Does the Potential to Merge Reduce Competition?*

Dirk Hackbarth[†] Bart Taub[‡]

November 11, 2020

Abstract

We study anti-competitive horizontal mergers in a dynamic model with noisy collusion. At each instant, firms either privately choose output levels or merge to form a monopoly, trading off the benefits of avoiding price wars against the costs of merging. The potential to merge decreases pre-merger collusion, as punishments effected by price wars are weakened. We thus extend the result of Davidson and Deneckere (1984), who analyzed the weakening of punishments post-merger, demonstrating that pre-merger collusion is weakened, in a fully stochastic model. Thus, although anti-competitive mergers harm competition ex-post, the implication is that barriers and costs of merging due to regulation should be reduced to promote competition ex-ante.

JEL Classification Numbers: D43, L12, L13, G34.

Keywords: Competition, imperfect information, industry structure, market power, mergers.

*We are grateful to two anonymous referees, Jacques Cremer, Jan Eeckhout, William Fuchs, Pavel Gapeev, Gerard Hoberg, Charles Kahn, David Levine, Evgeny Lyandres, Jianjun Miao, Volker Nocke, Mikhail Panos, Tomasz Piskorski, Sergey Popov, Michael Salinger, Yuliy Sannikov, Klaus Schmidt, Andrzej Skrzypacz, Chester Spatt, Ajay Subramanian, Lucy White, Andriy Zapchelnyuk, Mihail Zervos, and seminar participants at HSE in Moscow and the 2012 EFA Annual Meeting for useful comments and suggestions. This research was carried out in part during Bart Taub's stay at ICEF, Higher School of Economics, Moscow, and also at the European University Institute, Florence. He gratefully acknowledges financial support from the Academic Excellence Project 5-100 of the Russian Government and from the Fernand Braudel Fellowship at the European University Institute.

[†]Boston University Questrom Scool of Business, CEPR, and ECGI. Email: dhackbar@bu.edu.

[‡]University of Glasgow. Email: bart.taub@glasgow.ac.uk.

1 Introduction

According to the market power doctrine, the concentration of output among firms in an industry is a measure of market power in that industry. More market power is synonymous with monopoly: prices increase and output falls, to the detriment of consumers and to society at large.

The conventional view is that anticompetitive mergers increase industry concentration and hence increase market power, harm competition ex post, and therefore need to be carefully reviewed and possibly restricted by regulators. Hence, regulators, such as the Antitrust Division of the Department of Justice or the Federal Trade Commission, have the mandate to prevent situations that "excessively" transfer welfare from consumers to firms via buildups of dominant positions or firms with disproportionate market power, including mergers perceived to be anticompetitive.

This paper asks whether these policies are desirable or effective. To answer these questions, we build a dynamic, noisy collusion model that captures firms' optimal output strategies prior to a merger. Our model extends Sannikov's (2007) continuous-time model of tacit collusion, which built on the discrete-time models of Green and Porter (1984) and Abreu, Pearce, and Stacchetti (1986) In these models firms share a market and choose output levels on an ongoing basis. The firms would like to collude but neither firm can observe the actions of the other firm. Instead, they observe price, which is influenced by both firms, but which also is influenced by the noise in demand. As a result, firms cannot directly infer the action of the rival firm, but instead must indirectly infer it.[1]

To cleanly identify the effects of anticompetitive mergers, we abstract away from other common aspects of mergers that can obscure purely anticompetitive effects: these include operational, financial, or other synergies. Operational synergies can stem from higher growth or lower costs: for example, by combining hubs, routes and gate slots, two airlines might be able to operate more efficiently and reduce costs to consumers. Financial synergies can result from tax savings, increased debt capacity, or improved returns: for example, by pooling their portfolios of loans, two banks might better diversify risk and thus be able to offer lower interest rates to mortgage customers. Product mix synergies can improve as the result of a merger to the benefit of consumers. These synergies would bias a model in favor of mergers; by eschewing them we build in a bias against mergers. We thus

[1] In equilibrium, no firm deviates, but it can appear to have deviated, because random demand fluctuations can lower price. To maintain the equilibrium, the firms must nevertheless punish those apparent deviations by increasing their output in response.

focus only on the desire of firms to collude prior to merging or potentially to merge if collusion fails.

The conventional view fails to account for dynamics. Firms in our dynamic model are forward-looking, aware that they are in a dynamic cartel-like situation, but are unable to directly observe the actions of the rival firm, which would enable them to enforce the cartel. The inability of each firm to observe the other firm's output reflects the real world: regulators punish firms that directly track and coordinate with each other's actions for market power purposes.

Because they are blocked from observing each other directly, firms are unable to punish their rivals for directly perceived deviations from collusion, that is, for producing too much in order to realize temporarily higher profits at the expense of the other firm. The inability to directly observe and punish deviations therefore requires a tacit collusion arrangement, in which firms attempt to observe each other indirectly, via prices. This indirect observation is imperfect, however, because prices are affected by random influences, in addition to the effects of the firms' output choices.

Because of the random influences a firm can mistakenly appear to produce too much output, even though neither firm actually commits such an infraction in equilibrium. Under the tacit collusion arrangement this nevertheless triggers a punishment in which the rival firm increases output, thus driving down prices and so harming the firm that has apparently deviated: if continued, there is a price war, resulting in low profits for both firms. It is the fear of this price war that sustains the tacit collusion arrangement in the long run.

The potential to merge weakens those punishments, because it prematurely terminates them under terms that are an improvement over the price war for the firm that is being punished. Instead of the price war, the deviating firm gets a share in the monopoly that the firms form when they merge. Because the potential for punishment is concomitantly reduced, the trepidation about aggressively producing output in contravention of the interests of the cartel arrangement is reduced: there is more competition, resulting in more output and lower prices.

It is well known that weakening punishments weakens cooperation, which in the present context means a weakening of collusion. What is not so obvious is that mergers embody such a weakening, and how to model it; this is our central focus. We reverse the conventional view that mergers are harmful for society: making mergers more difficult (i.e., costlier for the firms) is actually harmful to society, because it strengthens the ability of firms to punish each other and enforce the cartel.

1.1 Related literature

The conclusion that mergers can weaken collusion by reducing the cost of punishment for deviations has previously been drawn in Davidson and Deneckere (1984) who present a model of horizontal mergers with tacit collusion. A related effect has been more recently shown in the context of vertical mergers in Nocke and White (2007) where it is termed the punishment effect.

Davidson and Deneckere's model is significantly different from our model. They posit that the merger consists of a post-merger cartel arrangement among firms that remain distinct after merging; [2] they do not analyze pre-merger play (which would be ill-defined in any case as their model is deterministic), but only whether a cartel is sustainable ex post of its formation. They consider two generic situations. In the first situation, all the firms in the industry form a cartel and evenly split the monopoly profits. In the second situation one of the firms has previously deviated, and so a full trigger-strategy punishment is imposed: in the first part of the paper they assume that the punishment phase is that the firms revert to Cournot-Nash collusion, for which total industry profits are lower than the full-collusion monopoly profits. They then compare the gain if one of the firms, which can be either an outside firm or the merged firm, because it acts as a single firm, deviates, with a subsequent permanent trigger-strategy punishment that is, Cournot-Nash profits. In the "merger" case the merger reduces the number of firms in the industry, and as a result the Cournot-Nash profits increase, but this has the effect of reducing the *relative* difference in profits as the result of the deviation, that is, it constitutes a weakening of the punishment from deviation. [3] We also have a weakening of punishments, but the weakening does not hinge on the number of firms in the industry. Our analysis focuses on *pre*-merger behavior, and the effects of the potential to merge on threats *prior* to the merger.

The paper of Thijssen (2008), like our paper, studies mergers using a continuous-time structure.

[2]Intuitively, think of OPEC with member countries inside the cartel, which Davidson-Deneckere would label a merged entity, and outside non-member countries competing against the cartel.

[3]Miller, Sheu, and Weinberg (2019) develop a more formal dynamic repeated-game model similar to that of Davidson and Deneckere, and calibrate it to data from the beer industry. They find evidence for collusive monopolisation, that is, markups exceeding those that would be expected in a Cournot-Nash equilibrium. They themselves note however that because their model is deterministic, punishments will not be exacted in equilibrium, and so are not directly measurable in the data. They also note that models with noisy observation, punishments might be observable, mentioning the example of Green and Porter (1984).

There are additional empirical studies with evidence for tacit collusion, including Porter (1983), who finds evidence for tacit collusion among railway shipping operators in the U.S. in the 19th century, and Knittel and Stango (2003), who find evidence for tacit collusion by credit card providers.

There are two firms that have potential gains from merging due to the improvement of diversification for the merged firm. Thijssen posits correlation in the profit processes of the firms, leaving optimization of any production function in the background, and so the model does not have any pre-merger or pre-acquisition collusion. The central optimization problem facing Thijssen's firms is when to merge with or acquire the other firm, which, since the model is couched in continuous-time terms, reduces to an optimal stopping problem, which in turn is expressed in terms of a boundary. In an acquisition, one firm hits its relevant boundary first, and then makes an acquisition offer designed to induce acceptance, given that the acquired firm can refuse. In a merger both firms hit their boundary simultaneously, and then bargain over the shares of the merged firm that will be paid to the shareholders of the separate firms. Because the boundaries are endogenous to the game, the model takes on a real option character. In our model, the ability of the firms to optimize pre-merger output on an ongoing basis, with the result that that path to the boundary, which is exogenously given, and also the division of the surplus from the merger, are entirely endogenous, and moreover, simultaneity also emerges endogenously. The potential exists for a firm to refuse to merge in our model as well, but, as in Thijssen's model, this does not happen in equilibrium.

1.2 Technical elements of our model

Our model builds on Sannikov (2007)'s continuous-time model of the tacit-collusion equilibrium of Abreu, Pearce, and Stacchetti (1986), and also on the unobservable action model of Fudenberg, Levine, and Maskin (1994). The continuous-time approach allows him to express the model in geometric form, which is far less tractable in discrete time. We assume that firm outputs are imperfect substitutes, as Sannikov similarly assumed, which we view as adding realism: airlines, for example, typically have different hubs but overlapping routes on which they compete.[4]

Our approach to determining the equilibrium differs from Sannikov's in that we explicitly treat each firm as a principal in an agency construct, treating the other firm as its agent, with the continuation-value of the agent as the state variable for the principal; correspondingly, each firm at the same time behaves as an agent reacting to what is effectively a contract set out by the other

[4] Sannikov and Skrzypacz (2007) for example show in a dynamic model similar in spirit to ours, that if information arrives continuously and firms are able to react quickly to that information, and also if the key assumption that the firms' goods are perfect substitutes is maintained, then collusion breaks down.

firm. The solution maps out marginal rates of substitution that can then be interpreted as prices or shadow prices in an equilibrium in which both firms' contracts are optimal.[5] The shadow prices in the contract internalize the external effect of a firm's change in output on the other firm's profits, enabling the firms to "steer" each other to maintain their ongoing tacit collusion.

The solutions of the agency contracts yield simultaneous differential equations that map out a one-dimensional manifold in the plane defined by the firms' continuation values. This manifold, identified by Sannikov as $\partial\mathcal{E}$, comprises the largest equilibrium set of the game. It is Markovian, i.e., the movement along the manifold depends only on the current state of the continuation values.

This Markovian property means that the equilibrium state can be mapped one-to-one to the equilibrium vector of public information that is driven by the output decisions of the firms and by noise. We begin by treating this information as a state variable for the firms, and we then carry out a transformation or mapping of this state to the space of continuation values using stochastic calculus by way of our agency construction. This then maps out the equilibrium manifold $\partial\mathcal{E}$ via Sannikov's main differential equation (Sannikov (2007), equation (24), p. 1309).

We extend this construction by including additional boundary conditions associated with the merger. Given the dynamic and stochastic nature of the model, the moment the firms merge is a stopping time, and we model each firm as independently choosing this stopping time as the *optimal* stopping time. Because *both* firms must choose the *same* optimal stopping time, there is a complication that goes beyond standard optimal stopping problems: the firms must somehow coordinate their stopping times. Characterizing this coordination problem is our central challenge.

We solve the coordination problem by applying a smooth-pasting condition to each firm's optimal agency problem. In the agency construct, both firms naturally choose the same smooth-pasting point, and hence the same stopping time, thus solving the simultaneity problem. It is the transformation to the agency construct that enables us to express the smooth-pasting condition. All of these elements—the manifold, the optimal stopping problem, the differential equations, and the marginal valuations using the agency construct—are extremely natural within the continuous-time technical framework, and would not have tractable analogues in discrete time.

[5] This construct, also known as the planner approach, has a long intellectual history and is an established solution technique in the dynamic contracting literature; see Miao and Zhang (2015) for a summary of the literature. It also has a long history in macroeconomics; a recent example is Alvarez and Jermann (2001), who reconstitute a growth model with defection constraints as a planner problem in which the partner country can defect from the contract.

1.3 Intuitive elements of the model

To illustrate the pro-competitive effect of mergers, Figure 1 plots the continuation values—the equilibrium discounted expected present value of profits—of two colluding firms. The continuation values are in turn influenced by the output quantities. These outputs are highly persistent, that is, firms never wildly oscillate between extremely high and extremely low output, but rather gradually adjust their outputs in response to the current state as indexed by the current locus of the equilibrium manifold.

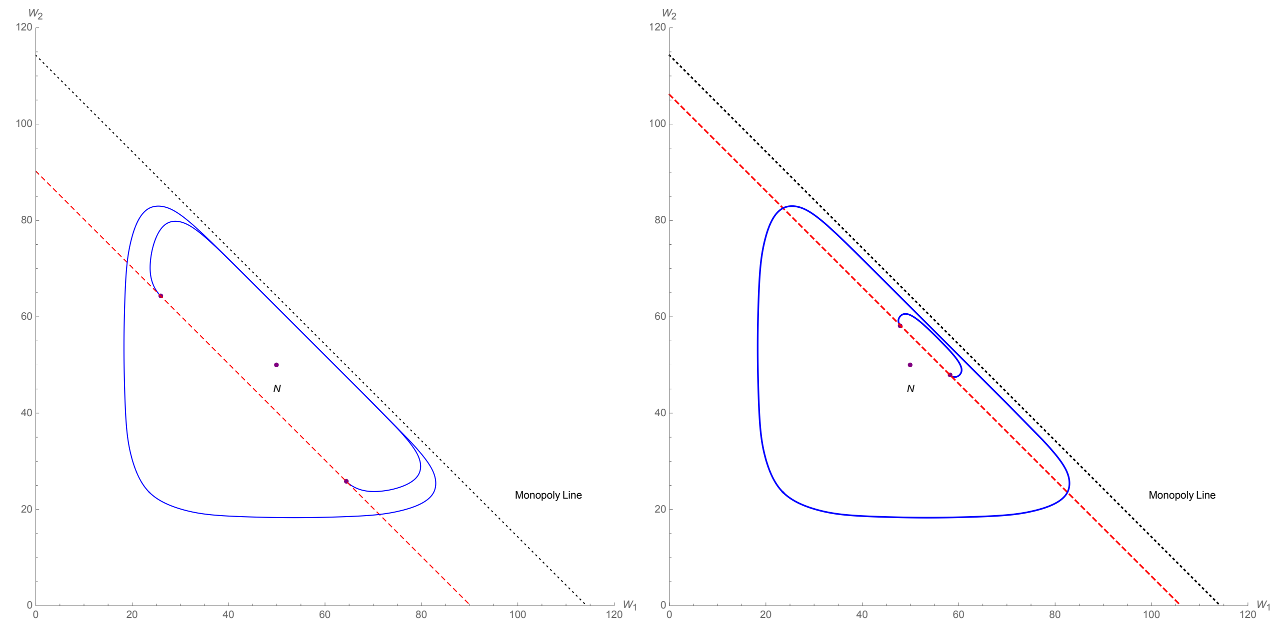
In each panel of Figure 1, the outer manifold is the maximal equilibrium manifold, $\partial\mathcal{E}$, found by Sannikov in his example of tacit collusion. In an equilibrium, at any moment the continuation values of the firms lie on a point of the manifold and are perturbed along the manifold by Brownian shocks. Thus, if one could observe the evolution of the continuation values dynamically, at any given time the continuation values would be located at a point on the manifold, like a bead on a wire, and would appear to jitter as if being jostled by invisible particles, exactly as real particles are jostled by the random motions of molecules in a medium in Brownian motion, but the equilibrium incentive requirements restrict the jostling to move the bead only along the wire.

The one-dimensional character of the manifolds reflects the mechanics of the punishments the firms mete out as the demand shocks perturb prices. The firms must punish *apparent* deviations of output—chiseling—from the agreed duopoly quantity to sustain the equilibrium, even though in equilibrium neither firm has deviated. There is only one tool available to the firms to punish each other: increasing output. As *both* firms are aware of the apparent deviation, the "offending" firm must cooperate in its own punishment in order to get back to the good graces of the other firm. Thus, as the punisher increases output, the offender must decrease output. This results in a movement along the manifold, which prescribes the direction of movement.[6]

If the continuation values lie in the northeast part of the manifold, the firms are colluding; they are producing at reduced rates and are effectively sharing monopoly profits, with some reduction due to the difficulty of coordination due to the noisy perturbations. If the continuation values lie in the southwest part of the diagram, the firms are in a price war. At this point both firms are producing

[6] The key feature of the equilibrium is that it is *optimal* for the firms to do this despite its self-destructive nature. In non-stochastic repeated games it is also optimal to punish oneself as well as the rival player, but in equilibrium this never happens—it is only a threat. Here, interestingly, the punishments do occasionally take place due to the noise.

Figure 1. Merger and no-merger equilibrium manifolds



This figure plots the merger equilibrium manifold from Figure 3 and the no-merger equilibrium manifold from Figure 2; the right hand panel has a plot similar to Figure 3 for the case when the cost of merging is small. The merger equilibrium manifold is contained entirely within the no-merger equilibrium manifold. The value in the collusive region of the merger manifold is therefore below the value in the collusive region of the no-merger manifold, expressing the reduced punishments and reduced competition of the merger manifold.

close to competitive amounts, driving their current profits down. Because the high output state is long lasting, the continuation values integrate over the resulting long-lasting low profits, whilst discounting limits the impact of the eventual reversion of the firms to the low-output, collusive state.

