IATSS Research 36 (2012) 75–82

Contents lists available at SciVerse ScienceDirect

IATSS Research

Modeling road traffic fatalities in India: Smeed's law, time invariance and regional specificity

Raj V. Ponnaluri *

Operations Management, Administrative Staff College of India, Bella Vista, Raj Bhavan Road, Khairatabad, Hyderabad 500 082, Andhra Pradesh, India

ARTICLE INFO

Article history:
Received 26 December 2011
Received in revised form 15 April 2012
Accepted 14 May 2012

Keywords:
Smeed's law
Andreassen
Fatality rate
Mathematical model
India
Panel data regression

ABSTRACT

Mathematical formulations linking road traffic fatalities to vehicle ownership, regional population, and economic growth continue to be developed against the backdrop of Smeed and Andreassen models. Though a few attempts were made, Smeed's law has not been fully tested in India. Using the 1991–2009 panel data from all states, this work (a) developed the generalized Smeed and Andreassen models; (b) evaluated if traffic fatalities were impacted by structural changes; and (c) examined if – in relation to the generalized model – the individual (time and regional) models are more relevant for application. Seven models (Smeed: original, generalized, time-variant, state-variant; and Andreassen: generalized, time-variant, state-variant) were developed and tested for fit with the actual data. Results showed that the per vehicle fatality rate closely resembled Smeed's formulation. Chow-test yielded a significant *F*-stat, suggesting that the models for four pre-defined time-blocks are structurally different from the 19-year generalized model. The counterclockwise rotation of the log-linear form also suggested lower fatality rates. While the new government policies, reduced vehicle operating speeds, better healthcare, and improved vehicle technology could be the factors, further research is required to understand the reasons for fatality rate reductions. The intercept and gradients of the time-series models showed high stability and varied only slightly in comparison to the 19-year generalized models, thus suggesting that the latter are pragmatic for application. Regional formulations, however, indicate that they may be more relevant for studying trends and tendencies. This research illustrates the robustness of Smeed's law, and provides evidence for time-invariance but state-specificity.

1. Introduction

More than four decades after achieving political independence, India began a massive deregulation of its economy in 1991. New policies of the Government of India encouraged private sector participation and allowed foreign direct investment in several sectors including transport. Table 1 shows that the country's gross domestic product achieved a compounded annual growth rate (CAGR) of 4.2% from 1970 to 1990 which almost doubled to 7.7% CAGR during the following 19 years. Road traffic crashes and fatalities also increased but at slower CAGRs of 2.9% and 4.5% as against 4.6% and 6.8% respectively during the previous 20 years. Investments in the transport sector appear to be more pronounced in the 1990–2009 period as evidenced from road length increase at 3.9% CAGR as against 2.6% CAGR from 1970 to 1990. Also, successive governments at the federal level have brought about structural and policy changes since 1991. Given the impact of post-liberalization, this work used the 1990–2009 data to develop mathematical models for studying road traffic fatalities. It is worth noting that from 1970 to 2009, while the death rate dropped from 103 to 11 fatalities per 10,000 vehicles — mainly due to a large increase in the number of automobiles, the per capita vehicle ownership increased phenomenally from 260 to 9902 vehicles per 100,000 persons (see Table 1).

Mathematical models linking road traffic fatalities to vehicle volumes, regional population, and economic growth continue to be developed against the backdrop of Smeed's [1,2] seminal work and further analysis by Andreassen [3,4] and others. In India, several road traffic crash studies [5–7] have been done but only a few applied Smeed's law or other methods, especially for developing region-specific or time-dependent models; Valli [8], for example, developed models for metropolitan cities. The main purpose of this work is to examine the relevance of Smeed's law to India and to develop time- and state-specific models. Using a 19-year panel data (1991–2009) from 29 regions across India, this work intends to (a) develop Smeed and Andreassen models and test their fit with actual data; (b) evaluate the shift in coefficients, possibly due to structural changes; and (c) develop and examine the statistical significance of state (cross-sectional) and time-series regressions. Given that India carries a significant 17% of the world's population and 16% of the global road traffic fatalities [9], it is imperative to develop reliable models for studying traffic fatalities and proposing crash mitigation measures.

* H.No. 8-7-105/37, Rd. 1, Balaji Encl., Mallikarjuna Colony, Old Bowenpally, Hyderabad 500 011, Andhra Pradesh, India. Tel.: +91 40 27957019, +91 9440887100; fax: +91 40 66496600.

E-mail addresses: hponnal@yahoo.com, rvp@asci.org.in.

0386-1112/$ – see front matter © 2012 International Association of Traffic and Safety Sciences. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.iatssr.2012.05.001

Table 1
GDP and road-related data from India (1970–2009).
Source: Reserve Bank of India & Ministry of Shipping, Road Transport & Highways, GoI.

	GDP (INR in 10 M)	Road crashes ('000)	Fatalities ('000)	Registered vehicles ('000)	Road length ('000 km)	Population ('000)	Fatalities/10,000 vehicles	Vehicles/100,000 persons
1970	474,131	114.1	14.5	1401	1188.7	539,000	103	260
1990	1,083,572	282.6	54.1	19 152	1983.9	835,000	28	2294
2009	4,464,081	486.4	125.7	114,951	4120.0[a]	1,160,813	11	9902
CAGR								
1970–1990	4.2%	4.6%	6.8%	14.0%	2.6%	2.2%		
1990–2009	7.7%	2.9%	4.5%	9.9%	3.9%	1.7%		
1970–2009	5.9%	3.8%	5.7%	12.0%	3.2%	2.0%		

[a] 2009 Road Length is estimated.

