1993

Speed Variance, Enforcement, and the Optimal Speed Limit

PHILIP E GRAVES, University of Colorado at Boulder
DWIGHT LEE, Southern Methodist University
ROBERT L SEXTON, Pepperdine University

Available at: https://works.bepress.com/philip_graves/64/
Speed variance, enforcement, and the optimal speed limit

Philip E. Graves*
Department of Economics, University of Colorado, Boulder, CO 80309, USA

Dwight R. Lee
Department of Economics, University of Georgia, Athens, GA 30602, USA

Robert L. Sexton
Social Science Division, Pepperdine University, Malibu, CA 90263, USA

Received 18 March 1993
Accepted 14 April 1993

Abstract

A model of the optimal speed limit is developed which explicitly recognizes the roles of average speed, speed variance, and the level of enforcement. An unusual result emerges, namely that a higher speed limit may be optimal when reducing the variance in highway speeds reduces accident externalities.

1. Introduction

When contemplating why speed limits are introduced in the United States and other countries, the impact of average highway speed on accident externalities comes first to mind. It would appear obvious that any reduction in highway speeds resulting from a lowering of the speed limit will reduce the number and severity of traffic accidents. But it may not be the reduction in average highway speed that generates the major benefit from a lower speed limit. It appears that a reduction in the speed limit makes its largest contribution to highway safety by decreasing the variance in highway speeds. For example, in a recent empirical investigation of the 55 mph speed limit, Forester et al. (1984) estimated that, "the greater concentration of speeds induced by the [55 mph] speed limit has reduced annual fatalities by 11,545 versus a reduction of 5327 deaths resulting from lower average speeds". Further support for this finding is provided by Lave (1985), although the result is considered controversial [see, for example, Fowles and Loeb (1989), Levy and Asch (1989), Snyder (1989), and the rejoinder by Lave (1989)].

It would appear then that recognizing the importance of variations in speed on highway safety would push in the direction of reducing the legally allowable speed, since this would reduce the variation in speed. Interestingly, however, this is not necessarily the case, even given that

* Corresponding author.
reductions in speed limits do reduce speed variance. Under reasonable assumptions on the costs and benefits of imposing an effective speed limit, it will be shown that the optimal speed limit (but not average speed) can increase if the negative relationship between the variance in highway speed and traffic safety becomes more pronounced.

In the next section a simple model of the optimal speed limit is presented which incorporates average speed, speed variance, and the level of enforcement. It is, ironically, by incorporating into the model the safety benefits from reducing the variance in speed that one may be able to justify increasing the speed limit above the 55 mph that still applies to most of the driving done in the United States. Furthermore, once the importance of speed variance is recognized, it is clearly possible that the optimal speed limit in the wide open western states (e.g. Montana) may be less than in the more congested eastern states (e.g. Maryland). This seemingly perverse result is shown to stem from more plausible assumptions than one would initially suspect. Some concluding remarks and caveats to the analysis are offered in the closing section.

2. Speed variance and the optimal speed limit

As discussed above, while reductions in average highway speed may be a benefit from the setting of speed limits, another benefit from a speed limit may come from a reduction in the variance of highway speeds. In this section we develop a model to account for both of these benefits.

For simplicity it is assumed that the accident externalities from both average speed, S, and speed variance, V, are separable functions of the speed limit set, L, and the level of policing activity, P. Considering first average highway speed, S(L, P), over relevant ranges of L and P, we assume that S increases at a decreasing rate with respect to L and decreases at a decreasing rate with respect to P. Letting subscripts represent partial derivations with respect to the indicated variable, we have

\[ S_1 > 0, \quad S_{11} < 0, \quad S_2 < 0, \quad S_{22} > 0. \]

Similarly, speed variance, V(L, P), would be expected to be increasing in L at a decreasing rate and decreasing in P at a decreasing rate, or

\[ V_1 > 0, \quad V_{11} < 0, \quad V_2 < 0, \quad V_{22} > 0. \]

The goal of those determining the speed limit and level of policing effort is to maximize net benefits, incorporating accident externalities that depend on both average speed and speed variance. The private net benefit realized from the average highway speed is given by the function