In each panel of Figure 1, the inner manifold depicts a merger equilibrium. When the firms merge they share monopoly profits, with the shares determined endogenously by the locus at which the manifold intersects the line depicting the monopoly profits attained by merging, less the cost of merging; in the right hand panel this cost is smaller.

The merger manifold satisfies the same differential equation as the no-merger manifold, but due to the boundary conditions the *entire* manifold is affected. Thus, the merger equilibrium manifold lies entirely inside the no-merger equilibrium manifold, and in the collusive region—the northeast part of the equilibrium manifold—the merger manifold lies to the southwest of the no-merger manifold, with this difference between the manifolds more clearly visible in the right hand panel of the figure. This southwest movement expresses the reduction of the firms' long run profits associated with the merger manifold. We will demonstrate that the collusive region is highly stable, so the figure illustrates how collusion is weakened by the potential to merge.

The rest of the paper is organized as follows. Section 2 outlines the model, and Section 3 outlines its solution (our main theoretical result). Section 4 derives the model's implications. Section 5 concludes. The appendices contain derivations, proofs and other technical results.

2 The model

Two firms compete in an industry with differentiated products that are imperfect substitutes by continuously taking private actions, namely by choosing output levels. Airlines, for example, typically have different hubs but overlapping routes and correspondingly different intensities of imperfect product market competition (Azar, Schmalz, and Tecu, 2018).

2.1 Actions, prices, information and payoffs

Each firm $i = 1, 2$ continuously chooses an action—that is, an output level—$A_t^i \in \mathcal{A}_i \subset \Re_+$ for all $t \in [0,\ \infty)$.[7] The firms observe the history of a vector of public price signals (i.e., price increments) dP_t, which, because the products are imperfect substitutes, depend on the actions of both firms. The instantaneous prices of firms 1 and 2 before and after the merger are given by the levels of the processes:[8]

$$dP_t^1 \ = \ \left(\Pi_1 \ - \ \beta_1 A_t^1 \ - \ \delta_1 A_t^2\right) dt \ + \ d\zeta_t^1 \tag{1}$$

and

$$dP_t^2 \ = \ \left(\Pi_2 \ - \ \delta_2 A_t^1 \ - \ \beta_2 A_t^2\right) dt \ + \ d\zeta_t^2 \tag{2}$$

[7]We depart from Sannikov's assumption that the action set is discrete and finite; we do however, assume boundedness. The boundedness assumption is anodyne in the sense that there is a finite output that maximizes the punishment of the rival firm by effectively minimizing the excess demand facing the rival.

[8] For intuition suppose that price is a deterministic process, with increments $dP = P(t)dt$. In a continuous time stochastic setting we add a stochastic process to the process dP and we are restricted to writing the process as $dP(t)$.

where ζ_t^1 and ζ_t^2 are correlated Brownian demand shock processes. Because we analyze collusion, we focus only on the case in which the β_i and δ_i are positive constants, reflecting that the goods in each market are substitutes in the rival's market. While the potential exists for the goods to be complements ($\delta_i < 0$) we don't examine this case.

2.1.1 Pre-merger information

Before the merger the firms cannot observe the rival's actions directly; they can learn about each other's actions indirectly by observing prices. The noise processes in the model have been constructed *a priori* so that, using the linearity of the prices in equations (1)–(2), the information processes can be isolated from those observations and expressed as a continuous process with independent and identically distributed increments. Thus, before the merger, firms observe a vector of signals X_t by inverting the price process vector:

$$dX_t^1 \;=\; \frac{(A_t^1 + A_t^2)}{2\,(\Pi_2 \,-\, \delta_2 A_t^1 \,-\, \beta_2 A_t^2)} dP_t^2 \,-\, \frac{(A_t^1 - A_t^2)}{2\,(\Pi_1 \,-\, \beta_1 A_t^1 \,-\, \delta_1 A_t^2)} dP_t^1 \;=\; A_t^1\, dt \,+\, \sigma_1\, dZ_t^1\,, \tag{3}$$

and

$$dX_t^2 \;=\; \frac{(A_t^1 + A_t^2)}{2\,(\Pi_2 \,-\, \delta_2 A_t^1 \,-\, \beta_2 A_t^2)}\, dP_t^1 \,-\, \frac{(A_t^2 - A_t^1)}{2\,(\Pi_1 \,-\, \beta_1 A_t^1 \,-\, \delta_1 A_t^2)}\, dP_t^2 \;=\; A_t^2\, dt \,+\, \sigma_2\, dZ_t^2\,, \tag{4}$$

where Z_t consists of two independent Brownian motions Z_t^1 and Z_t^2; the ζ_t^1 and ζ_t^2 processes are thus generated by reversing the inversion implicit in equations (3) and (4). We provide the details of the inversion, which is a straightforward matrix algebra operation, in Appendix A. The state space information vector Ω is thus characterized by all possible paths of X_t, and the public information filtration $\mathcal{F}_t$ is generated by X_t.

2.1.2 Payoffs

The instantaneous payoff functions are the product of output and price increments $A^i\, dP^i$ for $i = 1, 2$. The expected incremental payoffs of the firms are:

$$g_1\left(A^1, A^2\right) dt \;=\; E\left[A^1\, dP_t^1\right] \;=\; A^1 \left(\Pi_1 \,-\, \beta_1 A^1 \,-\, \delta_1 A^2\right)\, dt, \tag{5}$$

and similarly for firm 2. Discounted profits are integrals of these instantaneous profits. We also note for future reference that the functions g_i are by construction continuous and twice continuously differentiable.

If the action profiles A_t^i are measurable with respect to the public information filtration $\mathcal{F}_t$ and square-integrable, that is, $E\int_0^\infty e^{-rt}\left|A_t^i\right|^2 dt < \infty$, then firm 1's *expected* profit will include the expected value of the stochastic integral

$$E\left[\int_t^\infty e^{-r(s-t)} A_s^1 d\zeta_s^1\right]$$

The expected value of the stochastic integral is zero, [9] and only the drift terms survive in the profit calculation. Note that this conclusion holds even though the ζ_t^i processes are driven by the actions A_t^i (see Appendix A), as they are still linear in the underlying Z_t^i processes.

2.2 Post-merger monopoly

We begin our analysis by examining the relatively simple post-merger problem, which will then help us to derive the boundaries that apply in the more complicated pre-merger game.

When firms merge, they jointly control and observe the actions A^i and can therefore observe the demand shocks ζ^i directly. This eliminates their information problem and they can then choose monopoly outputs and share the monopoly profits available by acting as a single firm. A straightforward calculation determines the optimal monopoly outputs:

$$A^{1\star} \;=\; \frac{(\delta_1+\delta_2)\,\Pi_2 \,-\, 2\,\beta_2\,\Pi_1}{(\delta_1+\delta_2)^2 \,-\, 4\,\beta_1\,\beta_2}\,, \qquad A^{2\star} \;=\; \frac{(\delta_1+\delta_2)\,\Pi_1 \,-\, 2\,\beta_1\,\Pi_2}{(\delta_1+\delta_2)^2 \,-\, 4\,\beta_1\,\beta_2}\,. \tag{6}$$

These actions are square-integrable, so only the drift terms of the payoff functions survive when taking the expected value of the monopoly discounted profit, as the stochastic integrals have expectation zero. The resulting flow payoff function starting from the merger time $\mathcal{T}_m$ is then the

[9] The square integrability condition ensures that the stochastic integral is a martingale; intuitively, the variance of the stochastic integral is the integral of the square of the integrand and square integrability guarantees that this integral is finite; see Björk (2009), chapter 4.

static monopoly profit

$$\pi_m \;=\; r \int_{\mathcal{T}_m}^{\infty} e^{-r(s-t)} \sum_{i=1}^{2} g_i\left(A^{1\star}, A^{2\star}\right) ds \;=\; \frac{(\delta_1+\delta_2)\,\Pi_1\,\Pi_2 \;-\; \beta_1\,\Pi_2^2 \;-\; \beta_2\,\Pi_1^2}{(\delta_1+\delta_2)^2 \;-\; 4\,\beta_1\,\beta_2}\,, \tag{7}$$

which will be shared between the firms; we elaborate on the sharing rule below.[10]

2.2.1 The merger

A merger occurs when the firms simultaneously decide to merge at an equilibrium stopping time, $\mathcal{T}_m$, via a publicly observable signal S_t^i from the set {"do not merge," "agree to merge with share of the net monopoly profit $\xi^i_{\mathcal{T}_m}$"} such that $\xi^i_{\mathcal{T}_m} + \xi^{-i}_{\mathcal{T}_m} = 1$.

The merger entails fixed costs: these can include substantial legal fees that are necessary to obtain regulatory approval, due diligence measures, the generation of asset valuations, investment bank fees, and so on. As a practical matter, costly post-merger physical changes can be necessary as well: when two airlines merge, one of the fleets will need to be repainted. We express these costs as the one-time merger cost k, which is subtracted from the discounted monopoly profit resulting from the merger. Firm i's profit from the merger is therefore

$$\xi^i_{\mathcal{T}_m}(\pi_m - k)\,. \tag{8}$$

The potential ongoing profit share combinations, given by $\xi^i_{\mathcal{T}_m}(\pi_m - k)$ and $(1-\xi^i_{\mathcal{T}_m})(\pi_m - k)$, thus map out a line in value space; this is the red dashed "merger" line depicted in Figure 1.

We emphasize that the sharing rules ξ^i are endogenous and must be determined in equilibrium; we do not impose an *a priori* sharing rule such as equal shares. The stopping time, $\mathcal{T}_m$ must be optimal from the perspective of each firm. Because it is an *optimal* stopping time there really are *two* stopping times, ${\mathcal{T}_m}^1$, and ${\mathcal{T}_m}^2$, one for each firm; it is then a requirement of the equilibrium that the two stopping times be equal. As with the shares ξ^i, we emphasize that the stopping times are endogenous, and we do not impose equality of the stopping times ex ante; in theory, as with any model, there is the potential that an equilibrium with equal stopping times does not exist.

We demonstrate that equal stopping times are in fact possible and natural because at the

[10]The integral is discounted profit; pre-multiplying by r converts it into a flow, hence our designation of flow payoff.

moment of the merger both firms agree on the marginal value of the merger. We express this explicitly using a smooth pasting argument, but that smooth pasting argument depends on a reformulation of the model in terms of agency, and we discuss this reformulation below.

In contrast to the output decisions of the firms, at the moment that the smooth pasting conditions determining the merger are satisfied, the announcement of the willingness to merge is publicly observable. A firm might then decide to publicly refuse to merge in an attempt to extract surplus from the other firm. The jilted firm might then respond with a punishment for this refusal. For such a punishment path to work, however, it must itself be an equilibrium: both firms must agree to follow that path. We discuss the possibility of such alternative paths in Appendix K and also in Appendix L. We show that the refusal punishment paths would be suboptimal for both firms and hence cannot support equilibria in which a firm refuses to merge. For this reason we focus only on the merger equilibrium.[11]

2.2.2 Pre-merger payoffs

We assume that A_t^i is square-integrable; similar to Sannikov (2007), p. 1316, the pre-merger discounted expected payoff at time t is given by:[12]

$$r\,E\left[\int_t^{\mathcal{T}_m} e^{-r(s-t)} A_t^i dP_t^i \Big| \mathcal{F}_t\right] \tag{9}$$

The stochastic integral drops out, yielding the expected payoff

$$r\,E\left[\int_t^{\mathcal{T}_m} e^{-r(s-t)} A_t^i \left(\Pi_i \,-\, \beta_i A_t^i \,-\, \delta_i A_t^{-i}\right) dt \Big| \mathcal{F}_t\right] \tag{10}$$

2.2.3 Pre-merger continuation value

With the basic structure of the merger in hand we can state the objective of the firms prior to the merger. Define the pre-merger continuation value $W_t^i(\cdot)$ as the mapping, $W^i : \Re^2 \to \Re_+$, from the

[11]We emphasize that the nature of the punishment paths associated with the potential refusal of a merger that we address in the appendix are sharply distinct from the punishments underlying the mechanics of the pre-merger equilibrium and that are our focus going forward. In the ongoing product market competition game with noisy observations, pre-merger punishments are effected by changes in *output* in response to movements in *prices*.

[12] Note that we discount from time t, thus, the integral is the continuation value as of each time t.

set of state vectors $X_t \in \Re^2$, to firm i's time t payoff in the continuous-time game.

$$W^i(X_t) = \sup_{A^i \in H^2, \mathcal{T}_m^i, \xi^i_{\mathcal{T}_m}} \mathbb{E}\left[r \int_t^{\mathcal{T}_m^i} e^{-r(s-t)} A_s^i \left(\Pi_i - \beta_i A_s^i - \delta_i A_s^{-i}\right) ds + e^{-r(\mathcal{T}_m^i - t)} \xi^i_{\mathcal{T}_m^i} (\pi_m - k) \middle| \mathcal{F}_t, A^{-i} \right], \tag{11}$$

where H^2 is the space of square-integrable real-valued functions on the action set $\mathcal{A}_i \subset \Re_+$.

2.3 The formal problem and definition of equilibrium

The problem we consider is that of finding the maximal set of payoffs attainable in equilibrium in the repeated game between the two firms, subject to the constraint that players' continuation values can never fall below the merger line. This is because continuing to play the collusion equilibrium nets the firms more profit than merging due to the fixed cost of merging, insofar as every equilibrium point above the merger line dominates at least part of the merger line.

At the merger line, the continuation values are by definition equal in the merger and no-merger states and so the merging does not affect the instantaneous outcome. However the *marginal* impact of merging must also be accounted for, and this is expressed as the requirement that the *shares* garnered by each firm at the moment of the merger must be locally optimal for each, conditional on the other firm's strategy. More formally, we can define the game as follows.

Definition 1 *A duopoly Markov merger game is a repeated game with two firms $i \in \{1,2\}$ and stage game for every $t \in [0,\infty)$ that is a tuple*

$$\{(\mathcal{A}_i)_{i\in\{1,2\}}, (g_i)_{i\in\{1,2\}}, (P_t^i)_{i\in\{1,2\}}\},$$

where $(\mathcal{A}_i)_{i\in\{1,2\}}$, is the space of square-integrable action functions of player i, g_i is the instantaneous payoff of player i from (5), and $(P_t^i)_{i\in\{1,2\}}$ is the price process equations (1)–(2) adapted to the filtration $\mathcal{F}$ generated by (ζ_t^1, ζ_t^2) resulting in public information histories as determined by (3)–(4), with discounted expected payoffs in (9), and in addition a tuple

$$\{\mathcal{T}_m^1, \mathcal{T}_m^2, \xi^1, \xi^2, S^1, S^2\},$$

such that at any moment t, firms choose a publicly observable signal S_t^i from the set {"do not *merge," "agree to merge"*} *with shares* $\{\xi_t^i, 1-\xi_t^i\}$ *and where* $\mathcal{T}_m^1$ *and* $\mathcal{T}_m^2$ *are the stopping times defined as the first time the signal* $S^i =$ *"merge" is chosen by firm* i*.*

Thus, a Markov merger game is similar to Sannikov's game in the run-up to the merger, during which time the firms can be thought of as sending the "do not merge" signal (or at least one of them). At the moment of the merger, they both send the "merge" signal along with the choice of the sharing rule, and the merger takes place, and is irreversible.

We turn now to the definition of equilibrium, which is simply Sannikov's definition expanded to encompass the merger stopping time.

Definition 2 *A Markov merger game equilibrium consists of:*

- *a profile of public strategies* $A = (A^1, A^2)$ *such that* A^i *maximizes the expected discounted payoff of player* i *prior to the merger given the strategy* A^{-i} *of his opponent after all public histories. [Sannikov (2007) p. 1292], and in addition,*
- *firms merge only if both firms simultaneously play "agree to merge" with shares* $(\xi^1_{\mathcal{T}_m}, \xi^2_{\mathcal{T}_m})$ *such that* $\xi^2_{\mathcal{T}_m} + \xi^1_{\mathcal{T}_m} = 1$*;*
- *the stopping times* $\mathcal{T}_m^1$ *and* $\mathcal{T}_m^2$ *are optimal for firm 1 and firm 2 respectively;*
- *the stopping times are identical, that is* $\mathcal{T}_m^1 = \mathcal{T}_m^2 = \mathcal{T}_m$*;*
- *merging does not Pareto-dominate continuation prior to the merger.*

It is key that, combined with knowledge of the public history X_t, means that the player i knows the rival's recommended action A_t^{-i}, which is validated in equilibrium. Moreover, we note that, as with Fudenberg, Levine, and Maskin (1994), our equilibrium concept is a version of subgame perfection in that strategies are conditional on histories at every time t; we refer to it as Markovian due to the existence of a state process, but this requires additional exposition that we provide below.

While the largest equilibrium set is driven by the following sharing rule along the merger line,

$$W^1_{\mathcal{T}_m} + W^2_{\mathcal{T}_m} = (\pi_m - k) \tag{12}$$

it is unclear how to ensure that, in equilibrium, the firms agree to merge *simultaneously*; our definition does *not* impose this simultaneity. Our solution procedure transforms the model to characterize optimality, as is commonly the case with optimal stopping problems, by a smooth-pasting condition at the potential merger time that holds for both firms. It is then straightforward to show that the smooth-pasting condition is satisfied simultaneously, resulting in simultaneity of the stopping times.

The final item in the definition, concerning Pareto optimality, relates to the special character of the moment of the merger. Specifically, the decision to merge is public, with both firms undertaking the decision to merge simultaneously and at the same moment ceasing their private production decisions. One might ask on general principle, why don't firms simply merge at the start of the game? The answer to this question is evident from examining Figure 1: starting from any point on the equilibrium manifold that terminates in the merger, merging is not Pareto improving.

This is not true, however, for the no-merger manifold: it extends below the merger line and from any such lower point there is a selection of points on the merger line that can be attained by agreeing on a point and merging, with the merger target point somehow determined by a bargaining solution, such as the one in Thijssen (2008). This in turn destroys the equilibrium character of the manifold, because the continuation values on the manifold must discount the merger target point and thus must terminate on that point, contradicting the fact that the merger entails a jump to that point. We expand further on this point in Appendix L.

Thus, any equilibrium must continuously approach the merger line. This in turn requires that the smooth pasting condition be satisfied. We expand on these remarks in Appendix K.

