CHAPTER TWELVE

CONCLUSIONS

Assessing the Impact of Law by Observing Judicial Behavior

China's laws are consistent with global legal norms that allow and even oblige courts to support unilateral divorce petitions from plaintiffs claiming to be victims of domestic violence. China's laws also include a competing and highly subjective legal standard – the breakdown of mutual affection – that Chinese judges routinely apply largely according to mutual consent and the lack thereof. Pressure from clogged courts, a political ideology hostile to divorce, and performance evaluation systems that reward judges for volume and efficiency and punish them for social unrest and "extreme incidents" compel judges to deny divorce petitions. The breakdownism divorce standard provides convenient support and justification for judges to do so. By routinely denying divorce petitions when plaintiffs file them for the first time, judges extend a judicial process of a few months into a litigation repeat that typically drags on for over a year. When plaintiffs return to court for another try, judges tend to grant their refiled petitions. This routine practice, which I call the divorce twofer because it rewards judges in several ways, prolongs abused women's exposure to the source of their abuse. Some women stay with their abusers until the divorce is finalized and child custody and property division are resolved. Some women escape by participating in labor migration or otherwise going into hiding. Women who leave their children behind in the process of fleeing abuse often, as a consequence, lose their claims for child custody.

In their struggle to divorce, women have often sacrificed child custody and marital property in exchange for their husbands' consent to divorce. Over 40 years ago, Mnookin and Kornhauser (1979) argued

https://doi.org/10.1017/9781108768177.013 Published online by Cambridge University Press

that courts in the United States, in a process they called "bargaining in the shadow of the law," set the terms for out-of-court negotiations in divorce disputes. Even though courts account for only a small share of all divorces in China, they cast a long shadow over the entire landscape of divorce. Judges have rarely granted divorce petitions when defendants withheld consent. Simply by withholding consent, spouses – even abusive spouses – have mostly thwarted first-attempt divorce petitions. In the Civil Affairs Administration, where most divorces are processed, mutual consent is a sine qua non of divorce. Spouses of divorce-seekers therefore wield enormous leverage over the terms of divorce in both forums simply by withholding consent. Defendants, most of whom are men, use their consent as a bargaining chip. Plaintiffs, most of whom are women, must then use marital property and child custody as bargaining chips. Once they secure their spouses' consent to the divorce itself and to all terms of the divorce, often at great cost, divorce-seekers can go either to the Civil Affairs Administration or to court. Perhaps for this reason, and because Civil Affairs divorces are quicker, easier, and cheaper, most divorce-seekers whose first-attempt petitions were denied in court do not return to court for another try. If divorce litigation were less restrictive and judges attached less importance to mutual consent, fewer women would settle for raw deals in the Civil Affairs system.

Divorce-seekers' bleak prospects of success in China's courts has extended seamlessly to victims of domestic violence. Simply put, domestic violence has been unimportant to judges. The inherent legal ambiguity and flexibility of the breakdownism divorce standard has helped judges sideline the legal relevance of domestic violence allegations. As a forum of last resort for victims of marital abuse, courts have generally done less to protect vulnerable women than to empower and enable their abusers.

Women's outcomes have been worse than men's in other respects, too. Gender bias animates every stage of the litigation process. Among plaintiffs filing for divorce, women have been less successful than men at ending their marriages on the first try. Women, often under duress, have been more likely than men to withdraw with petitions. Courts have been more likely to issue adjudicated denials to female plaintiffs than to male plaintiffs. When they do grant divorces, courts have been more likely to grant child custody to fathers than to mothers. By subverting laws designed to deliver gender justice, courts are themselves a mechanism of gender injustice.

In this book, I set out to document and explain decoupling in China's divorce courts. "Decoupling" refers not only to the efforts of divorce-seekers to decouple from their spouses. It also refers to a yawning and widening gap between, on the one hand, China's judicial practices and, on the other hand, its domestic laws and international legal commitments. Working in the long tradition of gap studies in the field of law and society (Gould and Barclay 2012), I have sought to understand the gap in China between the law on the books that supports divorce and the law in action that restricts divorce.