1. See Graves et al. (1989) for greater detail on the role of enforcement in a simpler model which does not consider variance in the setting of the optimal speed limit.
2. The interesting possibilities that arise from the analysis are not dependent on this simplification.
3. Obviously the magnitude of these partials will depend upon the severity of the penalties imposed on those detected in violation of the speed limit. For example, the fine for going 65 mph in a 55 mph zone is $55 in Montana (with no points assessed) while in Maryland the fine would be $50 (with points) [see Newsweek, 21 July, 1986, p. 15, for information on other states]. In order to focus attention on the policy variables L and P, the penalty structure will be assumed fixed throughout the analysis, ignoring regional variations. The sign on the cross partial S_{12} will also be of significance in subsequent analysis. Since an increase in L will find more motorists obeying the speed limit voluntarily, it is reasonable to assume that increasing L will reduce the negative effect an increase in P has on S, or S_{12} = S_{22} > 0.
4. A priori reasons for believing the sign on V_{11} is positive are no more compelling than reasons for believing it is negative. With this justification, we make the simplifying assumption that V_{11} = 0.
\[ B(S). \text{ Since the purpose of a speed limit is to keep motorists from traveling as fast as they otherwise would, it is assumed that } B'(S) > 0 \text{ over the relevant range of speed with } B''(S) < 0. \text{ The costs to be considered by decision-makers are, in part, those of accident externalities, which are plausibly taken to be increasing at an increasing rate in average speed:} \]
\[ C'_1(S) > 0 \text{ and } C''_1(S) > 0. \]

Assume, for simplicity, that accident externalities increase at a constant rate, \( \alpha \), with variance.\(^5\)

In addition, the marginal and average costs of policing are taken to be given by the positive constant \( \theta \). Finally, it is assumed that there is some speed \( \bar{L} \) below which it is politically impossible to lower the speed limit.\(^6\) Recent legislative activity raising certain highway speed limits above 55 mph suggests that this speed limit may well be a reasonable approximation for \( \bar{L} \).

The objective of speed limit policy may now be expressed as solving for the \( L, P, \) and \( \lambda \) which maximize:
\[ Z(L, P, \lambda) = B[S(L, P)] - C_1[S(L, P)] - \alpha V(L, P) - \theta P + \lambda(L - \bar{L}). \]  

(1)

Omitting for brevity the non-negativity and complementarity conditions, the Kuhn–Tucker conditions for an extremum are\(^7\)
\[ [B'(S) - C'_1(S)]S_1 - \alpha V_1 + \lambda \leq 0 \]  

(2)
\[ [B'(S) - C'_1(S)]S_2 - \alpha V_2 - \theta = 0 \]  

(3)

The intuition behind conditions (2) and (3) is straightforward. If the constraint \( L \geq \bar{L} \) is not binding, then \( \lambda = 0 \) and (2) simply calls for reducing the speed limit until the \( MB \) from an increase in that limit rises to zero. If, however, the constraint \( L \geq \bar{L} \) is binding, then [as the LaGrangian (1) is set up], \( \lambda > 0 \) and it follows from (2) that, at \( L = \bar{L} \), the marginal benefit from increasing the speed limit is negative. The efficient speed limit would be less than \( \bar{L} \), except for the constraint. Condition (3) calls for an increase in policing until its marginal value, \( [B'(S) - C'_1(S)]S_2 - \alpha V_2 \), is equal to its marginal cost, \( \theta \).

In order to examine the implications of speed variance for the optimal speed limit policy, it is convenient to begin with the case where variance does not matter, \( \alpha = 0 \). When \( \alpha = 0 \), we know that \( B'(S) - C'(S) < 0 \) [from condition (3), with \( S_2 < 0 \)]. Hence condition (2) holds as an inequality, and \( L = \bar{L} \).

With standard assumptions on the smoothness or differentiability of the functions, it follows that over some positive range for \( \alpha \), \( B' - C' \) will remain negative and the optimal speed limit will remain \( \bar{L} \). Over this range, condition (2) can be ignored as condition (3) alone defines \( P \) as a function of \( \alpha \). Differentiating (3) through with respect to \( \alpha \), holding \( L \) constant at \( \bar{L} \) and recognizing that \( P \) is a function of \( \alpha \), yields
\[ \frac{dP}{d\alpha} = \frac{V_2}{(B' - C')S_{22} + (B'' - C'')S_2^2 - \alpha V_{22}}. \]

(4)

\(^5\) Reducing the variance in speeds can affect the private net benefit function by reducing the personal danger of increasing speed; \( B'(S) \) should increase as \( V \) is decreased. Since this would increase the likelihood of the 'pervasive' possibility that is the notable result of this paper, the effect of \( V \) on \( B(S) \) will be conveniently and harmlessly ignored.