3 Solution

This section discusses the solution to the dynamic game using stochastic calculus, in two stages. We motivate our approach by first stating an equivalent "planner" or agency problem.[13]

$$W^i\left(X_t\right) \;=\; \sup_{A_t^i,\mathcal{T}_m^i} \;\mathbb{E}\left[r\int_t^{\mathcal{T}_m^i} e^{-r(s-t)} g_i\left(A_s^i, A_s^{-i}\right) ds + e^{-r(\mathcal{T}_m^i-t)}\xi^i_{\mathcal{T}_m^i}\left(\pi_m-k\right) \middle| \mathcal{F}_t, A^{-i}\right], \qquad (13)$$

[13] See footnote 5.

subject to

$$w^{-i} = \sup_{A_t^{-i}} \mathbb{E}\left[r\int_t^{\mathcal{T}_m^i} e^{-r(s-t)} g_i\left(A_s^{-i}, A_s^i\right) ds + e^{-r(\mathcal{T}_m^i - t)}(1-\xi^i_{\mathcal{T}_m^i})(\pi_m - k) \middle| \mathcal{F}_t, A^i \right], \tag{14}$$

where w^{-i} is the promised utility to the rival firm at time t. Notice that the stopping time is chosen by the "principal," firm i. There is a symmetric problem for the rival firm $-i$. [14]

One could conventionally solve the principal's problem using stochastic calculus: state the Hamilton-Jacobi-Bellman equation associated with the objective (13) with the appropriate state and solve. What is not so obvious is how to express the rival firm's promised utility constraint (14) in stochastic calculus terms.

We take a two-stage approach. In the first stage we solve the maximization problem of the rival as expressed in (14) using stochastic calculus, treating the public information process X_t as the state variable for that problem, taking the rival's action profile as given. The solution of the rival's HJB equation can then be substituted into the Ito expansion of the rival's continuation value process dW_t^{-i}; this process then becomes the *state* process for the "principal," firm i in the second stage of the solution process.

The details of the two-stage solution procedure are as follows.

Stage 1

(i) Solve the conventional profit maximization problem in (14) for each firm, taking the other firm's action profile as given, using the public signal vector X_t as the state vector, thus satisfying *incentive compatibility*;

(ii) use the solution of the firm optimization problem to state the continuation-value process for each firm, W_t^i, in terms of the state vector X_t, that is, *promise-keeping*;

(iii) demonstrate that simultaneous promise-keeping implies a singular volatility matrix, or *enforcement*, restricting the structure of the continuation-value process vector W_t locally to a one-dimensional manifold in the space of processes adapted to $\mathcal{F}_t$.

The singularity of the volatility matrix is important in two senses. The first is that the con-

[14] We note that formally, both the action profile A^i is a *process* conditional on the public information history, and also that the stopping time $\mathcal{T}_{m^i}$ is itself a process, also conditional on the history, and that the optimization is over these processes. As a practical matter however, the solution of the model ends up using optimal control methods that pin these processes down at each time t and determine the optimal stopping time via smooth pasting.

tinuation values are restricted to a *locally* unique direction by the requirement of incentive compatibility. When a firm punishes the other firm due to a movement in prices, the other firm must agree to the punishment with a synchronous output adjustment. The result is that the movement of each firm's continuation value is restricted to movement in a single dimension, even though the continuation values occupy a two-dimensional plane. More concretely, a movement of one of the shock processes, say dZ_t^1, affects *both* continuation values via the volatility matrix as is evident in equation (B.15), but in a coordinated way due to the singularity of the matrix. The singularity of the matrix thus has the *mathematical* effect of reducing the dimension of movement of the continuation values locally to a line, and this line is eventually explicitly characterized by a differential equation. It also expresses the *economic* effect of requiring incentive compatibility in the equilibrium.

The second important sense is that it allows us to undertake the later transformation to the agency formulation because there is then a one-to-one mapping from the state vector X_t to the continuation value vector W_t that is essential for our agency construction.

Stage 2 (i) Using the single-dimensionality of the enforcement manifold, implicitly map the state vector X_t into the continuation-value vector W_t using calculus arguments, so that a firm's continuation value process W_t^i is implicitly expressed as a function of the rival firm's continuation value W_t^{-i}, i.e. as well as its own continuation value W_t^i, implicitly construct a mapping $\mathcal{M} : X_t \mapsto W_t$ and noting that $\mathcal{M}$ is invertible;

(ii) pose the profit maximization problem for each firm as an agency problem with the rival firm's transformed continuation value as the state process;

(iii) solve the principal's optimal stopping problem using value-matching and smooth-pasting conditions and verify that an optimum is attained;

(iv) characterize the manifold stemming from the main differential equation implied by the simultaneous solution of the principal's problem for both firms, and also the simultaneity of merger decisions;

(v) verify the equilibrium by noting that the inverse mapping $\mathcal{M}^{-1}$ implies that the firms' actions are optimal.

The ability to express each firm's continuation value as a function of the rival's continuation value, that is, the agency construction, is key. In that case, the smooth pasting condition that expresses the marginal value of stopping, is a derivative of the continuation value with respect to the rival firm's continuation value. Because the rival firm must similarly satisfy such a condition, the two smooth pasting conditions are guaranteed to be exactly inversely related. This is what enables us to establish simultaneity.

We provide the technical details of these steps in Appendix B.

3.1 Converting the main differential equation into geometric form

The nonlinearity of the model forces us to resort to numerical solutions. Like Sannikov (2007), we adopt a reformulation of the second-stage optimized Bellman equations in polar coordinates to facilitate the computation of numerical solutions in the next section. The details of these derivations, which were not provided by Sannikov, are presented in Appendix E.

We solve the resulting ordinary differential equation system (E.43) numerically to determine the (benchmark) equilibrium manifold, $\partial\mathcal{E}(r)$. If merging is possible, we solve for $\partial\mathcal{E}(r)$ subject to the boundary conditions (i.e., value-matching and smooth-pasting in equations (B.23) and (B.28)), which are not present in Sannikov (2007). These boundary conditions for the merger will have nontrivial effects on the firms' pre-merger strategies and values, which we study in the next section.

3.2 The impact of the merger on the equilibrium set

We next prove that the equilibrium manifold shrinks with the merger cost K, i.e., that the potential to merge results in a boundary condition that translates into reduced punishments and therefore reduced collusion. Denoting the boundary of the equilibrium manifold $\partial\mathcal{E}_{\text{M}}^{K}$ for the merger model, and $\partial\mathcal{E}_{\text{NM}}$ for the no-merger model, we obtain the following result (the proof is in Appendix I).

Proposition 1 *The Markov merger equilibrium set is strictly contained inside the no-merger Markov equilibrium set and shrinks with the merger cost K, that is, for $K' > K$,*

$$\partial\mathcal{E}_{M}^{K} \subset \partial\mathcal{E}_{M}^{K'} \subset \partial\mathcal{E}_{NM} \tag{15}$$

The result is illustrated most clearly in Figure 1: the fixed cost of merging is lower in the right hand panel, and it is apparent that the merger manifold shrinks relative to the high-cost case, with the key observation that in the low-cost case in the right hand panel, the merger manifold has moved away from the no-merger manifold.

The fact that the continuation values of both firms in the no-merger equilibrium manifold exceed the continuation valuations in the collusion region in the merger equilibrium might lead firms to want to conclude an agreement to never merge. This cannot be an equilibrium.

Corollary 1 *Never merging is not an equilibrium.*

Proof: In the no-merger equilibrium the price war region of the equilibrium manifold extends below the merger line. If the firms attain the price war state then there exists a point on the merger line entailing monopoly profit shares that make both firms better off relative to the price war. (See Appendix K and Appendix L for further discussion.) □

4 Implications

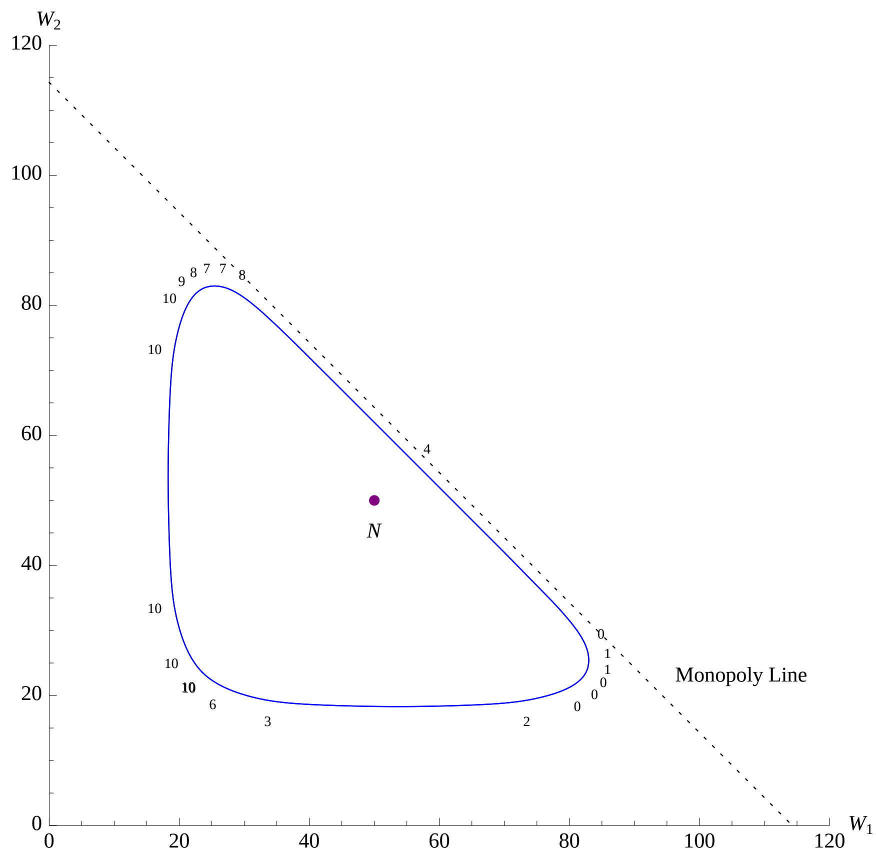
4.1 The no-merger equilibrium

To begin our numerical analysis, we solve for the equilibrium of the benchmark model without mergers. The benchmark model's solution to the differential equation (E.43) is characterized by an equilibrium set, $\partial\mathcal{E}(r)$, that forms a manifold in the space of continuation values, (W^1, W^2), as seen in Figure 2. We assume a baseline environment with symmetric demand functions and the following parameter values: $\Pi_1 = 30$, $\Pi_2 = 30$, $\beta_1 = 2$, $\beta_2 = 2$, $\delta_1 = 2$, $\delta_2 = 2$, $\sigma_1 = 1$, $\sigma_2 = 1$, and $r = 1$.[15] The static Nash equilibrium in the duopoly stage game is $(5, 5)$ in the baseline environment. This generates continuation values of $\pi_{d,i} = 50$ for each firm $i = 1, 2$ (or 100 for both firms).

Figure 2 displays the equilibrium manifold in the benchmark case without mergers along with firm 2's output choices. Due to the symmetric demand functions, we can rotate firm 2's output choices around the 45 degree line to obtain firm 1's output choices. In the northeast stretch of the

[15] Our symmetric example differs from Sannikov's (2007) asymmetric example, the parameter values of which are $\Pi_1 = 25$, $\Pi_2 = 30$, $\beta_1 = 2$, $\beta_2 = 2$, $\delta_1 = 1$, $\delta_2 = 2$, $\sigma_1 = 1$, $\sigma_2 = 1$, and $r = 1.5$. While the baseline parameter choices could be motivated in detail, we omit this for the sake of brevity. The model's results and implications vary quantitatively but not qualitatively with parameters.

Figure 2. No-merger equilibrium manifold and outputs



This figure plots the no-merger equilibrium manifold (blue, solid line). Firm 2's output choices are outside the equilibrium manifold. Due to the symmetric demand functions, we can rotate firm 2's output choices around the 45 degree line to obtain firm 1's output choices. The static Nash equilibrium's output choices of (5,5) are depicted by N in terms of the continuation values of $(50, 50)$. We use the baseline environment in which $\Pi_1 = 30$, $\Pi_2 = 30$, $\beta_1 = 2$, $\beta_2 = 2$, $\delta_1 = 2$, $\delta_2 = 2$, $\sigma_1 = 1$, $\sigma_2 = 1$, and $r = 1$.

equilibrium manifold the firms cooperate, with output levels around 4 and hence are collusive. On the opposite side of the equilibrium manifold, that is, in the southwest stretch, they engage in a price war, in which output levels are around 10 and hence non-collusive.

In the upper right region of $\partial\mathcal{E}(r)$ in Figure 2, output is low for both firms in that it sums up to about 8. That is, the firms approach the monopoly output, which is, according to equations (6), $a_i^\star = 3.75$ for each of the two firms (or 7.5 for both firms). Their continuation values are consequently higher, (around (56,56) at the midpoint of the market-sharing region) than in the static duopoly's Nash equilibrium, which corresponds in terms of continuation values to $(50, 50)$, depicted by N in the figure. The monopoly value of the two firms is $\pi_m = 112.5$ and is depicted by the (dotted) *monopoly line* for all (feasible) sharing rules in the unit interval. Clearly, the monopoly value is unattainable in either the dynamic or the static duopoly. It is evident that, in this region, when a firm's continuation value increases, its market share also increases. Therefore, firms are tempted to overproduce, moving away from the center of the market-sharing region.

In the upper left segment of $\partial\mathcal{E}(r)$ in Figure 2, firm 2 obtains the maximal continuation value of almost 83, while the continuation value of firm 1 equals about 25. At that point, firm 1 underproduces, while firm 2 overproduces relative to the duopoly and monopoly quantities. Output is asymmetric: at the lower right, for example, firm 1's output is high (i.e., 10) and firm 2's is low (i.e., 0). In the lower right segment of the equilibrium manifold, firms thus display similar strategies with the roles of firm 1 and 2 reversed, namely with firm 1 the incumbent and firm 2 as the entrant.

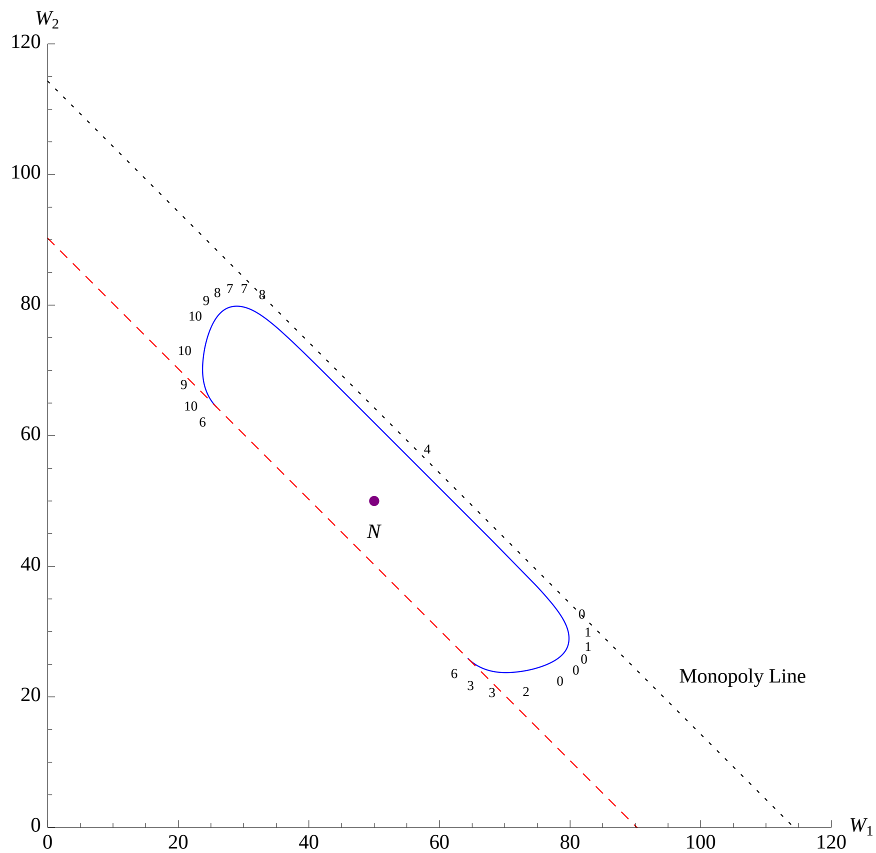
At the intersection of $\partial\mathcal{E}(r)$ with the 45 degree line, firms engage in a price war in that both firms aggressively overproduce. Their output levels are (10,10) and substantially exceed the (static) duopoly outputs of (5,5), which leads continuation values to drop well below (25,25).

4.2 The merger equilibrium

We continue our analysis by solving the differential equation (E.43) for the equilibrium manifold, $\partial\mathcal{E}(r)$, with mergers by incorporating the boundary conditions for value-matching and smooth-pasting in equations (B.23) and (B.28). When the merger occurs, both firms share (net of the merger cost) the value of the resulting monopoly stage game without imperfect information. This yields the *merger line* in equation (B.26), which corresponds to the monopoly line, π_m, minus the cost of merging, k, and is represented by the red, dashed line in the figure for all feasible sharing rules in the unit interval. If firm 1, for example, captures more of the merger gains, then the merger point will more likely lie on the lower right section of the line; conversely, if firm 2 captures more of the gains, then the merger point is more likely to be on the upper left section of the line.

Figure 3 illustrates how firms anticipate the impending merger. As we demonstrated in Proposition 1 and in our discussion of Figure 1 in the introduction, the equilibrium manifold with mergers is entirely contained inside the original no-merger equilibrium manifold in Figure 2. This means that some of the collusion profits attainable in the no-merger equilibrium are not attainable in the merger equilibrium, while some of the non-collusion costs (due to potential price wars) are

Figure 3. Merger equilibrium manifold and outputs



This figure plots the merger equilibrium manifold (blue, solid line), the monopoly line (black, dotted line), and the merger line (red, dashed line) for a merger cost of $k = 24$. Firm 2's output choices are outside the equilibrium manifold. Due to the symmetric demand functions, we can rotate firm 2's output choices around the 45 degree line to obtain firm 1's output choices. The static Nash equilibrium's output choices of (5,5) are depicted by N in terms of the continuation values of $(50, 50)$. We use the baseline environment in which $\Pi_1 = 30$, $\Pi_2 = 30$, $\beta_1 = 2$, $\beta_2 = 2$, $\delta_1 = 2$, $\delta_2 = 2$, $\sigma_1 = 1$, $\sigma_2 = 1$, and $r = 1$.

avoided in the merger equilibrium. Intuitively, this stems from the weaker punishments inherent in the equilibrium with mergers. The punishments are weaker because the opportunity to merge eliminates the severe punishments in the price-war regime.