Although it is not supported by any law, the divorce twofer, as a highly institutionalized practice, has assumed a law-like and policy-like quality. Throughout this book I have referred to as "endogenous" the salient institutional forces competing with and undermining the exogenous force of law. The law in action is endogenously shaped in a couple of respects. First, judges' incentives to *deny* first-attempt divorce petitions are rooted in now-familiar local institutional influences on judicial decision-making that compete with and neutralize its domestic laws and international legal commitments calling on judges to *grant* first-attempt divorce petitions. Second, law is endogenous to organizational prerogatives and practices. As Lauren Edelman's (2016) legal endogeneity theory would lead us to expect, legal ambiguity gives China's basic-level courts "wide latitude to construct the meaning of law and compliance with law" (p. 14) and enables Chinese judges to interpret and apply the law in ways that are at odds with the law on the books. In contrast to views of law as "an exogenous, coercive, downward force on organizations" (Edelman 2016:41), legal endogeneity theory views law as endogenously created by the very organizations subject to its control.

Michael Lipsky's ([1980] 2010) classic theory of street-level bureaucracy provides a complementary framework for explaining how and why Chinese judges created bottom-up legal and policy substitutes for the top-down domestic laws and international legal commitments they sidelined and subverted. Just as legal ambiguity is at the center of Edelman's (2016) explanation for "the gap between ideal and actual" (p. 104), space for discretionary judgment is at the center of Lipsky's (2010) explanation for "the gap between the realities of practice and service ideals" (p. xvi), "the gap between policy as written and policy as performed" (p. xvii), and "the gap between public promises and performance" (p. 214).

CHINESE JUDGES AS STREET-LEVEL BUREAUCRATS

Street-level bureaucracies are public service agencies composed of a sizeable share of street-level bureaucrats, the key defining characteristics of whom include direct interaction with citizens seeking public services and a high degree of discretion in how they carry out their work (Lipsky 2010:3). As such, street-level bureaucrats are "frontline workers" serving as gatekeepers to state resources, services, and opportunities (Lipsky 2010; Maynard-Moody and Portillo 2010). Street-level bureaucracies include "the schools, police and welfare departments, lower courts, legal services offices, and other agencies whose workers interact with and have wide discretion over the dispensation of benefits or the allocation of public sanctions" (p. xi). Lipsky focuses on police officers, teachers, and social workers, but explicitly extends the scope of his theory to "judges, public lawyers and other court officers, and many other public employees who grant access to government programs and provide services within them" (p. 3).

As a consequence of their considerable discretion to interpret and implement rules, the individual decisions of street-level bureaucrats "become, or add up to, agency policy" (p. 3). By redefining law and policy, street-level bureaucrats decouple their decision-making from the judicial ideal. In a practical sense, therefore, street-level bureaucrats *make* law and policy. Lipsky (2010:24) calls them "de facto policy makers," while Maynard-Moody and Portillo (2010:260) call them the "ultimate policymakers."

In China, basic-level courts are quintessential street-level bureaucracies, and their frontline judges are quintessential street-level bureaucrats. China's judiciary is undifferentiated from the rest of the state bureaucracy. Although judges enjoy a real measure of decision-making autonomy in their everyday work, the judiciary as a whole is subordinate and beholden to the needs and interests of the party-state (Fu 2014; Kinkel 2015; Peerenboom 2010). As civil servants without tenure, judges are duty-bound to support the party-state and its political priorities. At the same time, however, owing to their heavy caseloads and the highly discretionary nature of their work, judges develop creative methods of complying – or faking compliance – with shifting mandates and directives from above (Li, Kocken, and van Rooij 2016; Paneyakh 2014). According to Lipsky (2010:xiii), "the decisions of street-level bureaucrats, the routines they establish, and the devices they invent to cope with uncertainties and work pressures, effectively

become the public policies they carry out" (emphasis in original). Judges' routine coping strategies, which so frequently diverge from the judicial ideal, "add up to street-level policy" (Lipsky 2010:86).