\(^6\) We let \( L \) be sufficiently low so that if a speed limit of \( L \) were perfectly enforced, the result would be an average highway speed of \( S \), \( S < L \), and \( V \), where \( B'(S) - C'_1(S) < 0 \).

\(^7\) The sufficient conditions are guaranteed by earlier text restrictions.
Over the range of \( \alpha \) being considered, both the numerator and the denominator on the right-hand side of (4) are negative and \( dP/d\alpha \) is unequivocally positive. As \( \alpha \) increases, thereby increasing the impact of speed variance on accident externalities, and, in consequence, the value of policing, policing should be increased.

As \( P \) increases, with \( L \) remaining constant at \( \bar{L} \), average highway speed will decline. Increasing \( \alpha \) will eventually result in sufficient policing to push \( B' - C' \) above zero (recall footnote 6). With \( [B' - C']S_1 \) positive and increasing in \( \alpha \) (recall from footnote 3 that \( S_1 \) increases in \( P \) and therefore in \( \alpha \)), it is possible for this term to increase to \( \alpha V_l - \lambda \), at which point condition (2) holds as an equality. Once (2) becomes an equality, then (2) and (3) jointly define both \( L \) and \( P \) as functions of \( \alpha \). The task is now to investigate the simultaneous influence \( \alpha \) has on these two policy variables.

Differentiating (2) and (3) with respect to \( \alpha \), while recognizing that \( L \) and \( P \) are both functions of \( \alpha \), yields the system of equations where \( a_{12} = a_{22} = \frac{dL}{d\alpha} = [B' - C']S_1 \), \( a_{11} = a_{21} = \frac{dP}{d\alpha} \), \( V_1 \), and \( V_2 \), \( S_1 \), \( S_2 \), \( B' - C' \) \( S_1 \), and \( S_2 \), \( \lambda \) are all positive, whereas \( a_{22} \) \( S_1 \), \( S_2 \), \( B' - C' \) \( S_1 \), and \( S_2 \) are all negative. The matrix is negative definite with \( a_{11} < 0 \) and \( a_{22} < 0 \). Solving (5),

\[
\begin{bmatrix}
  a_{11} & a_{12} \\
  a_{21} & a_{22}
\end{bmatrix}
\begin{bmatrix}
  \frac{dL}{d\alpha} \\
  \frac{dP}{d\alpha}
\end{bmatrix}
=
\begin{bmatrix}
  V_1 \\
  V_2
\end{bmatrix},
\]

we obtain

\[
\frac{dL}{d\alpha} = \frac{V_1 a_{22} - V_2 a_{12}}{D}
\]

and

\[
\frac{dP}{d\alpha} = \frac{V_2 a_{11} - V_1 a_{21}}{D},
\]

where \( D \) is the positive determinant of the matrix in (5). Since both \( V_1 a_{22} < 0 \) and \( V_2 a_{12} < 0 \), the possibility that \( dL/d\alpha \) is negative cannot be ruled out. It is likely, however, that \( |V_1 a_{22}| \) \( V_2 a_{12} \), particularly if \( V_1 \) is small relative to \( V_2 \), in which case \( dL/d\alpha > 0 \). From (7) it is easily seen that \( dP/d\alpha > 0 \) if \( V_2 a_{11} > V_1 a_{21} \). The interesting result here, however, is that the optimal speed limit can increase as the negative safety impact of variance in speeds becomes more pronounced.

---

Footnote 2: Two caveats should be noted. First, as \( B' - C' \) becomes positive, \( dP/d\alpha \), as expressed in (4) is no longer unequivocally positive. Presumably, however, policing will continue to increase in \( \alpha \) beyond the point where \( B' - C' \) is positive. Secondly, the more positive \( V_l \) the less likely, all other things equal, (2) will be established as an equality as \( \alpha \) increases. It is certainly possible, however, the \( V_l \) is only weakly positive and \( B' - C' \) \( S_1 \) will increase to \( \alpha V_l \) as \( \alpha \) increases (a possibility, of course, even if \( V_l \) is rather large). At a strictly intuitive level the sign on \( V_l \) is not completely obvious. For example, it is conceivable that a reduction in the speed limit, policing constant, could increase the variance in highway speeds as some people slow down but others do not. We continue, however, to assume that \( V_l > 0 \). But as will become clear, the possibilities highlighted in this paper are more likely if \( V_l \) is rather small in comparison with the absolute value of \( V_l \).