The (upper right) market-sharing region of the equilibrium manifold, which now reflects the possibility of a merger, is slightly less stretched out in Figure 3 compared with Figure 2. The firms trade off being the production leader in this noisy duopoly against being punished for deviating. As in the no-merger equilibrium, total output stays low at around 8, which is again close to the optimal monopoly output of 7.5. In other words, the market-sharing regime, in which firms' optimal output levels are highly collusive, is similarly large relative to one without an anti-competitive merger, as in the previous figure. Moreover, as the entrant and incumbent regime is approached at the upper left region, total output increases to 10 and finally to 12 and 13 in the contestability region. But then total output declines slightly again to 12 just before the merger line is smoothly pasted to the equilibrium manifold. Thus, compared with the previous figure's no-merger equilibrium manifold, total output tends to be lower in the worst stages of the dynamic game.

In practice, merger gains are often split asymmetrically between the merging firms. The model predicts this: the firm that is being punished in the contestability region gets a smaller share of the merged entity's value, because it has a smaller continuation value and hence it appears to be taken over by the overproducing firm that has a larger continuation value in the contestability region. It is therefore reasonable to designate them target and acquirer. The firm that overproduces at the right time will be rewarded by the larger share in the merged entity if the merger boundary is reached. This asymmetry is not driven by any inherent asymmetry in the demand functions, noise parameters, or other parameters, which are all symmetric: it is driven solely by the state of product market competition that the firms have attained as a result of cumulative play of the noisy duopoly game.

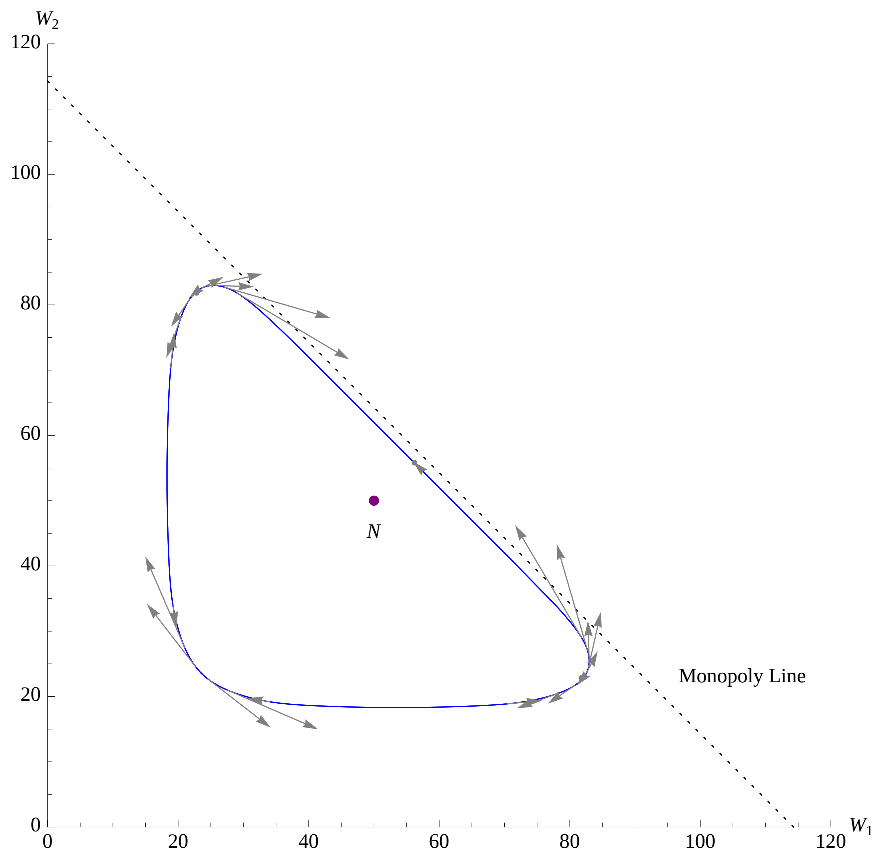
4.3 Collusion and the dearth of mergers

We next examine the stability of the no-merger and merger equilibria.[16] The arrows in Figures 4 and 5 provide information about the stability of the regions. The length of each arrow, which corresponds to the volatility-scaled drift of the continuation value process represented in equation (B.15), indicates the strength of stability. The direction of the arrows conveys information about the local stability of each region.

We can analyze these stability properties using the agency approach. Because we have expressed the equilibrium in terms of continuation values, the equilibrium manifold expresses the trade-off between the continuation value of the principal and the agent from the perspective of each firm in the implicit agency contract. The slope of the manifold is the shadow price or marginal value of increasing the continuation value of the rival. One can therefore view the continuation value of the rival as an "asset" that can be spent or saved at this price.

[16] Sannikov (2007) studies stability for the partnership example (see his Figure 2), but not for the duopoly example.

Figure 4. No-Merger equilibrium manifold and stability

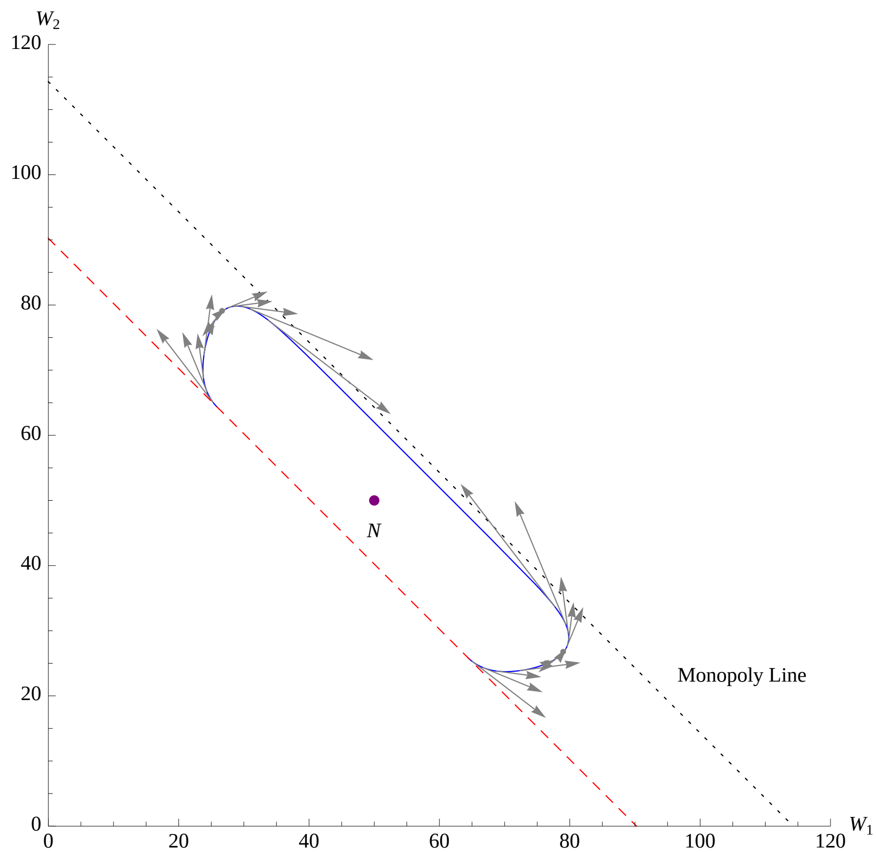


This figure plots the no-merger equilibrium manifold (blue, solid line). The gray arrows indicate the stability of the dynamic game, where the length of each arrow is the scaled drift of the value state vector. The static Nash equilibrium's output choices of (5,5) are depicted by N in terms of the continuation values of $(50, 50)$. We use the baseline environment in which $\Pi_1 = 30$, $\Pi_2 = 30$, $\beta_1 = 2$, $\beta_2 = 2$, $\delta_1 = 2$, $\delta_2 = 2$, $\sigma_1 = 1$, $\sigma_2 = 1$, and $r = 1$.

Continuing with this asset interpretation, the "Stage 2" agency problem for firm i is expressed in the objective (B.22) with state variable W^{-i} and state process equation (B.21). Examining (B.21) the "asset" interpretation is evident in the drift function: the state W^{-i} earns "interest" at rate r, but subtracting "consumption" g_{-i}, that is, the instantaneous profit flow of the rival, and these profits are reduced by increases in firm i's output A^i. Thus the first fundamental trade-off is that firm i implicitly must reward the rival firm $-i$ with an *increase* in firm $-i$'s continuation value if it reduces firm $-i$'s current profits by increasing output.

There is a second channel that affects its optimum via the term $-r(g_{-i} - W^{-i})\tilde{W}^i_{W^{-i}}$ in the HJB equation, (B.24): the shadow price—the slope of the equilibrium manifold—is expressed in

Figure 5. Merger equilibrium manifold and stability



This figure plots the merger equilibrium manifold (blue, solid line), the monopoly line (black, dotted line), and the merger line (red, dashed line) for a merger cost of $k = 24$. The gray arrows indicate the stability of the dynamic game, where the length of each arrow is the scaled drift of the value state vector. The static Nash equilibrium's output choices of (5,5) are depicted by N in terms of the continuation values of $(50, 50)$. We use the baseline environment in which $\Pi_1 = 30$, $\Pi_2 = 30$, $\beta_1 = 2$, $\beta_2 = 2$, $\delta_1 = 2$, $\delta_2 = 2$, $\sigma_1 = 1$, $\sigma_2 = 1$, and $r = 1$.

the partial derivative $\tilde{W}^i_{W^{-i}}$; in the regions of interest, the price war region and the collusion region, we know this shadow price is negative. This has the effect of reinforcing the trade-off facing firm i: by reducing its own output it increases its own continuation value and vice versa.

These trade-offs help to clarify the stability properties of the equilibrium manifold. Consider Figure 4's no-merger manifold and focus on the upper part of the collusion region. The arrows indicate that the firms will tend to move to the southeast, that is, toward the center of the collusion region. Examining the dynamics of the continuation value process vector in equation (B.21), we see that firm 2's value state is high, and firm 1's value state is low. Firm 1 "spends" the asset—the rival's continuation value—and reduces its own output, which allows firm 2 to choose high output, resulting in a negative drift; firm 1, because its current profit is significantly reduced, has positive drift as it earns "interest;" the result is that both states drift toward the center point of the collusion region. Similarly, to the left of the price war point, firm 1's continuation value state is again low. It "saves" by reducing its own output, allowing firm 2's continuation value to accumulate.

Even though the collusive region is smaller in Figure 5 than in Figure 4, because the option to merge weakens the punishments that enforce collusion and thus weakens collusion, it remains highly stable in the merger case. In addition, observe that there are two nodes where the stability flips between the collusion node and the merger node: the unstable one in the entrant and incumbent region of the equilibrium manifold, and the additional stable node in the contestability region nearer the price war or merger node. [17] Comparing the stability diagrams, in both figures the instability of the contestability region makes it likely that the firms will get back to cooperating if they stray into this region. In the unlikely event that the cusp in the contestability region is crossed, a price war (or a merger if there is the potential for it) is unstable, thus making a price war (or a merger) unlikely.[18]

We can therefore conclude that mergers are *rare*. This stems from *two* sources, both the instability of the merger nodes (which dynamically drives the equilibrium path away from the node, although not with certainty), and the collusion zone, which is far away from the merger nodes and which is stable.

4.3.1 Implications for market power

As we demonstrated in Proposition 1, as the merger cost falls, the equilibrium manifold changes generally. It flattens, reflecting that the firms are increasingly acting like a shadow monopoly in terms of output, with the main issue being the equity shares in the merged entity. Outsiders unaware of the potential for a merger attempting to value the companies would find output choices diminished relative to the theoretical prediction of the static Nash equilibrium. In addition, regulators would find greater collusion than would seem warranted by that same benchmark. This collusion will be strongest when the merger is most remote. For practical purposes, the merger will be a phantom,

[17] As will be apparent in the simulations of the model, the extra stable node in the contestability region has little impact on the actual dynamics of the equilibrium.

[18] However, if the merger state is approached, then in terms of corporate practice, this corresponds to mergers being "imminent" or "anticipated" just before they are announced. For example, Edmans, Goldstein, and Jiang (2012) and Cornett, Tanyeri, and Tehranian (2011) provide empirical evidence for this anticipation.

seemingly unrelated and hidden from the firms' current actions. While Andrade, Mitchell, and Stafford (2001), for example, point out that stronger antitrust laws and stricter enforcement have provided challenges for anti-competitive mergers, this model's solution implies that the dearth of market-power-increasing mergers need not imply more competition in a dynamic duopoly game, which is designed for the companies to compete.

The dynamic model thus suggests that tests pointing to rejection of the market power doctrine might be ill-posed: there is not much to deter if the anti-competitive effects of horizontal mergers are anticipated in merging and rival firms' product market strategies prior to merger announcements (or likely challenges by regulators). Consistent with our dynamic model's insight, Eckbo (1992) even concludes the following on p. 1005:

> While the U.S. has pursued a vigorous antitrust policy towards horizontal mergers over the past four decades, mergers in Canada have until recently been permitted to take place in a virtually unrestricted antitrust environment. The absence of an antitrust overhang in Canada presents an interesting opportunity to test the conjecture that the rigid market share and concentration criteria of the U.S. policy effectively deters a significant number of potentially collusive mergers. The effective deterrence hypothesis implies that the probability of a horizontal merger being anti-competitive is higher in Canada than in the U.S. However, parameters in cross-sectional regressions reject the market power hypothesis on samples of both U.S. and Canadian mergers. Judging from the Canadian evidence, there simply isn't much to deter.

In sum, the model is consistent with several regularities in the mergers and acquisitions literature that have heretofore led to rejection of the market power doctrine. The solution suggests an alternative interpretation of the literature's empirical tests. According to our dynamic model with the possibility of anti-competitive mergers, it is not surprising but rather inevitable that the evidence for the market power doctrine is weak when using capital market data and short-term announcement return methods to gauge changes in competition (or concentration) that have already taken place prior to the announcement return window when firms optimize dynamically.

5 Conclusion

We have studied mergers in a dynamic noisy collusion model, building on the models of Green and Porter (1984), Abreu, Pearce, and Stacchetti (1986), and especially Sannikov (2007). At each instant, firms either privately choose output levels or merge, which trades off benefits of avoiding price wars against the costs of merging. Mergers are optimal when collusion fails. Long periods of collusion are likely, because colluding is dynamically stable. Therefore, mergers are rare. Lower merger costs decrease pre-merger collusion, as punishments by price wars are weakened. This suggests that, although anti-competitive mergers harm competition ex-post, barriers and costs of merging due to regulation should potentially be reduced to promote competition ex-ante. We discuss the welfare implications more fully in Appendix J, which makes the case that there is an unambiguous welfare improvement from the potential to merge.

Our equilibrium solution combines an "agency" approach with a stochastic calculus technical approach. This method results in an interpretation of equilibrium behavior in terms of shadow prices that internalize the externalities the firms impose on each other in the duopoly.

We close by noting areas that warrant future research in this class of dynamic models. First, we have restricted attention to two firms. Extending our analysis to three or more firms would be informative about the impact of mergers on non-merging rivals as examined in many of the empirical studies. Moreover, Salant, Switzer, and Reynolds (1983), for example, show that the presence of three or more firms in an industry can deter mergers, and that mergers can be welfare-enhancing, even in the absence of scale economies or synergies. Perry and Porter (1985) examine this result further with a more fine-grained treatment of the allocation of costs in the merged firm and moved the conclusion back in the classical direction. Farrell and Shapiro (1990) establish that quantity competition in a post-merger industry raises prices if there are no scale economies or synergies, but still find cases where mergers are deleterious to potential merging firms. Other researchers, such as Deneckere and Davidson (1985) and Gaudet and Salant (1992), analyze welfare and policy implications in extensions of these models, again finding some counterintuitive results. The technical challenge in expanding the model to multiple firms is significant, however, in that equilibrium manifolds would reside in higher-dimensional spaces, with a concomitant increase in the computational difficulty of numerical solutions.

Second, as we show, the model is related to agency. While we treat firms as black boxes that are able to hide information, one might reinterpret this as more like a standard agency construct in which managers hide information from rival firms. With agency explicit, a merger might not eliminate all information asymmetries: we could ask whether the increase in market power effected by the merger is strengthened or weakened, and how pre-merger collusion is affected.

Finally, we note that our model does not include costs of production. Costs would vastly complicate the model, for two reasons. First, it would make sense to make those costs privately observable to each firm, adding an additional source of noise to the information structure of the model. Incorporating additional unobservable stochastic cost shocks here would add a signal-extraction element to the model, because firms would not be able to invert price signals to impute equilibrium actions as they can here; this in turn would introduce a signal-jamming dimension to the model, and concomitant additional complexity. We don't have signal jamming because firms "know" the recommended equilibrium action of the other firm. Second, cost shocks are by their nature private-value shocks, whilst demand shocks (as in the present structure of the model) are common-value shocks. This turns out to have major implications for how firms engage in, and react to, signal jamming. We refer readers to the paper of Bernhardt and Taub (2015) for a detailed treatment of this issue.

References

Abreu, Dilip, David Pearce, and Ennio Stacchetti, 1986, Optimal cartel equilibria with imperfect monitoring, Journal of Economic Theory 39, 251–269.

Alvarez, Fernando, and Urban Jermann, 2001, Quantitative asset pricing implications of endogenous solvency constraints, Review of Financial Studies 14, 1117–1151.

Andrade, Gregor, Mark Mitchell, and Erik Stafford, 2001, New evidence and perspectives on mergers, Journal of Economic Perspectives 15, 103–120.

Azar, José, Martin C. Schmalz, and Isabel Tecu, 2018, Anti-competitive effects of common ownership, Journal of Finance 73, 1513–1565.

Bernhardt, Dan, and Bart Taub, 2015, Learning about common and private values in oligopoly, RAND Journal of Economics 46, 66–85.

Björk, T., 2009, Arbitrage Theory in Continuous Time (Oxford University Press: New York).