Judges in China's basic-level courts illustrate five defining hallmarks of street-level bureaucrats. First, owing to pervasive legal ambiguity, judges exercise enormous discretion at every stage of the civil litigation process. Chinese-language scholarship about China's civil justice system is replete with the words "discretionary" (裁量权), "subjective" (主观), and "arbitrary" (随意) for describing almost every type of judicial decision, and the words "flexible" (弹性 and 灵活) and "ambiguous" (笼统, 模糊, and 含糊) for describing decision-making rules. Throughout this book we have seen judges exercise discretion in the interpretation and implementation of ambiguous rules concerning: whether to issue a public notice to a defendant whose whereabouts are alleged to be unknown; whether to apply the simplified civil procedure (solo judge) or the ordinary civil procedure (collegial panel); whether to admit or exclude evidence; whether to affirm a litigant's claim of domestic violence on the basis of admitted evidence; whether to affirm a litigant's claim of a two-year physical separation; whether to affirm the breakdown of mutual affection; how to protect the best interests of the child in child custody determinations; and whether and how to broker informal compromises, settlements, and concessions between litigants. A judge in Anhui succinctly asserted that, compared to those who work on other kinds of civil cases, "Judges in domestic relations trials are distinguished by the relatively great deal of discretion they wield" (Zhou and Qiu 2018). Legal ambiguity allows judicial practice to decouple from laws that champion the freedom of divorce, gender equality, and the best interests of the child.

Second, the problem of "many cases, few judges" that characterizes China's court system extends to street-level bureaucracies everywhere. China's basic-level courts exemplify the universal challenges of public service agencies that face "extraordinary demand for resources relative to the supply" and are "under-staffed relative to the demands on them" (Lipsky 2010:132). "Street-level bureaucracies are labor-intensive in the extreme" (p. 5) because "demand for services is practically inexhaustible relative to supply" (p. 55). Toiling under the weight of heavy caseloads, street-level bureaucrats "cannot do the job according to the ideal conceptions of the practice" (Lipsky 2010:xvii). "Resources are chronically inadequate relative to the tasks workers are asked to perform" (p. 27n1) owing to "the scarcity of resources relative to the

demands made upon them" (p. 83). "Street-level bureaucrats characteristically have very large case loads relative to their responsibilities. The actual numbers are less important than the fact that they typically cannot fulfill their mandated responsibilities with such case loads" (p. 29). The working conditions of China's basic-level courts, like those of other street-level bureaucracies, lead to worker burnout and attrition (p. xv).

Third, owing to their heavy caseloads, frontline public service workers such as Chinese judges are too busy to give each individual case the full consideration demanded by formal procedures. "The fact that street-level bureaucrats must exercise discretion in processing large amounts of work with inadequate resources means that they must develop shortcuts and simplifications to cope with the press of responsibilities" (Lipsky 2010:18). Their innovative "coping behaviors" and "coping mechanisms" allow them to complete their tasks, albeit often in ways that are contrary to organizational goals and service ideals (Lipsky 2010:xvii–xviii). China's basic-level courts have coped with their crushing workloads by developing various methods of optimizing scarce judicial resources: deputizing assistant judges and putting them to work on a greater share of cases; assigning a greater share of cases to solo judges, which entailed increasing simplified civil procedure utilization rates; expanding the pool of lay assessors and increasing their rates of participation on collegial panels; and, of course, clamping down on divorce by denying a greater share of first-attempt divorce petitions. Such routines, simplifications, and shortcuts "did not merely facilitate work; they determined outcomes divergent from the stated policy objective" (Lipsky 2010:84). The right of abused women to divorce is a clearly "stated policy objective" of the Chinese party-state that is fully supported by law but undermined by judges' routine practice of denying first-attempt divorce petitions.