2. A review of road traffic fatality models

Based on his study of 1938 data from 20 countries, Smeed [2] approximated the per vehicle fatality rate vis-à-vis the per capita vehicle ownership model as follows:

$$F/V = C(V/P)^{\beta} \quad (1a)$$

where F = number of road traffic fatalities; V = number of vehicles; P is the population; C = constant (0.0003); and $\beta = -0.67$ or $-2/3$.

Eq. (1a) can be rewritten as

$$Ln(F/V) = \alpha + \beta Ln(V/P) \text{ or } (F/V) = e^{\alpha}(V/P)^{\beta} \quad (1b)$$

where 'α' represents the constant term (e^{α}, to be precise).

Adams [10], as have other authors, recounted Andreassen's [3] critique of Smeed and referred to evidence [11] for 'how well the 'apparent relationship' has stood the test of time despite the passage of 42 years.' Adams also discussed the efficacy of road safety measures and 'behavioral adjustments', who seems to be supported by the risk homeostasis hypothesis developed by Wilde [12]; Wilde was later critiqued by O'Neill and Williams [13]. The complex human decision-making patterns have been studied extensively in such diverse areas as travel behavior [14], road safety [15,16] and behavioral finance [17]. Smeed himself may have suggested the original per vehicle death or fatality rate formulation on the basis of his grassroots understanding of driver behavior, group psychology, and 'fatalistic view of traffic flow' as described by Dyson [18].

Interestingly, Jacobs and Hards [19] found that Smeed's equation remained unchanged when the analysis was repeated for the same 20 countries using the 1950, 1960, and 1970 data. Later, Mekky [20] found close approximation to Smeed's work with actual data. However, a review of two studies [21,22] from 30 to 34 developing countries respectively, showed a 'clockwise rotation' [23] of the Smeed relationship, suggesting a higher fatality rate for a given level of vehicle ownership. The study [23] of 12-year data from 17 Asia-Pacific countries and showed a clockwise relationship in the 1980–1987 period followed by a 'leveling off' in the years 1988–1992. The new generation models have begun including parameters beyond vehicle volumes and population to explain the variability in road traffic fatalities. For example, Jacobs and Cutting [24], in addition to explaining 'slope rotation', studied the role of other characteristics such as gross national product, and road and vehicle densities. Smith [25] considered modal split and suggested the use of cross-sectional data across a time-series for model development.

Elvik [26] observed that, given the varying viewpoints, a synthesis of accident models has not emerged. Indeed, literature points to numerous studies including the world-encompassing works such as those of Kopits and Cropper [27], Oppe [28], and Navin et. al's [29] extension of Smeed to include a third dimension, 'deaths per vehicle', to explain the fatality rates. Valli [30] developed mathematical models for metropolitan cities in India while Al-Matawah and Jadaan [31] developed prediction models for the United Arab Emirates, Jordan, and Qatar. Akgüngör and Doğan [32] applied Smeed and Andreassen models, and developed artificial neural network formulations for metropolitan cities in Turkey while Ackaah and Salifu [33] developed a framework for road crashes on rural highways in Ghana.

3. Theoretical approach and application methods

3.1. Data description

The Government of India routinely collects information from various sources on all aspects of the economy. The collated data from the Ministries of Finance, Home Affairs, and Road Transport and Highways were synthesized [34] and analyzed. A complete set of data was available for 26 states and 3 union territories. As a result, a 'balanced panel' comprising a matrix of 29 cross-sectional units (states and union territories) by 19 time-points (years) was prepared for analysis and model development. Thus, the 1991–2009 panel data were studied to understand the relationship between road traffic fatalities, vehicle ownership, and population. A review of the panel dataset helped observe that almost all states across the country have reported complete road accident information to the federal government, as evidenced from the availability of data for individual states during the study period. Trend analysis, not presented here, showed that each state had its own natural progression in road accident incidence. Author's experience also corroborates that a few states began implementing road safety systems after noticing such trends. In terms of road accidents, a few states consistently ranked within the top three to five. The registered motor vehicle data also showed a similar trend in that all states now appear to report detailed information such as vehicle type and age. While the road accident and motor vehicle data are reported by individual states, the population figures are estimated by the government and its various Ministries. Population data also displayed trends which varied by the state. The dataset also included other parameters such as the gross state domestic product.

The juxtaposing of the 551 (19×29) row elements (comprising 'F', 'V', 'P', 'F/V', and 'V/P') yielded 33 invalid data points and missing cell values, which reduced the usable dataset to 518 elements. The F/V parameter was then regressed over the independent V/P variable.

3.2. Scenario development

While building the various scenarios, as expected, the scatter plot (Fig. 1) showed the classic reciprocal form and yielded a relationship similar to that of Smeed's; the results will be discussed in Section 4. Eq. (1a)'s modified natural logarithmic form was then utilized to draw meaningful conclusions regarding the change in per vehicle fatality rate over a time-series (Eq. (2)) and regional cross-section (Eq. (3)). Detailed statistical analyses were then performed and individual equations were derived for time and regional variances. The