Cornett, Marcia Millon, Basak Tanyeri, and Hassan Tehranian, 2011, The effect of merger anticipation on bidder and target firm announcement period returns, Journal of Corporate Finance 17, 595–611.

Davidson, Carl, and Raymond Deneckere, 1984, Horizontal mergers and collusive behavior, International Journal of Industrial Organization 2, 117–132.

Deneckere, Raymond, and Carl Davidson, 1985, Incentives to form coalitions with bertrand competition, Rand Journal of Economics 16, 473–486.

Dixit, Avinash, 1993, The Art of Smooth Pasting (Routledge: Abingdon).

Eckbo, B. Espen, 1992, Mergers and the value of antitrust deterrence, Journal of Finance 47, 1005–29.

Edmans, Alex, Itay Goldstein, and Wei Jiang, 2012, The real effects of financial markets: The impact of prices on takeovers, Journal of Finance 67, 933–971.

Farrell, Joseph, and Carl Shapiro, 1990, Horizontal mergers: An equilibrium analysis, American Economic Review 80, 107–126.

Fudenberg, Drew, David Levine, and Eric Maskin, 1994, The folk theorem with imperfect public information, Econometrica 62, 997–1039.

Fudenberg, Drew, and Jean Tirole, 1991, Game Theory (MIT Press).

Gaudet, Gerard, and Stephen Salant, 1992, Towards a theory of horizontal mergers, The New Industrial Economics: Recent Developments in Industrial Organization, Oligopoly and Game Theory, edited by George Norman Manfredi and La Manna pp. 137–157.

Green, Edward J., and Robert H. Porter, 1984, Noncooperative collusion under imperfect price information, Econometrica 52, 87–100.

Hurewicz, Witold, 1958, Lectures on Ordinary Differential Equations (Cambridge: MIT Press).

Knittel, Christopher R., and Victor Stango, 2003, Price ceilings as focal points for tacit collusion: Evidence from credit cards, American Economic Review 93, 1703–1729.

Miao, Jianjun, and Yuzhe Zhang, 2015, A duality approach to continuous-time contracting problems with limited commitment, Journal of Economic Theory 159B, 929–988.

Miller, Nathan H., Gloria Sheu, and Matthew C. Weinberg, 2019, Oligopolistic price leadership and mergers: The united states beer industry, Discussion paper, Georgetown University.

Nocke, Volker, and Lucy White, 2007, Do vertical mergers facilitate upstream collusion?, American Economic Review 97, 1321–1339.

Perry, Martin, and Robert Porter, 1985, Oligopoly and the incentive for horizontal merger, American Economic Review 75, 219–227.

Porter, Robert H., 1983, A study of cartel stability: The joint executive committee, 1880-1886, Bell Journal of Economics 14, 301–314.

Salant, Stephen, Sheldon Switzer, and Robert Reynolds, 1983, Losses from horizontal merger: The effects of an exogenous change in industry structure on cournot-nash equilibrium, Quarterly Journal of Economics 98, 185–199.

Sannikov, Yuliy, 2007, Games with imperfectly observable actions in continuous time, Econometrica 75, 1285–1329.

——— , and Andrzej Skrzypacz, 2007, Impossibility of collusion under imperfect monitoring with flexible production, American Economic Review 97, 1794–1823.

Thijssen, J. J., 2008, Optimal and strategic timing of mergers and acquisitions motivated by synergies and risk diversification, Journal of Economic Dynamics 32, 1701–1720.

A Inverting prices to obtain the information vector

In this appendix we detail the linear transformation from the underlying state processes X_t^1 and X_t^2 to the price processes. The X_t processes are

$$dX_t^1 \;=\; A_t^1\,dt \,+\, \sigma_1\,dZ_t^1\,, \tag{A.1}$$

and

$$dX_t^2 \;=\; A_t^2\,dt \,+\, \sigma_2\,dZ_t^2\,, \tag{A.2}$$

We asserted that

$$dX_t^1 \;=\; \frac{(A_t^1 + A_t^2)}{2\,(\Pi_2 \,-\, \delta_2 A_t^1 \,-\, \beta_2 A_t^2)}dP_t^2 \,-\, \frac{(A_t^1 - A_t^2)}{2\,(\Pi_1 \,-\, \beta_1 A_t^1 \,-\, \delta_1 A_t^2)}dP_t^1 \tag{A.3}$$

and

$$dX_t^2 \;=\; \frac{(A_t^1 + A_t^2)}{2\,(\Pi_2 \,-\, \delta_2 A_t^1 \,-\, \beta_2 A_t^2)}\,dP_t^1 \,-\, \frac{(A_t^2 - A_t^1)}{2\,(\Pi_1 \,-\, \beta_1 A_t^1 \,-\, \delta_1 A_t^2)}\,dP_t^2 \tag{A.4}$$

Write this as the matrix formulation

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -(A_t^1 - A_t^2) & A_t^1 + A_t^2 \\ A_t^1 + A_t^2. & -(A_t^2 - A_t^1) \end{pmatrix}\begin{pmatrix} \frac{dP_t^1}{(\Pi_1 - \beta_1 A_t^1 - \delta_1 A_t^2)} \\ \frac{dP_t^2}{(\Pi_2 - \delta_2 A_t^1 - \beta_2 A_t^2)} \end{pmatrix} \tag{A.5}$$

Now invert the coefficient matrix to obtain

$$\begin{pmatrix} \frac{dP_t^1}{(\Pi_1 - \beta_1 A_t^1 - \delta_1 A_t^2)} \\ \frac{(dP_t^2}{(\Pi_2 - \delta_2 A_t^1 - \beta_2 A_t^2)} \end{pmatrix} = \frac{2}{(A_t^1 - A_t^2)(A_t^2 - A_t^1) - (A_t^1 + A_t^2)^2}\begin{pmatrix} (A_t^1 - A_t^2) & A_t^1 + A_t^2 \\ A_t^1 + A_t^2 & (A_t^2 - A_t^1) \end{pmatrix}\begin{pmatrix} dX_t^1 \\ dX_t^2 \end{pmatrix} \tag{A.6}$$

Simplification yields

$$\begin{pmatrix} \frac{dP_t^1}{(\Pi_1 - \beta_1 A_t^1 - \delta_1 A_t^2)} \\ \frac{(dP_t^2}{(\Pi_2 - \delta_2 A_t^1 - \beta_2 A_t^2)} \end{pmatrix} = -\frac{1}{{A_t^1}^2 + {A_t^2}^2}\begin{pmatrix} (A_t^1 - A_t^2) & A_t^1 + A_t^2 \\ A_t^1 + A_t^2 & (A_t^2 - A_t^1) \end{pmatrix}\begin{pmatrix} dX_t^1 \\ dX_t^2 \end{pmatrix} \tag{A.7}$$

Multiplying out on the right hand side yields

$$-\frac{1}{{A_t^1}^2 + {A_t^2}^2}\begin{pmatrix} (A_t^1 - A_t^2)dX_t^1 + (A_t^1 + A_t^2)dX_t^2 \\ (A_t^1 + A_t^2)dX_t^1 + (A_t^2 - A_t^1)dX_t^2 \end{pmatrix} \tag{A.8}$$

Thus, the noise terms can be fleshed out, yielding

$$-\frac{1}{{A_t^1}^2 + {A_t^2}^2}\begin{pmatrix} ((A_t^1 - A_t^2)A_t^1 + (A_t^1 + A_t^2)A_t^2)dt \\ ((A_t^1 + A_t^2)A_t^1 + (A_t^2 - A_t^1)A_t^2)dt \end{pmatrix} - \frac{1}{{A_t^1}^2 + {A_t^2}^2}\begin{pmatrix} (A_t^1 - A_t^2)\sigma_1 dZ_t^1 + (A_t^1 + A_t^2)\sigma_2 dZ_t^2 \\ (A_t^1 + A_t^2)\sigma_1 dZ_t^1 + (A_t^2 - A_t^1)\sigma_2 dZ_t^2 \end{pmatrix} \tag{A.9}$$

$$-\frac{1}{{A_t^1}^2 + {A_t^2}^2}\begin{pmatrix} ({A_t^1}^2 + {A_t^2}^2)dt \\ ({A_t^1}^2 + {A_t^2}^2)dt \end{pmatrix} - \frac{1}{{A_t^1}^2 + {A_t^2}^2}\begin{pmatrix} (A_t^1 - A_t^2)\sigma_1 dZ_t^1 + (A_t^1 + A_t^2)\sigma_2 dZ_t^2 \\ (A_t^1 + A_t^2)\sigma_1 dZ_t^1 + (A_t^2 - A_t^1)\sigma_2 dZ_t^2 \end{pmatrix} \tag{A.10}$$

which simplifies to

$$\begin{pmatrix} \frac{dP_t^1}{(\Pi_1 - \beta_1 A_t^1 - \delta_1 A_t^2)} \\ \frac{(dP_t^2}{(\Pi_2 - \delta_2 A_t^1 - \beta_2 A_t^2)} \end{pmatrix} = -\begin{pmatrix} dt \\ dt \end{pmatrix} - \frac{1}{{A_t^1}^2 + {A_t^2}^2}\begin{pmatrix} (A_t^1 - A_t^2)\sigma_1 dZ_t^1 + (A_t^1 + A_t^2)\sigma_2 dZ_t^2 \\ (A_t^1 + A_t^2)\sigma_1 dZ_t^1 + (A_t^2 - A_t^1)\sigma_2 dZ_t^2 \end{pmatrix} \tag{A.11}$$

Multiplying by the denominators on the left hand side then yields the price process equations (1) and (2) with $d\zeta_t^1$ and $d\zeta_t^2$ implicitly defined as linear combinations of the dZ_t^1 and dZ_t^2 processes. Thus, the noise processes in price are correlated across the two firms, but these noise processes can be inverted from the price signal.

B Detailed derivations of the equilibrium

In this appendix we provide the details of the first and second stages of the solution procedure outlined in the main text.

B.1 First stage: Deriving the rival's optimized value process

The Hamilton-Jacobi-Bellman equation associated with the rival's problem in (14) with the state process in (3)-(4) is

$$rW^{-i} \quad = \quad \max_{A^{-i}} \left\{ r\, g_{-i}(A^{-i}, A^i) \,+\, A^{-i} W^{-i}_{X^{-i}} \,+\, A^i W^{-i}_{X^i} \,+\, \tfrac{1}{2}\sigma^2_{-i} W^{-i}_{X^{-i}X^{-i}} \,+\, \tfrac{1}{2}\sigma_i^2 W^{-i}_{X^i X^i} \right\}, \tag{B.12}$$

where the arguments of W^1 have been suppressed on the right hand side to avoid clutter, and where the cross-partial terms have dropped out given that the noise terms are uncorrelated.

Invoking our assumption that the action space is a continuum, we can use a conventional derivative to generate the optimality condition:

$$r\, g_{-i\,A^{-i}} \,+\, W^{-i}_{X^{-i}} \;=\; 0\,. \tag{B.13}$$

where we have made use of our assumption that g_i is differentiable. We also note again that it is not necessary to know the rival's policy function before deriving this condition.

We remark that a central assumption is that firm $-i$ cannot observe firm i's action A^i, yet A^i appears in the objective. An interpretation of this is as follows. Because the public information state is observable, in equilibrium each firm takes a "recommended" action A^i, and this recommendation is known to the rival firm—it is the "agency" aspect of the model. Thus, it is a requirement of equilibrium that each firm's conjecture about the rival's recommended action be correct. [19]

We can now characterize the firms' optimized continuation-value processes:

Lemma 1 *The firms' value states follow the processes*

$$dW^{-i} \;=\; r\left(W^{-i} - g_{-i}\right)dt - \sigma_{-i} r g_{-i\,A^{-i}} dZ_t^{-i} \,+\, \sigma_i W^{-i}_{X^i} dZ_t^i\,. \tag{B.14}$$

with a similar equation for firm i.

This is *promise-keeping.*

Proof: This follows from substituting the optimality conditions (B.13) and the HJB equation (B.12) into the Ito expansion of W_t^i. The detailed derivations are provided in Appendix C. □

B.1.1 Enforcement

Combining the continuation value processes for the two firms and denoting the volatility matrix by B_t yields the vector process:

$$\begin{aligned} dW_t \;&=\; r\,(W_t - g(A_t))dt \,+\, \begin{pmatrix} -\sigma_1 r g_{1\,A^1} & \sigma_2 W^1_{X^2} \\ \sigma_1 W^2_{X^1} & -\sigma_2 r g_{2\,A^2} \end{pmatrix} dZ_t \\ &=\; r\,(W_t - g(A_t))dt \,+\, B_t\, dZ_t\,. \end{aligned} \tag{B.15}$$

The volatility matrix B_t contains cross-partial derivatives that we can partially characterize.

Proposition 2 *The volatility matrix is singular.*

[19] We explore the equivalence of our stochastic calculus approach with Sannikov's approach in Appendix D.

Proof: The optimality conditions (B.13) can be multiplied to yield

$$r^2 g_{1\,A^1} g_{2\,A^2} \;=\; W^1_{X^1} W^2_{X^2} \tag{B.16}$$

Using the chain rule, we can write

$$W^1_{X^1} = W^1_{W^2} W^2_{X^1} \qquad \text{and} \qquad W^2_{X^2} = W^2_{W^1} W^1_{X^2} \tag{B.17}$$

and the condition (B.16) can then be written

$$\begin{aligned} r^2 g_{1\,A^1} g_{2\,A^2} - W^1_{X^1} W^2_{X^2} &= r^2 g_{1\,A^1} g_{2\,A^2} - W^1_{W^2} W^2_{X^1} W^2_{W^1} W^1_{X^2} \\ &= r^2 g_{1\,A^1} g_{2\,A^2} - W^2_{X^1} W^1_{X^2} = 0 \end{aligned} \tag{B.18}$$

This is the determinant of the volatility matrix B_t, which is therefore singular as asserted. □

Corollary 2 *The continuation value process maps out a one-dimensional manifold.*

Proof: The volatility matrix of the continuation value vector process is singular by Proposition 2, so the error process vector is mapped into a single effective stochastic process. Increments to this single process are added to the evolution determined by the drift functions, which is along a one-dimensional manifold. □

The singularity of the volatility matrix is the property of *enforcement*. It arises from the simultaneity of the firms' satisfying incentive compatibility. [20]

In equilibrium the firms react to their signals, altering their output in response to the signal, which is constantly stochastically perturbed. The resulting actions by both firms push them along the one-dimensional equilibrium path in the two-dimensional space of continuation values. Sannikov labeled this set $\partial\mathcal{E}$.

Viewing enforcement as requiring that the firms stay on the equilibrium manifold can be interpreted as the operation of a constraint, and the $W^{-i}_{X^i}$ elements of the volatility matrix are analogous to Lagrange multipliers, an interpretation that Sannikov did not provide. (The Lagrange multiplier interpretation holds in the additional sense that the multipliers are equal to derivatives with respect to the state, X_t, as we would expect.) We can interpret the multipliers as shadow prices, and these shadow prices provide incentives *beyond* the direct profit incentives in the g_i functions: they internalize the externality that firms have on each other.

The one-dimensionality of the enforced process allows us to carry out a change of variables and this will further enable our Lagrange multiplier interpretation.

B.2 Second stage: Agency reformulation

In the first stage formulation we generated the first order conditions for each firm, taking the other firm's policy rule as fixed. From firm i's perspective, firm $-i$'s policy rule can be viewed as a kind of *contract* against which firm i chooses its own actions. Knowing this, firm $-i$ wants to choose the *optimal* contract. The second stage optimization solves this problem.

B.2.1 Transforming the X_t state to the W_t state

We begin by implicitly mapping the X_t process to the continuation-value process W_t. We use the the cross-coefficients in the volatility matrix, $W^1_{X^2}$ and $W^2_{X^1}$ in equation (B.15), to determine the partial derivatives $W^1_{W^2}$ and $W^2_{W^1}$. Combining them we obtain slopes of W^1 in terms of W^2 and vice versa, and thus we obtain W^1 as an implicit function of W^2. This is roughly analogous to finding the slope of an indifference curve by taking the ratio of the marginal utilities.

[20] We explore the equivalence of our stochastic calculus approach with Sannikov's approach in Appendix D.

Substituting from the optimality condition (B.13) into equation (B.17) results in the transformations,

$$W^2_{X^1} \; = \; -rW^2_{W^1} g_{1A^1} = -r\frac{1}{W^1_{W^2}} g_{1A^1}\,, \tag{B.19}$$

and, similarly,

$$W^1_{X^2} = -rW^1_{W^2} g_{2A^2}\,. \tag{B.20}$$

To distinguish this transformed system we use a tilde notation, that is, $W^i_t = \tilde{W}^i(W^{-i}_t) = W^i(X_t)$, and similarly $\tilde{\xi}^i(W_t) = \xi^i(X_t)$, and so on; this approach has the usual abuse of notation in the sense that W^i_t denotes a process, whilst $W^i(X_t)$ is a function of the process X_t.

B.2.2 The state equation

To formulate the equivalent agency problem we first characterise the state variable process. Normalizing $\sigma_1^2 = \sigma_2^2 = 1$ and $\sigma_{12} = 0$, and eliminating X^1_t as an argument by substituting from equation (B.19) into equation (B.14), the continuation value process for firm -i in terms of W^{-i}_t and W^i_t is:

$$d\tilde{W}^{-i} \; = \; r\left(\tilde{W}^{-i} - g_{-i}\right)dt \; - \; rg_{-iA^{-i}}dZ^{-i}_t \; + \; (-r)\,\frac{1}{\tilde{W}^i_{W^{-i}}} g_{iA^i} dZ^i_t. \tag{B.21}$$

and similarly for firm 1.