Footnote 3: If, as soon as (2) holds as an equality, \( dL/d\alpha \) is negative, then increasing \( \alpha \) would make it desirable again to reduce the \( L \) below \( L \). This would mean that the constraint \( L > L \) is once again binding and that (2) once again holds as a strict inequality. In other words, if \( dP/d\alpha \) is not positive over some range, then increasing \( \alpha \) from zero will increase the desirable \( L \) (the one that would be chosen in the absence of a constraint) to a maximum of \( L \) at which point further increases in \( \alpha \) reduce the desirable \( L \). Obviously, this would be a most unusual situation.
This possibility can be usefully illustrated diagrammatically. In Fig. 1 the marginal value of policing speed limit $\bar{L}$ (in terms of reduced externalities from reduced average speed), $MVP(P, \bar{L})$, is shown, along with the marginal costs of policing with the additional marginal benefit from policing the speed limit that comes from reducing variance in speeds. For a given $\alpha$ this marginal benefit is the absolute value of $\alpha V_2(L, P)$. In Fig. 1 this marginal benefit is deducted from the marginal cost of policing to get a net marginal cost of policing, $\alpha V_2(L, P) + \theta$. For a sufficiently large $\alpha$ this net marginal cost will be negative when it intersects the curve $MVP(P, \bar{L})$, as is the case in Fig. 1. Since $V_2 < 0$, the net marginal cost curve slopes upward, eventually crossing the horizontal axis and approaches $\theta$ asymptotically, one would expect that $V_i \to 0$ as $P \to \infty$ and full compliance is realized.

The intersection between $MVP(P, \bar{L})$ and $\alpha V_2(L, P) + \theta$ calls for a speed limit policy of $\bar{L}, \hat{P}$, which satisfies condition (3) but which might not satisfy condition (2). Since $MVP(\hat{P}; \bar{L}) < 0$, it follows that $B' - C' > 0$ at $L, \hat{P}$ and, therefore, it is possible that

$$[B'(S) - C'(S)]S_1 > \alpha V_1 - \lambda$$

at $\bar{L}, \hat{P}$. If this is the case, the marginal benefit from increasing the speed limit can exceed the marginal cost resulting from greater variance in speeds, and the optimality conditions call for a speed limit in excess of $L$. As the speed limit increases, the relevant $MVP$ curve becomes less steeply sloped than $MVP(P; \bar{L})$, falling below $MVP(P; \bar{L})$ over some initial range of $P$ but lying

$^{10}$The value of $L$ in $V_2(L, P)$ is not specified because of the earlier assumption that $V_{i2} = 0$ (footnote 4).
above \( MV(P; \hat{L}) \) thereafter.\(^{11}\) At some speed limit \( L^* > \hat{L} \), the curve \( MV(P; L^*) \) will intersect \( aV_r + \theta \) at policing level \( P^* \), as shown in Fig. 1, with condition (2) holding as a strict equality at \( L^*, P^* \).\(^{12}\) Condition (3) is, of course, also satisfied at \( L^*, P^* \) by virtue of the intersection.

Though it at first seems implausible that increasing the importance of speed variance on highway safety could require an increase in the optimal speed limit, the intuition behind this result is straightforward. Beginning with a situation in which the safety benefit from reducing speed variance is ignored, we saw that optimality called for a policy that put the major burden of controlling highway speeds on the speed limit with little burden being shouldered by policing [see Graves et al. (1989) for further details]. Given this low speed limit–low policing policy, it is easy to see why, once the benefit from reducing variance in highway speeds is introduced, it will be an increase in policing that is called for. But an increase in the policing of a given speed limit will reduce the average highway speed and, by doing so, increase the marginal net benefit from a higher average speed. This can find it desirable to allow an increase in highway speeds by increasing the speed limit while relying primarily on policing to keep the variance in highway speeds under control.\(^{13}\)

On the highly congested highway systems in the east (and, of course, the west coast), speed variance is a significant safety consideration. Therefore, the analysis developed here suggests that the optimal speed limit is probably lower in Montana than it is in Maryland, and with certainty is not higher. This runs quite counter to recent legislation.