Adjudicated denials in first-attempt divorce cases have significantly lightened judges' workloads while simultaneously supporting China's political priorities of marital preservation and stability maintenance. For these reasons, judges have denied divorce petitions across the board, regardless of whether they involved statutory wrongdoing. Although a fault-based claim of domestic violence can and should be lawful grounds for divorce, divorce petitions involving such a claim – which accounted for about 30% of all first-attempt divorce petitions – were no more likely to be granted than those which did not involve such a claim. When judges denied divorce petitions, they frequently

and improperly misrepresented allegations of domestic violence as ordinary family conflict, mutual hitting, or aberrations from otherwise healthy marital behavior; cited abusers' contrition and unwillingness to divorce as evidence of mutual affection and reconciliation potential; and disaffirmed the admissibility and validity of supporting evidence. By doing so, judges routinely gaslighted domestic violence victims. Recasting as fundamentally healthy and hopeful – as a pretext for forcibly preserving – a toxically abusive marriage from which the victim seeks to exit is the very epitome of gaslighting.

Fourth, street-level bureaucrats triage and ration scarce resources by privileging help-seekers they deem to be more "deserving" and "worthy" of public services. They classify, sort, and differentiate applicants into categories of deservingness and worthiness shaped by cultural stereotypes (Lipsky 2010). Their "routines and simplifications are subject to workers' occupational and personal biases, including the prejudices that blatantly and subtly permeate society. The biases expressed in street-level work may be expected to be manifested in proportion to the freedom workers have in defining their work life and the slack in effective controls to suppress those biases" (Lipsky 2010:85). Chinese judges, like other street-level bureaucrats, make judgments about the credibility and truthfulness of applicants (Lipsky 2010:74), some of whom must bear "persistent assumptions of fraud and dishonesty" (p. 93). Empirical patterns presented throughout this book suggest that judges have regarded women (1) as more likely than men to file frivolous and impulsive divorce petitions, make bogus claims of the unknown whereabouts of the spouses they want to divorce, and exaggerate and fabricate claims of domestic violence, and therefore (2) as less deserving than men of divorce on the first attempt, the full due process afforded by the ordinary civil procedure, and child custody.

Fifth, street-level bureaucracies measure and evaluate the performance of their workers. Because the extent to which a worker's performance advances organizational goals generally eludes measurement, street-level bureaucracies set performance targets that are easier to measure. These "surrogate measures" take on a life of their own and "guide future performance" as street-level bureaucrats orient their behavior toward them (Lipsky 2010:52). In their efforts to maximize their professional rewards and minimize their professional sanctions, street-level bureaucrats carry out their work tasks according to how they are measured and evaluated. In the words of Lipsky (2010:51), "the behavior of workers comes to reflect the incentives and sanctions

implicit in those measurements" and "behavior in organizations tends to drift toward compatibility with the ways the organization is evaluated." Although an organizational goal of courts everywhere is to deliver justice according to the law, judicial performance in China is evaluated primarily according to measures of case processing efficiency and litigants' acceptance of outcomes.

Endogenous institutional dynamics – including the conditions of work in street-level bureaucracies – help explain the limited impact of exogenous laws and legal norms. Let us now consider the theoretical and methodological implications of institutional decoupling in China's divorce courts.

THEORETICAL IMPLICATIONS

My empirical findings on marital decoupling in China shed theoretical light on institutional decoupling – the extent to which and reasons why legal systems that bear the symbolic hallmarks of global legal norms subvert them in practice. Decades of research on local compliance with global norms documents a ubiquitous gap between doctrinal law and on-the-ground practices, and the ubiquity of hollow and symbolic commitments to world society values. Do previous explanations for the extent and character of decoupling between global legal commitments and local legal practices in general help us make sense of decoupling in the specific context of Chinese divorce litigation? China's ratification of CEDAW, which scholars argue helps explain international variation in the gap between promises and practices (Englehart and Miller 2014; Htun and Weldon 2018; Wang and Schofer 2018), appears to have had little impact on the protection of Chinese domestic violence victims unilaterally seeking divorce in court. Of course, we have no way of knowing whether the plight of Chinese women seeking freedom from their abusive husbands would have been even worse in its absence. Furthermore, divorce litigation is only one piece of the larger puzzle of gender violence, and perhaps China's ratification of CEDAW has helped women in other institutional contexts.