B.2.3 The agency contract objective

The second-stage objective for firm i is

$$\tilde{W}^i\left(\tilde{W}^{-i}_t\right) = \sup_{\mathcal{T}^i_m, A^i(\cdot)} \mathbb{E}\Bigg[r\int_t^{\mathcal{T}^i_m} e^{-r(s-t)} g_i\left(A^i_s, A^{-i}_s\right)\right) ds$$
$$+ e^{-r(\mathcal{T}^i_m - t)}\tilde{\xi}^i(\tilde{W}^i_{\mathcal{T}^i_m}, \tilde{W}^{-i}_{\mathcal{T}^i_m})(\pi_m - k)\big|\mathcal{F}_t\Bigg], \tag{B.22}$$

with state process (B.21), and with the boundary condition

$$\tilde{W}^i(\tilde{W}^{-i}_{\mathcal{T}_m}) = \tilde{\xi}^i(\tilde{W}^i_{\mathcal{T}^i_m}, W^{-i}_{\mathcal{T}^i_m})(\pi_m - k). \tag{B.23}$$

and taking as given the other firm's control process A^{-i}. The Hamilton-Jacobi-Bellman equation for (B.22) is

$$0 = \max_{A^i}\Big\{r(g_i - \tilde{W}^i) - r(g_{-i} - \tilde{W}^{-i})\tilde{W}^i_{W^{-i}} + \tfrac{1}{2}\left(-r\tilde{W}^{-i}_{W^1} g_{iA^i}\right)^2 \tilde{W}^i_{W^{-i}W^{-i}} + \tfrac{1}{2}\left(rg_{-iA^{-i}}\right)^2 \tilde{W}^i_{W^{-i}W^{-i}}\Big\}. \tag{B.24}$$

This equation differs from the first-stage HJB equation, equation (B.12). It is however consistent with (B.12) in that it takes the solution of (B.12) as an implicit constraint.

The optimality condition is

$$g_{iA^i} - g_{-iA^i} W^i_{W^{-i}} = 0 \tag{B.25}$$

Observe that $W^1_{W^2}$ looks like a Lagrange multiplier on the indirect effect of A^i on firm $-i$'s payoff; this expresses the externality. That is, the marginal value of the action on the rival firm's payoff is equal to the multiplier. This is where "enforcement" becomes explicit.

We can say a bit more: the "constraint" is the drift of the rival's continuation value, which can be "steered" by firm i's action A^i_t; the multiplier expresses the shadow price of this constraint. And with the agency formulation we see that the shadow price is the trade-off between increasing firm i's own continuation value W^i_t in terms of firm $-i$'s continuation value W^{-i}_t; it is, as previously conjectured, the (local) slope of the equilibrium manifold.

B.2.4 The merger boundary

There are two boundary conditions. The first is the trite requirement of no jumps at the merger boundary. Firm i's share of the net payoff from merging at time $t = \mathcal{T}_m^i$ is given by:

$$\tilde{W}^i_{\mathcal{T}_m}(\tilde{W}^{-i}_{\mathcal{T}_m}) \;=\; (\pi_m - k) - \tilde{W}^{-i}_{\mathcal{T}_m}(\tilde{W}^i_{\mathcal{T}_m}) \;=\; \tilde{\xi}^1(\tilde{W}^i_{\mathcal{T}_m}, W^{-i}_{\mathcal{T}_m})(\pi_m - k). \tag{B.26}$$

This implicitly defines the endogenous function $\tilde{\xi}^i$ as a function of $\tilde{W}^i_{\mathcal{T}_m^i}$ and $\tilde{W}^{-i}_{\mathcal{T}_m^i}$. Solving for the share of firm i, $\tilde{\xi}^i$, yields:

$$\tilde{\xi}^i(\tilde{W}^i_{\mathcal{T}_m^i}, \tilde{W}^{-i}_{\mathcal{T}_m^i}) \;=\; 1 - \frac{\tilde{W}^{-i}_{\mathcal{T}_m}(\tilde{W}^i_{\mathcal{T}_m^i})}{\pi_m - k} \qquad \text{and therefore} \qquad \tilde{\xi}^i_{\tilde{W}^{-i}_{\mathcal{T}_m}} \;=\; -\frac{1}{\pi_m - k} \tag{B.27}$$

This is the value-matching condition.

The second condition is the smooth-pasting condition. This is found by differentiating the boundary function (B.27):

$$\tilde{W}^i_{\tilde{W}^{-i}}(\tilde{W}^{-i}_{\mathcal{T}_m}) \;=\; \tilde{\xi}^i_{\tilde{W}^{-i}_{\mathcal{T}_m}}(\pi_m - k) \;=\; -1\,, \tag{B.28}$$

with a similar condition for firm $-i$.[21] We begin with a lemma about the smooth-pasting condition. We show that the smooth-pasting condition locally satisfies the second-order condition for the firm solving the agency problem.

Lemma 2 *The smooth-pasting condition is necessary for an optimum with respect to the action A^1.*

Proof: See Appendix H. □

We can now address the challenge we posed about the simultaneity of the optimal stopping times.

Proposition 3 *The smooth-pasting condition implies equal stopping times: $\mathcal{T}_m^1 = \mathcal{T}_m^2$.*

Proof: The proof follows from two observations. First, for the value-matching condition to be met, that is, for the terminal point to be on the merger line, the value-matching condition is necessarily met for both firms simultaneously. Second, the smooth-pasting condition (in the stage 2 agency formulation) entails the condition

$$\tilde{W}^1_{W^2}(\tilde{W}^2_{\mathcal{T}_m}) = -1, \tag{B.29}$$

for firm 1; inverting the equation yields

$$\tilde{W}^2_{W^1}(\tilde{W}^1_{\mathcal{T}_m}) = -1, \tag{B.30}$$

which is the smooth-pasting condition for firm 1. Thus, satisfying the smooth-pasting formula for firm 1 necessarily satisfies the smooth-pasting formula for firm 2. □

It is worth noting the economic interpretation of the smooth-pasting condition. As the firms are driven to the merger line by the realizations of the noise, they stay on the equilibrium manifold by trading current payoffs against future "promise-keeping" payoffs; this trade-off is explicit in the sense of the shadow price $W^i_{W^{-i}}$ in equation (B.25). The smooth-pasting condition (B.28), which expresses incentive compatibility at the merger point, reflects—like Sannikov's (2007) incentive-compatibility condition (9) in what for us is the pre-merger play—the trading of utility between the two firms. However, the rate of exchange is fixed by the slope of the merger line. We summarize with the following proposition:

Proposition 4 *The action $A^i(\cdot)$ and $\tilde{W}^i(\cdot)$ that solve* (B.24), (B.23)*, and smooth-pasting condition* (B.28)*, taking as given A^{-i}, solve the optimal action and stopping problem* (B.22).

Proof: See Appendix H. □

[21] We again draw attention to our notation: $\tilde{W}^2_{\mathcal{T}_m}$, viewed by firm 1 as a state variable, denotes firm 2's continuation value evaluated at the stopping time $\mathcal{T}_m$, whilst $\tilde{W}^1_{W^2}(\tilde{W}^2_{\mathcal{T}_m})$ denotes the partial derivative of firm 1's continuation value as a function of that state at the stopping time.

B.3 Equilibrium and characterization

In an equilibrium of the game, firms must choose optimal contracts in their role as principals, optimally reacting to the other firm's contract in their role as agents, and the contracts must be identical; furthermore, they must agree on an identical stopping time.

We begin with the following definition, adapted from Sannikov (2007):

Definition 3 *A set $\mathcal{W}$ in the space of continuation values is self-generating if any initial pair of value processes $(W_0^1, W_0^2) \in \mathcal{W}$ then for every $t > 0$, $(W_t^1, W_t^2) \in \mathcal{W}$ and (ii) satisfy enforcement, that is, they satisfy* (B.21).

A manifold in continuation-value space satisfies enforcement, which we only defined previously in the context of the filtration generated from the X_t process, follows because our mapping $\mathcal{M}$ from X_t to $\tilde{W}_t$, implicitly defined in equations (B.19)-(B.20) is invertible. Because enforced paths satisfy optimality by construction, they are candidate equilibria.

The self-generating manifolds in Sannikov's non-merger analysis are equilibria of the non-merger game, because they are closed loops. In our model, the self-generating sets terminate at the merger line, however they still qualify as self-generating given that the merger point is a terminus. However, self-generation is not in itself sufficient to determine an equilibrium: it is possible to construct self-generating manifolds that satisfy the value-matching and smooth-pasting conditions, but which lie entirely *below* the merger line; these manifolds fail as equilibria precisely because both firms are better off by jumping to the merger line rather than evolve toward it via the self-generating manifold. We explore this in more detail in Appendix K. Thus, candidate self-generating sets must lie above the merger line.

Lemma 3 *For any manifold that is self-generating, satisfies the value-matching and smooth-pasting conditions, and which lies entirely above the merger line, merging prior to reaching the merger line does not make both firms better off.*

Proof: Because the merger announcement is public, the firms could mutually agree to merge prior to attaining the merger line, that is, they could agree to jump to some point on the merger line prior to attaining it via evolution along the manifold. By hypothesis, the manifold lies above the merger line, so for at least one of the firms the jump to the merger would reduce its continuation value and it would be not be individually rational to agree to the merger. □

Proposition 5 *The self-generating manifold consisting of the value-function processes that solve* (B.24)-(B.28) *for both firms simultaneously, such that the resulting manifold lies above the merger line, constitute a Markov merger equilibrium.*

Proof: We first observe that the actions A_t^i are adapted to $\mathcal{F}_t$ by construction. To demonstrate square integrability, that is, $E\int_0^\infty e^{-rt}\frac{1}{2}\left|A_t^i\right|^2 dt < \infty$, it suffices to demonstrate that they are bounded. The continuation values of the firms along the equilibrium manifold are such that in punishment mode the punishing firm i exerts a maximum punishment, and this is carried out by maximizing output A^i; this achieves the minimum continuation value for firm $-i$. Clearly A^i is bounded, as the continuation value is positive for the punished firm $-i$.

We next observe that because they are self-generating, the value function process W_t in (B.21) maps to the $\mathcal{F}_t$-adapted state process (B.14) that satisfies enforcement. The optimality of the associated A_t process is then implicit. Because the smooth-pasting condition is satisfied, by Proposition 4 an optimum is attained. □

C Derivation of the value process in the first stage

Here is the derivation of Lemma 1.

Proof: We establish the result for firm 1. We first apply Ito's lemma to $W^1(X_t^1, X_t^2)$ in generate the stochastic continuation value process of the state:

$$dW^1 \quad = \quad \left(A^1 W^1_{X^1} + A^2 W^1_{X^2} + \tfrac{1}{2}\sigma_1^2 W^1_{X^1X^1} + \tfrac{1}{2}\sigma_2^2 W^1_{X^2X^2}\right) dt \ + \ \sigma_1 W^1_{X^1} dZ_t^1 \ + \ \sigma_2 W^1_{X^2} dZ_t^2 . \quad \text{(C.31)}$$

Notice the resemblance of the terms in the drift to the stage-game payoffs in the Bellman equation. Substituting equation (B.12) into (C.31) yields a simpler expression for the continuation value process:

$$dW^1 = \left(r\, W^1(X^1, X^2) - r\, g_1(A^1, A^2)\right) dt + \sigma_1 W^1_{X^1} dZ_t^1 + \sigma_2 W^1_{X^2} dZ_t^2 . \quad \text{(C.32)}$$

We further modify this equation by using the optimality condition (B.13) to eliminate the $W^1_{X^1}$ term, replacing $W^1_{X^1}$ with $-rg_{1\,A^1}$ (i.e., the envelope condition):

$$dW^1 \ = \ \left(r\, W^1(X^1, X^2) - r\, g_1(A^1, A^2)\right) dt - \sigma_1 r g_{1\,A^1} dZ_t^1 + \sigma_2 W^1_{X^2} dZ_t^2 . \quad \text{(C.33)}$$

Dropping the arguments, we find that W^1 evolves according to:

$$dW^1 \ = \ r\left(W^1 - g_1\right) dt \ - \ \sigma_1 r g_{1\,A^1} dZ_t^1 \ + \ \sigma_2 W^1_{X^2} dZ_t^2 . \quad \text{(C.34)}$$

This eliminates the explicit influence of the state variable X_t^1 from the equation. □

D Connecting our approach to Sannikov's approach

We explore the equivalence of our approach with Sannikov's approach here.

D.1 Formulating the model using the history of public information (Sannikov's approach)

Our key departure from Sannikov's formulation is our positing that the public information process X_t can be treated as a state variable. This approach depends on the assumption that the rival firm chooses its actions based on the public information state; it is then optimal for the firm to react to the information state as well. In Sannikov's formulation, the continuation values are direct functions of the entire history of the public information process. We can establish informally that the two approaches lead to the same first order conditions.

Using Sannikov's approach, define the continuation value process for firm 1 as

$$dW^1 = (rW^1_t - g^1(A^1_t, A^2_t))dt + W^1_{X^1} dX^1_t + W^2_{X^2} dX^2_t \quad \text{(D.35)}$$

This is simply the differential form of Sannikov's equation (5) (Sannikov p. 1296), with the structure of the drift term accounting for discounting. Substituting from the definition of the X_t process,

$$\begin{aligned} & g^1(A^1_t, A^2_t)dt + W^1_{X^1}(A^1_t dt + \sigma_1 dZ^1_t) + W^2_{X^2}(A^2_t dt + \sigma_2 dZ^2_t) \\ =& (g^1(A^1_t, A^2_t) + W^1_{X^1}A^1_t + W^2_{X^2}A^2_t)dt + (W^1_{X^1}\sigma_1 dZ^1_t + W^2_{X^2}\sigma_2 dZ^2_t) \end{aligned} \quad \text{(D.36)}$$

Now substitute into (B.12), that is,

$$r\, W^1(X^1, X^2)dt = \max_{A^1} \left\{\mathbb{E}\left[r\, g_1(A^1, A^2)\, dt + W^1_{X^1}A^1_t + W^2_{X^2}A^2_t)dt + W^1_{X^1}(\sigma_1 dZ^1_t + W^2_{X^2}\sigma_2 dZ^2_t)\right]\right\} \quad \text{(D.37)}$$

the incentive condition is then

$$r\, g_{1\,A^1} \ + \ W^1_{X^1}(X^1, X^2) \ = \ 0 . \quad \text{(D.38)}$$

which is identical to equation (B.13).

We can then substitute from the optimality condition to re-express the drift of W^1, and so in optimized form equation (D.37) can be written as

$$dW^1 \ = \ r\left(W^1 - g_1\right)dt \ - \ \sigma_1 r g_{1\,A^1} dZ_t^1 \ + \ \sigma_2 W^1_{X^2} dZ_t^2 \,. \tag{D.39}$$

Compare with equation (6) of Sannikov (2007), p. 1296. Thus, we end up in the same place as with our state variable approach.

D.2 Connecting the stochastic calculus approach to Sannikov's martingale representation argument

Sannikov develops the "promise-keeping" and "enforcement" arguments using the martingale representation theorem. He uses the following discrete-time analogy: in a dynamic programming model we could write the *optimized* Bellman equation,

$$W_t = (1-\delta)g(A_t) + \delta E[W_{t+1}(y_t)|A_t]$$

where W_t is the value function, δ is the discount factor, and so on.

A heuristic way to develop the continuous time stochastic HJB equation is as follows. The objective is

$$V(X_0) = \max_{\{A_t\}} E\left[\int_0^\infty e^{-rt} f(A_t, X_t)dt\right]$$

subject to

$$dX_t = \mu(A_t, X_t)dt + \sigma(A_t, X_t)dZ_t$$

We write this in discrete dynamic programming form

$$\begin{aligned} V(X_0) &= E\int_0^{dt} f(X_t, A_t)dt + \max_a E\int_{dt}^\infty e^{-rt} f(X_t, A_t)dt \\ &= E\left[f(X_t, A_t)dt + e^{-r\,dt}V(X_{dt})\right] \end{aligned}$$

An algebra step yields:

$$E\left[e^{-r\,dt}V(X_{dt})\right] - V(X_0) = -f(X_t, A_t)dt$$

which we can write as

$$E\left[e^{-r\,dt}V(X_{dt})\right] = V(X_0) - f(X_t, A_t)dt$$

This is analogous with Sannikov's unnumbered equation at the top of page 1297, except for the addition rather than subtraction on the left hand side. However this is an artifact of the different approach to discounting used by Sannikov: if we treat δ in the usual way, his equation becomes

$$\delta E[W_{t+1}(y_t)|A_t] = W_t - g(A_t)$$

The left hand side is

$$e^{-r\,dt}V(X_{dt}) = V(X_0) + de^{-r\,dt}V$$

We can now proceed with the Ito expansion of the left hand side, yielding

$$\left(-rV + \mu(A_t, X_t)V_X + \tfrac{1}{2}\sigma(A_t, X_t)^2 V_{XX}\right)dt + \sigma V_X dZ_t + V(X_0) = V(X_0) - f(X_t, A_t)dt$$

Now we just need to remember that we can substitute from the optimized Bellman equation, completing the analogy with the Sannikov approach. Specifically, we can draw a more direct connection with the development of the enforcement matrix B using stochastic calculus, versus Sannikov's martingale representation theorem approach.

D.3 Relationship with Sannikov's treatment of enforcement

Sannikov's notion of enforcement is as follows: a volatility matrix B *enforces* a profile (A^1, A^2) if it satisfies incentive compatibility. He expresses this as the instantaneous drift for player i dominating the drift with an alternative profile, that is,

$$g_i(a) + \beta^i \mu(a) \geq g_i(a_i', a_{-i}) + \beta^i \mu((a_i', a_{-i}))$$

where we recall that $\mu(A)$ is the drift of the signal process.

Sannikov then makes a somewhat complicated argument in which he constructs and characterizes the β^i functions. We have employed a more direct approach using stochastic calculus.

This is simply reflecting the optimality requirement: in our case, rather than state an inequality reflecting optimization, we generate a conventional first order condition. After substituting this first-order condition and the Bellman equation into the Ito expansion of the continuation value W we obtain

$$dW^i \;=\; r\left(W^i - g_i\right)dt - \sigma_i r g_{i\,A^i} dZ_t^i \;+\; \sigma_{-i} W^i_{X^{-i}} dZ_t^{-i}\,.$$

The elements corresponding the the μ drift function are $\sigma_i r g_{i\,A^i}$ and $\sigma_{-i} W^i_{X^{-i}}$. If we look back at the HJB equation (B.12),

$$rW^1(X^1, X^2) \;=\; \max_{A^1} \Big\{ r\, g_1(A^1, A^2) + A^1 W^1_{X^1} + \alpha^2(X^1, X^2) W^1_{X^2} + \tfrac{1}{2}\sigma_1^2 W^1_{X^1X^1} + \tfrac{1}{2}\sigma_2^2 W^1_{X^2X^2} \Big\}, \quad \text{(D.40)}$$

we see that the element that is optimized with respect to A^i is

$$r\, g_1(A^1, A^2) + A^1 W^1_{X^1}$$

which corresponds exactly to Sannkov's optimization of $g_i(a) + \beta^i \mu(a)$, because $\mu(A) = A^i$ and $\beta^i = W^1_{X^1} = g_{i\,A^i}$. Thus, we recover the equivalence of the optimization of the HJB, conditional on the equilibrium play (optimization) of the rival firm, with Sannkov's approach.