It should be emphasized that this is not the same as saying that the optimal average speed is likely lower in Montana than in Maryland. Indeed, this emphatically is not the case. If the optimal speed limit is higher in Maryland than in Montana, it is because the optimal level of policing is high in Maryland relative to Montana, and this reduces the optimal average speed in Maryland below that in Montana. This is true even if it is assumed that the \( B(S) \) and \( C(S) \) functions are the same in Maryland as in Montana with the only relevant difference between the two states being reflected in the importance of speed variance on highway safety (the value of \( \alpha \) in our simplified model). And the actual state differences in these functions surely call for a higher average speed in Montana than in Maryland when differences in \( \alpha \) are ignored.\(^{14}\) The higher \( \alpha \) for Maryland will serve to increase this difference in the optimal average speed, as opposed to the optimal speed limit.

\(^{11}\) At the higher speed limit there is less benefit over small \( P \) from additional policing since increasing \( L \) will reduce the negative effect an increase in \( P \) has on average speed (recall footnote 3). However, as policing is increased and average speed is greater under the higher limit than it would have been under the lower one, the marginal value of policing the higher limit will eventually become greater than the marginal value of policing the lower limit. It is obvious, for example, that \( MV(P; L) \) will decline to zero at a small \( P \) the lower is \( L \).

\(^{12}\) This intersection necessarily occurs below the axis since the satisfaction of (2) as an equality required \( B' - C' > 0 \).

\(^{13}\) It might be argued that the problem of variance in speeds would be best handled by a minimum speed limit, something our formal analysis does not consider. But the inclusion of a minimum speed limit, while complicating the analysis a bit, would not alter the results developed here. For the same reason the posted speed limit is set as low as politically feasible in the model in section 2, as soon as variance in speeds becomes a consideration the smallest politically feasible difference between the maximum and minimum speed limits would be called for. As the importance of variance on accident externalities increases, the optimal policing of both limits will also increase in order to reduce the actual, as opposed to the legal, variance in highway speeds. With the maximum speed limit set at \( L \), increasing policing in response to an increasing importance in speed variance will eventually result in an average highway speed that is too low. At some point it will pay to increase the maximum speed limit even if doing so increases the minimum politically feasible difference between the maximum and minimum speed limits. In other words, it is entirely possible that the optimal response to an increase in the importance of controlling variance in highway speeds is to increase the legally allowed variance in speeds and also to increase the maximum legal speed.

\(^{14}\) Assuming, as seems reasonable, that \( \alpha \) is close to zero in Montana, the optimal speed limit in Montana will be \( \hat{L} \), quite independently of the \( B(S) \) and \( C(S) \) functions.
3. Conclusion

Most studies of the speed limit consider the 55 mph speed limit and, after making a benefit–cost comparison, conclude that this speed limit is too low. The appropriate conclusion of such studies is not that the 55 mph speed limit is too low, but instead that the average highway speed is too low. As discussed in Graves et al. (1989), once it is recognized that average highway speed depends on both the speed limit and policing, an obvious advantage can be seen in allowing highway speeds to increase by reducing policing (which has real resource costs) rather than increasing the speed limit.

By ignoring the important role policing plays in enforcing the speed limit, we show here that existing work has also failed to notice interesting policy implications flowing from the connection between speed variance and traffic safety. Existing empirical work indicates that reducing the speed limit reduces speed variance, which reduces accident externalities. It would appear then that the influence of speed variance on safety would exert downward pressure on the optimal speed limit. However, when policing is taken into consideration, the present analysis suggests that the effect of speed variance can be to increase the optimal speed limit, not reduce it. Indeed, it may well be that because speed variance has a greater impact on highway safety in Maryland than in Montana, it is possible that the optimal speed limit is higher in Maryland than in Montana.

At a more general level, all regulations require enforcement if they are to be effective. Although this is a rather obvious point, much economic analysis proceeds as if somehow regulations are self-enforcing. This can lead to policy conclusions that are questionable, as in the present exploration of the role of speed variance in the setting of optimal speed limits.

References