durat^{MDR} + (1 - MDRcase\_share) * TBisolat\_durat^{DS}$$

$$= 0.03 * \left(\frac{16}{52}\right) + (1 - 0.03) * \left(\frac{8}{52}\right)$$

$$= 0.158\ years/case$$

References

MoHFW. 2018. Report of the First National Anti-Tuberculosis Drug Resistance Survey India 2014-16. Ministry of Health & Family Welfare, Government of India. URL: https://tbcindia.gov.in/showfile.php?lid=3315.

Sweeney S, Cunnama L, Laurence Y, Garcia Baena I; Kairu A, Minyewelet M, Eyob H, Chatterjee S, Toshniwal M, Chikovani I, Shingelia N, Capeding TJ, Vassall A. 2021. Value TB Dataset: costs per intervention. Harvard Dataverse. DOI: https://doi.org/10.7910/DVN/QOI6IR.

WB. 2021. Electronic data. World Development Indicators (WDI) database. Washington DC: World Bank. URL: https://databank.worldbank.org/source/world-development-indicators.

WHO. 2021. Electronic data on TB incidence and mortality by age in India. Online github repo database annex to Global Tuberculosis Report 2021. World Health Organization: Geneva. URL: https://github.com/GTB-TME/gtbreport2021.