E Converting to polar coordinates

The normal and tangent to the manifold are, in terms of the angle θ, respectively:

$$\mathbf{N}(\theta) = (\cos(\theta), \sin(\theta))\,, \qquad \text{and} \qquad \mathbf{T}(\theta) = (-\sin(\theta), \cos(\theta))\,. \quad \text{(E.41)}$$

We can then express the value process in polar coordinate form: the curvature $\kappa(W)$ of the equilibrium manifold, which is determined by the transformed form of the optimized Bellman equations of the firms, is the derivative of the polar coordinate with respect to movement along the equilibrium manifold. For example, κ would be zero if the manifold were locally a straight line. Thus, angle θ and arc length ℓ, we have along the manifold that:

$$\frac{d\theta}{d\ell} \;=\; \kappa\,, \qquad \text{and} \qquad \frac{d\ell}{d\theta} \;=\; \frac{1}{\kappa}\,, \quad \text{(E.42)}$$

The left-hand side derivative can be expressed in terms of the derivatives of the continuation values using polar coordinates, which leads us to the differential equation we solve numerically:

$$\begin{pmatrix} \frac{dW^1(\theta)}{d\theta} \\ \frac{dW^2(\theta)}{d\theta} \end{pmatrix} \;=\; \begin{pmatrix} -\frac{\sin(\theta)}{\kappa(\theta)} \\ \frac{\cos(\theta)}{\kappa(\theta)} \end{pmatrix} \;=\; \frac{\mathbf{T}(\theta)}{\kappa(\theta)}\,. \quad \text{(E.43)}$$

We can express the first-order cross-partial of W^1 in trigonometric form:

$$W^2_{W^1} = \frac{dW^2}{dW^1} = -\frac{\cos(\theta)}{\sin(\theta)}\,. \tag{E.44}$$

Hence the second-stage Bellman equation (B.24) becomes:

$$0 = \max_{A^1}\Big\{r\left(g_1 - W^1\right) - r(g_2 - W^2)\left(-\frac{\sin(\theta)}{\cos(\theta)}\right) + \tfrac{1}{2}\left(-r\frac{\cos(\theta)}{\sin(\theta)}g_{1\,A^1}\right)^2 W^1_{W^2W^2} + \tfrac{1}{2}\left(rg_{2\,A^2}\right)^2 W^1_{W^2W^2}\Big\}, \tag{E.45}$$

which then leads to:

$$W^1_{W^2W^2} = -\max_{A^1}\left\{\frac{\left(g_1 - W^1\right) - (g_2 - W^2)\left(-\frac{\sin(\theta)}{\cos(\theta)}\right)}{r\left(\left(-\frac{\cos(\theta)}{\sin(\theta)}g_{1\,A^1}\right)^2 + (g_{2\,A^2})^2\right)}\right\}. \tag{E.46}$$

After some algebra, the equation can be restated as follows:

$$W^1_{W^2W^2} = -\max_{A^1}\left\{\frac{\frac{1}{\cos(\theta)}\left(\cos(\theta)(g_1 - W^1) - (g_2 - W^2)\left(-\sin(\theta)\right)\right)}{r\cos(\theta)^2\left(\left(\frac{g_{1\,A^1}}{\sin(\theta)}\right)^2 + \left(\frac{g_{2\,A^2}}{\cos(\theta)}\right)^2\right)}\right\}, \tag{E.47}$$

or

$$W^1_{W^2W^2} = -\max_{A^1}\left\{\frac{1}{\cos(\theta)^3}\frac{\cos(\theta)(g_1 - W^1) + \sin(\theta)(g_2 - W^2)}{r\left(\left(\frac{g_{1\,A^1}}{\sin(\theta)}\right)^2 + \left(\frac{g_{2\,A^2}}{\cos(\theta)}\right)^2\right)}\right\}. \tag{E.48}$$

The numerator term is $\mathbf{N}(g - W)$, and the denominator term is $r|\phi|^2$, just as in Sannikov's formula. Notice that this equation has a curvature on the left-hand side. The fact that it is a curvature will later be used in the numerical solution of the model. We repeat the exercise with firm 2 and obtain:

$$W^2_{W^1W^1} = -\max_{A^2}\left\{\frac{1}{\sin(\theta)^3}\frac{\cos(\theta)(g_1 - W^1) + \sin(\theta)(g_2 - W^2)}{r\left(\left(\frac{g_{1\,A^1}}{\sin(\theta)}\right)^2 + \left(\frac{g_{2\,A^2}}{\cos(\theta)}\right)^2\right)}\right\}. \tag{E.49}$$

Notice that the denominators in equations (E.48) and (E.49) are the same.

As shown in Appendix G, the second-order partial derivatives of the continuation values are weighted expressions of the curvature of the equilibrium manifold in the direction of the normal vector, $\kappa(W)$:

$$\cos(\theta)^3 W^1_{W^2W^2} \;=\; \sin(\theta)^3 W^2_{W^1W^1} \;\equiv\; \tfrac{1}{2}\,\kappa(W)\,. \tag{E.50}$$

We can add the two curvature values in equations (E.48) and (E.49) and denote $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2$ to obtain an expression for the curvature:

$$\kappa(W) \;=\; \max_{A\in\mathcal{A}\setminus\mathcal{A}^N} \frac{2\,\mathbf{N}\,(g - W)}{r|\phi|^2}\,, \tag{E.51}$$

which is Sannikov's (2007) optimality equation. Note that the maximization in equation (E.51) is over both A^1_t and A^2_t (excluding the set of pure strategy Nash equilibria, $\mathcal{A}^N$, as mentioned earlier). This is innocuous here because the numerator and denominator are each additively separable in A^1_t and A^2_t, so separate maximization for each firm taking its turn as the "principal" is satisfied. ■

Sannikov demonstrates that the differential equation system in (E.51) satisfies Lipschitz conditions and therefore has a solution:

> ...the right-hand side of the optimality equation [(E.51)] is Lipschitz-continuous in w and T. This property guarantees the existence of solutions to the optimality equation from any initial conditions and helps us characterize the set $\mathcal{E}(r)$. [Sannikov p. 1320]

Sannikov then continues with his Lemma 3, culminating in his Proposition 4, establishing existence. (See also Hurewicz (1958), p. 28.)

F Why maximizing the Bellman equation is equivalent to maximizing a ratio in the curvature ODE

Having established that the "agency" optimization problem in equation (B.24) is equivalent to the ODE in equation (E.48) for the curvature of the equilibrium set boundary $\partial\mathcal{E}$, we need to show why the optimization of the Bellman equation is equivalent to the optimization of the ratio in the ODE.

Consider the abstract problem:

$$\max_x \; \{f(x) \,+\, A\, g(x)\} \,. \tag{F.52}$$

The first-order condition is:

$$f'(x) \,+\, A\, g'(x) \;=\; 0\,, \tag{F.53}$$

or:

$$A \;=\; -\frac{f'}{g'}\,. \tag{F.54}$$

Now consider the maximization problem:

$$\max_x \; \frac{f(x)}{g(x)}\,. \tag{F.55}$$

The first-order condition can be written as:

$$\frac{f}{g} \;=\; \frac{f'}{g'}\,, \tag{F.56}$$

and therefore, at the maximum, we have that:

$$\max_x \; \frac{f}{g} \;=\; \frac{f'}{g'}\,. \tag{F.57}$$

Therefore,

$$A \;=\; -\frac{f'}{g'} \;=\; -\,\max_x \; \frac{f}{g}\,. \tag{F.58}$$

Thus, the maximization of the ratio generates the same optimum (adjusted for the sign) as the Bellman equation. ■

G Curvature equality

We want to show that:

$$\cos(\theta)^3\, W^1_{W^2W^2} \;=\; \sin(\theta)^3\, W^2_{W^1W^1}\,. \tag{G.1}$$

To begin, note that:

$$\frac{d}{dW^1} W^1_{W^2} = W^1_{W^2W^2}\frac{dW^2}{dW^1} \;=\; -\,W^1_{W^2W^2}\frac{\cos(\theta)}{\sin(\theta)}\,. \tag{G.2}$$

This is equal to:

$$\frac{d}{dW^1}\frac{1}{W^2_{W^1}} \;=\; -\,\frac{1}{\left(W^2_{W^1}\right)^2} W^2_{W^1W^1} \;=\; -\,W^2_{W^1W^1}\left(\frac{\sin(\theta)}{\cos(\theta)}\right)^2\,. \tag{G.3}$$

Equating the two terms and performing algebra yields the result. ■

H Optimality of the smooth-pasting condition

In this section we prove Lemma 2 and Proposition 4, which establish optimality of the smooth-pasting condition. Our strategy differs a bit from the more well-known approaches such as Dixit (1993); we will show that if the smooth-pasting condition is satisfied then the second-order condition associated with the Bellman equation is locally satisfied at the boundary point characterised by the the smooth-pasting condition. That is, we provide a verification of the sufficiency of local optimality at the merger point. More concretely, the steps in this demonstration are as follows:

(i) Calculate the second-order condition for the Bellman equation from the optimality condition;

(ii) Evaluate the "ratio" version of the Bellman equation at the value-matching point, using the value-matching condition and also the smooth-pasting condition, resulting in a reduced-form expression for the second partial derivative $W^1_{W^2W^2}$;

(iii) Substitute this reduced-form expression for $W^1_{W^2W^2}$ into the second-order condition, as well as the value-matching condition, establishing that the second-order condition is negative.

Proof: (Of Lemma 2) Commencing step (i), the reprise of the Bellman equation is

$$0 = \max_{A^1} \left\{ r(g_1 - W^1) - r(g_2 - W^2)W^1_{W^2} + \tfrac{1}{2}\left(-rW^2_{W^1}g_{1A^1}\right)^2 W^1_{W^2W^2} + \tfrac{1}{2}\left(rg_{2A^2}\right)^2 W^1_{W^2W^2} \right\}. \tag{H.4}$$

with optimality condition

$$rg_{1A^1} - rg_{2A^1}W^1_{W^2} + \left(\left(-r^2W^2_{W^1}g_{1A^1}\right)g_{1A^1A^1} + \left(r^2g_{2A^2}\right)g_{2A^2A^1}\right)W^1_{W^2W^2} = 0 \tag{H.5}$$

The second-order condition can make use of the quadratic structure of g_1 and g_2: the third derivatives are zero, so we have

$$rg_{1A^1A^1} - rg_{2A^1A^1}W^1_{W^2} + r^2\left(-W^2_{W^1}(g_{1A^1A^1})^2 + (g_{2A^2A^1})^2\right)W^1_{W^2W^2}$$

Now we can substitute the smooth-pasting condition: $W^1_{W^2} = -1$:

$$rg_{1A^1A^1} + rg_{2A^1A^1} + r^2\left(-W^2_{W^1}(g_{1A^1A^1})^2 + (g_{2A^2A^1})^2\right)W^1_{W^2W^2}$$

The same holds for the other player: $W^2_{W^1} = -1$:

$$rg_{1A^1A^1} + rg_{2A^1A^1} + r^2\left((g_{1A^1A^1})^2 + (g_{2A^2A^1})^2\right)W^1_{W^2W^2}$$

Now recall the structure of the stage game payoff function g_1:

$$g_1(a_1, a_2) = a_1(\Pi_1 - \beta_1 a_1 - \delta_1 a_2),$$

and similarly for g_2, so that

$$g_{1A^1A^1} = -2\beta_1 - \delta_1 \qquad g_{1A^2A^1} = -\delta_2,$$

which are both negative by assumption.

Carrying out step (ii), we next transform the system of Bellman equations to ratio form to isolate the second partials. Appendix F demonstrates by straightforward algebra that the maximization of the following ratio is equivalent to the original maximization in (B.24):

$$W^1_{W^2W^2} = -\max_{A^1}\left\{\frac{(g_1 - W^1) - r(g_2 - W^2)\frac{1}{r}W^1_{W^2}}{r\left(\left(W^2_{W^1}g_{1A^1}\right)^2 + \left(g_{2A^2}\right)^2\right)}\right\}. \tag{H.6}$$

Substituting $\frac{1}{W^1_{W^2}}$ for $W^2_{W^1}$ yields:

$$W^1_{W^2W^2} = -\max_{A^1}\left\{\frac{(g_1 - W^1) - r(g_2 - W^2)\frac{1}{r}W^1_{W^2}}{r\left(\left(\frac{1}{W^1_{W^2}}g_{1A^1}\right)^2 + \left(g_{2A^2}\right)^2\right)}\right\}, \tag{H.7}$$

which, along with the first-order condition in A^1, is an ordinary differential equation (ODE) in W^1. Thus, we have converted the Bellman equation from a partial to an ordinary differential equation.

Finally, carry out step (iii), substituting for $W^1_{W^2W^2}$ from the ratio reformulation of the Bellman equation in equation (H.7), evaluated at the value-matching and smooth-pasting point:

$$\begin{aligned} W^1_{W^2W^2} = -\frac{(g_1 - W^1) - r(g_2 - W^2)\frac{1}{r}W^1_{W^2}}{r\left(\left(\frac{1}{W^1_{W^2}}g_{1A^1}\right)^2 + \left(g_{2A^2}\right)^2\right)} &= -\frac{(g_1 - W^1) + (g_2 - W^2)}{r\left(\left(\frac{1}{(-1)}g_{1A^1}\right)^2 + \left(g_{2A^2}\right)^2\right)} \\ &= -\frac{(g_1 - W^1) + (g_2 - W^2)}{r\left(\left(g_{1A^1}\right)^2 + \left(g_{2A^2}\right)^2\right)} \end{aligned} \tag{H.8}$$

If we can demonstrate that the numerator of this expression is negative or zero then we will have demonstrated that the second-order condition holds at the smooth-pasting point.

Lemma 4 *At the smooth-pasting point, we have that:*

$$(g_1 - W^1) + (g_2 - W^2) = 0\,.$$

Proof: The expressions $(g_1 - W^1)$ and $(g_2 - W^2)$ are the drifts of the continuation value processes for W^1 and W^2, respectively (see equation (B.15)). At the smooth-pasting point, these drifts necessarily are equal and of the opposite sign in order to point along the merger line, which has a slope of -1, and they therefore sum to zero. (Equivalently, the curvature at the smooth-pasting point, $W^1_{W^2W^2}$, is zero.) □

This completes step (iii), establishing the result. □

We can use this result to prove global optimality.

Proof: (of Proposition 4) The optimality of the main action A^1_t as the solution of the HJB equation (B.24) follows directly from conventional optimal control considerations. Therefore the optimal value state path satisfies equation (E.48); further derivations lead to the differential equation (E.43), combined with equation(E), as demonstrated in E. Sannikov (Sannikov (2007) Theorem 2, p. 1309) establishes that any optimal path must satisfy this differential equation system.

The remaining issue is to provide two boundary conditions for the second-order differential equation, (E.48), that the optimal path necessarily satisfies; the value-matching condition (B.23) provides one boundary condition. By Lemma 2, the smooth-pasting condition is locally optimal. Suppose that an alternative optimal path exists that does not satisfy the smooth-pasting condition. In that case, there is a kink at the merger boundary, and thus local optimality cannot be satisfied. Any equilibrium path must satisfy optimality, and hence any equilibrium path satisfies smooth pasting.

The kink is "one-sided:" the local curvature at the kink point is *positive* infinity, but cannot be negative infinity because the differential equation can approach the merger line from above. Therefore it is only necessary only to observe that the violation of the relevant inequality rests on the positive infinity property. □

We remark that our results have not established uniqueness of the equilibrium path. We know from numerical experiments that two manifolds that satisfy the smooth-pasting condition exist, with a smaller one fully contained within the larger one that we have analyzed here.

The "standard" proof of the optimality of the smooth-pasting point such as in Dixit (1993) uses a Taylor expansion of the solution of a second-order differential equation, exploiting the structure of the solution stemming from the assumption of fixed coefficients. Our model does not lead to an equation with fixed coefficients, so it is not possible to use Dixit's approach directly. Notice, however, that the smooth-pasting condition is *locally* optimal at the smooth-pasting point; by continuity of the solution that obeys the main differential equation this optimality argument must also hold in a neighborhood of the smooth-pasting point.

I Proof of the inclusion result, Proposition 1

Proof: The proof has three main parts.

(i) In the first part we observe that the boundary of the equilibrium set $\mathcal{E}$, $\partial\mathcal{E}$, is described by the same ordinary differential equation, equation (E.43), for any merger model with merger cost K and for the no-merger case. If the inclusion were violated there would be some point $(\tilde{W}^1, \tilde{W}^2)$ at which the differential equation would necessarily hold identically for both K and K' at $(\tilde{W}^1, \tilde{W}^2)$. However non-inclusion would necessarily imply that the local differential would be different at $(\tilde{W}^1, \tilde{W}^2)$, a contradiction. Thus, the merger manifold must lie either entirely inside the no-merger manifold or entirely outside it.

(ii) Now consider the case where the merger manifold is entirely inside the no-merger manifold, that is, $\partial\mathcal{E}^K_{\mathrm{M}} \subset \partial\mathcal{E}_{\mathrm{NM}}$. To begin, note that the equilibrium manifold is locally characterized by the angle θ, the slope of the manifold for each θ, and the curvature κ. The curvature κ is characterized by equation (E.51), which we reproduce here:

$$\kappa(\tilde{W}) \;=\; \max_{A\in\mathcal{A}\setminus\mathcal{A}^N} \frac{2\,\mathbf{N}\,(g-\tilde{W})}{r|\phi|^2} \tag{I.9}$$

The key part of this equation is the numerator term

$$-\cos(\theta)(g_1-\tilde{W}^1)-\sin(\theta)(g_2-\tilde{W}^2) \tag{I.10}$$

(See equations (E.48) and (E.49).)

We will use the standard convention for measuring the angle θ, that is, $\theta = 0$ along the x-axis.

- Suppose that we solve the differential equation for the manifold with initial point at the middle of the collusion point. In that case, $\theta = \frac{\pi}{4}$, so that $\cos(\theta) = \sin(\theta)$, resulting in equal weights on the numerator term in (I.10). Given that the smooth-pasting manifold must lie inside the no-merger manifold, if we reduce $\tilde{W}_1$ and $\tilde{W}_2$, that is, reduce the initial point of the differential equation, then the absolute value of the curvature must increase. Therefore the path of the differential equation will move away from the no-merger path toward the interior.
- Next, consider $\theta = -\frac{\pi}{4}$, corresponding to the "contestability" region. In this region $\sin(\theta) = -\cos(\theta)$, so that there is a negative weight, $\sin(\theta)$, on the $(g_2 - W^2)$ term, and a positive weight, $\cos(\theta)$, on the $(g_1 - W^1)$ term. Now suppose we increase $\tilde{W}_2$ and decrease $\tilde{W}_1$ along the negative-45 degree line, so that the sum $\tilde{W}^1 + \tilde{W}^2$ is constant. In that case, the negative weight on $\tilde{W}_1$ and the positive weight on $\tilde{W}^2$ means that the net effect of the movement is to increase the curvature. Therefore as we move inward, the curvature increases.

This argument implicitly holds the actions A^i fixed; however, recall that the curvature is defined by the formula in (E.51), which has a maximization with respect to the vector A on the right hand side. Therefore, adjustments of A that are driven by our thought experiments with $\tilde{W}$ will not affect our argument, as the effects of A drop out due to the envelope condition.

It is clear that this logic holds for all other angles: moving toward the center increases the curvature. The result is that the manifold "curls up" near the merger line, so that a tangency to the merger line becomes possible; this is apparent in the numerical simulations presented in the figures.

The remaining possibility is that the inclusion is reversed. (This possibility is illustrated in Figure 10 of Appendix G; clearly it is possible for the ODE to have a solution lying outside the no-merger manifold.)

We can see that the tangency condition cannot be satisfied locally where the no-merger boundary $\partial\mathcal{E}_{\mathrm{NM}}$ crosses the merger line. By continuity of the *derivative* of the manifold satisfying the ODE at that point, it cannot flip over and become tangent.

As we move away from that point, the logic we developed above for the "inside" manifold, establishes that the slope continues to flatten, so it moves away from a potential tangency, and the smooth pasting condition therefore cannot be satisfied.

However, we need to rule out the possibility that the smooth pasting conditions are satisfied, but pointing the wrong way. However this would require that the curvature of the manifold boundary reverse itself, at some locus, which in turn would imply that the ODE (48) was locally linear at the locus where this reversal occurs, that is, the local second derivative would be zero. But the structure of the ODE precludes this: it would imply that the differential equation is "stuck" at the zero-curvature point; because κ would be negative the linear part of the manifold would be infinitely long (see equation (E.43)).

(iii) The third part of the proof is to establish that the first inclusion holds, namely that increasing the merger cost K increases the equilibrium set. But this case is similar to the previous one, in that it would require that the merger manifold would not be tangent as we decrease K. □

J A remark on social welfare

The continuation values in the model are discounted profits, so we can infer that when these profits increase, the firms are colluding to a greater extent, and consumer surplus is concomitantly reduced. Thus, social welfare is inversely related to the firms' joint profits.

The smooth-pasting equilibrium has lower firm profits when the firms are in the collusion phase than they do if they can never merge, that is, the equilibrium manifold moves in the southwesterly direction relative to the no-merger manifold.

However, the merger cost itself is part of the deadweight loss.[22] How do we incorporate this deadweight loss in the accounting? The answer is that we can ignore it, because the firms' discounted profits (their continuation values) in the collusion phase are still close to the monopoly line, despite the discounted merger cost. Thus, this deadweight cost is, from the firms' perspectives, just an alternative cost like the cost they would face if they entered into the price war in the no-merger model.

K Refusals to merge, punishment and jumps

In this appendix[23] we explore the impact of assumptions about the structure of the underlying game, especially the details of the moment of the merger, on outcomes. We focus on one issue: the consequences of the fact that the decision to merge is observable to the rival firm. We emphasize that our discussion is informal.

One of the key notions in the main text is that the firms find it optimal to merge when the value states attain a point on the merger line. Merging at that point is a binary decision, and each firm can observe whether or not the other firm has merged with it because in the event of the merger they subsequently share monopoly profits.

Because the decision by each firm to merge is observable, if the firms do not commit to the merger in advance, the possibility that a firm would refuse to merge can also be admitted as a strategy, and the continuation of the game must be specified in that case. Because the refusal to merge is observable, the structure of the game at that moment and the associated continuation are potentially very different from the noisy collusion game that has been played up to that point.

[22] We are grateful to Mikhail Panov for raising this question.

[23] The ideas and contributions of Mikhail Panov, Yuliy Sannikov, and Andrzej Skrzypacz led to the discussion in this appendix.

If one firm refuses to merge, then, given the noncooperative environment of the game, it is appropriate to consider how the other firm would subsequently punish it. If the punishment is effective then the firm contemplating refusal would be be deterred from refusing, and the merger would take place, but potentially without necessarily satisfying smooth pasting. One can broadly describe this punishment: it would be to revert to a continuation phase with the worst possible outcome for the firm that has refused to merge.

The punishment phase must must itself be an equilibrium. Given that the continuation game would not differ from the collusion game in the key respect of having actions that are obscured by noise, the punishment phase would be constituted from the same elements as the pre-merger (or pre-refusal) game. If such an equilibrium exists, we refer to it as a refusal-punishment (R-P) equilibrium.

What we will establish here is that credible punishment phases cannot in fact themselves be equilibria, and so alternative equilibria that are sustained by such punishments cannot exist.

The refusal-punishment construction

For a refusal-punishment equilibrium to function, there must be agreement between the firms about the punishments in the punishment game. These punishments, and in essence the structure of the continuation game, must be agreed to in a pre-play phase of the game the firms play in advance of the noisy collusion game. Thus, there are really three phases of the overall game: (i) the pre-play phase in which the punishments that will be coordinated upon in the event of a refusal to merge; (ii) the pre-merger phase of the noisy duopoly game; (iii) the punishment phase in the event that a firm refuses to merge, which itself must be an equilibrium of a pre-merger phase of the game, including the potential for repeated refusal-punishment phases. Thus, in this final stage, any boundary conditions entailed by the potential for future mergers, or the lack thereof, must be delineated.

In the sequel we examine the simplest case: that if a merger has been refused and punishment invoked, there is no further potential to merge: the punishment is permanent.

If there is no potential to merge, then the game reverts to the collusion game described by Sannikov (2007): the value states of the firms stochastically and continuously transit around an egg-shaped manifold with payoffs that are strictly bounded by the monopoly line. The initial point on the manifold is determined by the extreme punishment: if firm 1 (with value state on the horizontal axis) has refused to merge, the worst continuation on the no-merger manifold is the leftmost point on it (point B in the figure); similarly if firm 2 (with value state on the vertical axis) has refused to merge, the continuation commences on the lowest point on the manifold (point B' in the figure).

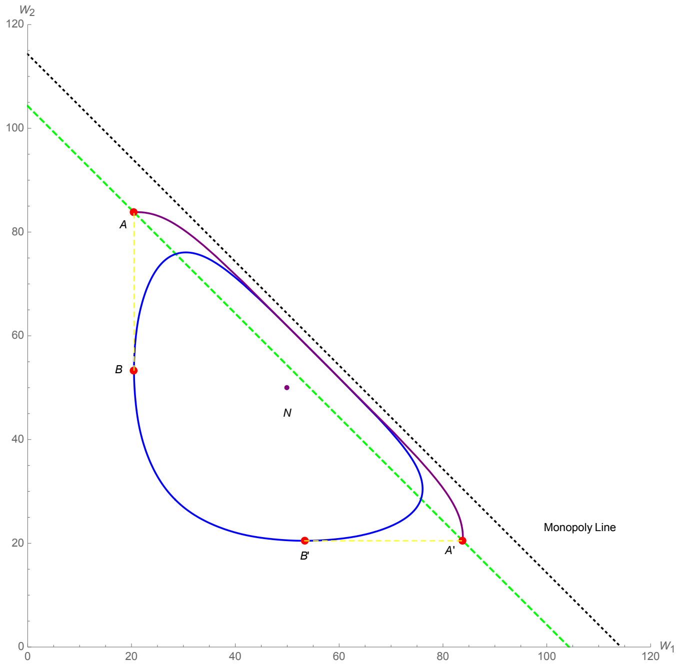
Punishments alter the boundary conditions of the pre-merger game. At the merger point, the punishment must be weakly worse than the worst state at the moment of the merger. Because the no-merger manifold is fixed in size and location, the leftmost and bottom-most punishment points on the manifold are fixed. This in turn dictates the corresponding upper and lower merger points on the merger line (points A and A' in Figure 6). Finally, because the equilibrium pre-merger manifold must be continuous and obey the underlying main differential equation, the pre-merger manifold must land on these two points, and is thus fully and uniquely determined.[24]

Because the landing points are determined by potential punishments, the smooth-pasting condition no longer determines the boundary conditions in this situation. The resulting manifold is semicircular and is not tangent at the landing points. We will refer to this manifold as the refusal-punishment (R-P) manifold. In addition to not being tangent at the merger line, the R-P manifold lies entirely between the no-merger manifold and the monopoly line. This is in contrast to the smooth-pasting manifold, which lies entirely inside the no-merger manifold.

This construction has *two* ingredients. The first is that the refusal and punishment are *observable*, unlike the pre-merger play, and also unlike the no-merger manifold that the firms jump to if the punishment is carried out. The second is that the punishment entails a *jump*. This is key: in the other stages of the game,

[24] Of course in equilibrium the punishments are not carried out. If they were carried out, the equilibrium manifold would then have to terminate on the punishment point, and this would in turn eliminate the original manifold from consideration. Thus, our example is already vitiated in the sense that it cannot be an equilibrium, but we continue with it to underscore an additional point.

Figure 6. The refusal-punishment equilibrium



In this figure one of the firms can refuse to merge; thereafter it is assumed that no merger is possible. If a firm refuses to merge, say at point A, then a punishment is initiated via a jump to point B, with symmetric possibilities at points A' and B'. The punishment manifold is simply the original no-merger manifold, with play evolving along this manifold as in the original Sannikov no-merger model.

jumps are not possible because actions, and the payoffs they generate, which are a flow, cannot generate discrete jumps in the continuation values, which are the integrals of these expected future payoffs and as such are continuous. But the *observability* of the refusal means that the game is no longer in its noisy mode at that instant, and a more conventional, static, full-information game can be played. Because there is a jump if the punishment is invoked, the punishment is far more severe than the incremental punishments of the pre-merger game. As the cost shrinks, the size of the jump interval actually increases, making the jump punishments even stronger. This is why the R-P manifold has the potential to achieve better cooperation than the no-merger game alone can achieve, if it is an equilibrium.

The punishments are not equilibria

The above reasoning breaks down for two reasons.

First, in equilibrium the punishments cannot be carried out with the proposed structure. If they were carried out, the equilibrium manifold would then have to terminate on the punishment point, because the continuation values are the integrals of the expected future payoffs, including the point at which the punishment commences, and as such are continuous and cannot accommodate jumps. This would in turn eliminate the original manifold from consideration.

Second, the punishment phase of the equilibrium would entail the firms traversing the punishment manifold that lies *below* the merger line. Every point on the punishment manifold has the property that it is worse for both firms relative to some point on the merger line. One of the firms could propose to move to such a point from the punishment manifold by merging, and the other firm could accept this proposal and improve its payoff.[25] This logic works because the proposal to merge, unlike the production action, is publicly observable. Therefore the punishment manifold is not immune to renegotiation in this sense, so in turn the punishment cannot be an equilibrium.[26]

This leaves only the manifolds that satisfy the smooth-pasting property as potential equilibria: because the smooth-pasting manifolds lie *above* the merger line, it is not better for both firms to immediately jump to the merger line. Thus, the R-P manifolds are ruled out. [27]

L Ruling out never merging

In this appendix we again examine the potential for the firms to never merge, but focus on how these equilibria are ruled out by Pareto-improving mergers. Unlike the discussion in Appendix K, in this appendix we examine the opposite situation: instead of being punished by refusing to merge via a discrete and negative jump that reduces the continuation values of one or both firms, we examine the situation in which the firms have the potential to achieve a discrete positive improvement in their continuation values by merging.

Never merging might be viewed as a kind of coordination-failure continuation path where firms do not merge. Because the action space is such that both firms need to say they are ready to merge at the same time, there is a nominal potential for both firms never say they are ready to merge because each firm perceives that the rival will never agree to merge. In that case either firm refusing to merge seems to be a best response, since no matter what happens there is never going to be a merger.

If the firms never merging is an equilibrium then there is only one maximal equilibrium set and its associated boundary associated with this equilibrium, namely the manifold described by Sannikov, which we have replicated as the outer manifold in Figures 1 and 2 of the main text; we will refer to this manifold as the no-merger manifold. A key property of this manifold is that it dips below the merger line (except in the case where the merger cost is so high that the merger line is everywhere below the no-merger manifold, which we have not depicted).

We can focus on an arbitrary point on the no-merger manifold that lies below the merger line. For purposes of discussion denote this point by P. It is evident that there is a continuum of Pareto-improving mergers attained by jumping to the merger line, and the improvement in the continuation values is discrete. We can focus on one of these Pareto-improving jumps, namely the one such that the starting point P on the no-merger line is at right angles to the merger point.

Because there is a discrete improvement in the continuation payoffs, and because the improvement is the same for both firms, we can label this improvement as $\gamma > 0$ for both firms. Of course the magnitude of γ depends on the locus of the continuation values on the no-merger manifold, but the key property is that γ is positive.

We can now pose this situation faced by the firms at point P as an auxiliary game. We present the payoff matrix of the game in Table 1. The payoffs from agreeing to merge are the improvements in the continuation values, whilst the zero payoffs reflect the lack of improvement by refusing to merge or by both firms refusing. Another important detail of the auxiliary game is that it is a full-information game, unlike the tacit-collusion game that is our main object of study.

A key property of the payoff matrix is that there is no net punishment for refusing, and therefore the game is properly called a pure coordination game, also known as an assurance game: if the players play different strategies they both get zero. If they play the same strategy then there is a positive reward,

[25] The target locus would need to be decided in the pre-play phase of the game via a separate bargaining solution that we do not model. However any standard bargaining solution such as Nash bargaining or Kalai-Smorodinsky would have Pareto optimality as axiomatic.

[26] We further elaborate on such Pareto-improving jumps as non-cooperative coordination games, demonstrating that undertaking the merger in these situations is a dominant strategy under subgame perfection.

[27] It also bears emphasis that a Pareto-improving jump is not feasible in the pre-merger phase of the game because the continuation values are the integrals of these expected future payoffs and as such are continuous. The locus of the continuation values before the jump must be coterminous with the target point of the jump, simply because the continuation value prior to the jump is the expected discounted value of the continuation value at the target point.

Table 1. Merger coordination game between firms

	Merge	Refuse
Merge	γ, γ	0, 0
Refuse	0, 0	0, 0

however the reward is higher if they coordinate on the better option. This is a variant of the so-called stag hunt game that was originally described by Jean Jacques Rousseau: if the two players coordinate on hunting stags their reward is large, and if they take the safe option and hunt hares they get a smaller reward (see, e.g., Chapter 1 of Fudenberg and Tirole, 1991). The key feature of the game is that is a pure coordination game, so no punishments are available for players who choose to hunt hares. Therefore hunting stags is an equilibrium but so is hunting hares. Our game is an extreme version of this game in that there is a zero reward or punishment for hunting hares.

This multiplicity of equilibria of the auxiliary game can be broken by extending the Nash equilibrium concept to subgame perfection. The subgame perfect approach focuses on the extensive form and examines dominant moves if play is sequential. The extensive form of the merger assurance game is depicted in Figure 7. Working from the end of the game with a move by firm 2, it is a dominant strategy to play Merge if player 1 has played Merge; in the penultimate stage where player 1 moves, it is then a dominant strategy to play Merge. Thus, {Merge, Merge} is the only equilibrium of the auxiliary game.[28]

This reasoning demonstrates that if the firms traverse the no-merger manifold to our arbitrary point P the firms then jump to the merger manifold. If we considered this as equilibrium play then the equilibrium manifold would terminate at our jump point, however it would then fail as an equilibrium in two senses: (i) it would fail to satisfy the main differential equation locally because at the jump point the differential equation is violated, and (ii) because the continuation values are defined as the integrals of future payoffs the jump point would incorrectly represent the discounted value of jumping to the merger.

In addition to these observations we can also note that our reasoning applies to any point on the no-merger manifold attained prior to our arbitrary jump point P, and so P will never be attained, as the jump will occur at an earlier point. Of course this reasoning applies at every point below the merger line, so no point below the merger line can survive as an equilibrium jumping off point.

By formally appending the assurance game as an auxiliary game at each point on the no-merger manifold and adopting subgame perfection as the equilibrium concept for this auxiliary game we then rule out the no-merger equilibrium. There is one additional detail however, which is that, given that no point on the no-merger manifold can survive a Pareto-improving jump, it is still possible for the upper part of the no-merger manifold to locally satisfy the main differential equation, and to terminate in a merger where it intersects the merger line. However it is clear that such a terminal point fails to satisfy the smooth pasting condition and as such is suboptimal for both firms. We can therefore rule out such candidate manifolds as equilibria.

We note that the subgame perfection mechanism that we have described here is the same underlying mechanism that we applied to rule out punishment manifolds in Appendix K, eliminating any candidate manifold in which the continuation values fall below the merger line.

[28] For further discussion see Fudenberg and Tirole (1991) chapter 3.

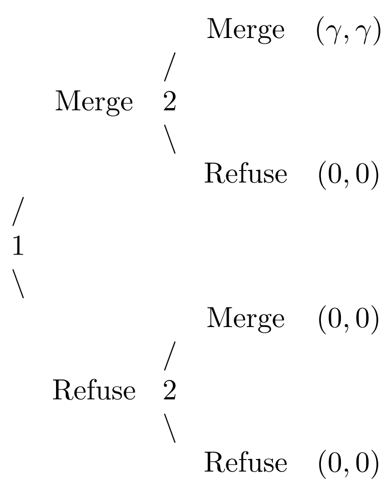


Figure 7. Extensive form of assurance game