Can we make sense of China's divorce twofer as a function of bureaucratic capacity (Cole 2015)? While increasing the supply of judges and other court personnel, by alleviating some of the pressures of crushing dockets that incentivize judges to deny first-attempt divorce petitions, might help women seeking divorce, the institutional roots of the divorce twofer also lie with the endogenous institutional logics of a

political ideology of marital preservation and performance evaluation systems that incentivize adjudicated denials of divorce petitions. Can we attribute gender injustice in China's divorce courts to the absence of a strong and autonomous domestic feminist women's movement (Htun and Weldon 2018)? We can only speculate about the hypothetical ability of such a movement to diminish the impact of the patriarchal cultural beliefs and gender stereotypes shaping judges' decisions.

My Henan–Zhejiang comparison has illuminated contextual similarities and differences. The same gender injustices emerge in high relief from both provincial samples. Empirical patterns show remarkable consistency between the two provinces in the extent of and reasons for gender inequality in case outcomes. That the empirical patterns I presented are so consistent across two subnational contexts that differ in other ways suggests the power of the endogenous institutional logics at the heart of my argument.

With respect to regional differences, adjudicated divorce was far more restrictive in Zhejiang than in Henan. Zhejiang, as a coastal province adjacent to Shanghai in one of the most prosperous parts of China, is much more integrated with and proximate to world society than Henan is. Adherents of world society theory would have been hard-pressed to predict a judicial clampdown on divorce unfolding at precisely a place and time of intensifying global integration such as China in the mid-2000s. Nor would they have predicted the judicial clampdown on divorce to be earliest and most severe in an area of China closest to world society. Contrary to the expectations of world society theory, adjudicated divorce became increasingly difficult to obtain over time as court dockets across China swelled, and have been considerably more difficult to obtain in Zhejiang than in Henan in no small measure because court dockets have been so much heavier in Zhejiang. The contrast between Henan and Zhejiang in this regard also reflects a more general effect of urbanization: in both provinces, urban courts, which had heavier dockets than rural courts, were more inclined than rural courts to deny first-attempt divorce petitions.

At the same time, however, empirical findings partially support world society theory by revealing that gender injustices in both provinces were concentrated in rural areas further away from world society. In urban courts, female plaintiffs were at no disadvantage vis-à-vis male plaintiffs with respect to the probability of obtaining a divorce on the first try, and mothers enjoyed an advantage vis-à-vis fathers with respect to the probability of winning child custody when divorces

were granted. Judges' biases toward female litigants appear to diminish with urbanization. Over time, the ongoing process of urbanization may help to exert gender-equalizing pressures on divorce litigation practices. During the time period covered by the court decisions I analyze in this book, however, predominantly urban areas accounted for only a small share of all divorce adjudications.

In the context of Chinese divorce litigation, the direction of institutional "drift" (Schofer and Hironaka 2005) has been away from at least as much as toward global norms. The "paradox of empty promises" is not courts' incremental enforcement of international legal commitments that hitherto were widely ignored, as Hafner-Burton and Tsutsui (2005) might expect. On the contrary, it is the extent to which courts are a primary obstacle to the realization of the promises enshrined in laws intended to protect women's freedom of divorce and to offer relief to victims of marital violence. The real paradox is that Chinese judges are *expected* to subvert and stretch beyond recognition domestic laws that are "constructed out of a common and universalistic world cultural frame" (Boyle and Meyer 1998:214) and are *rewarded* for doing so. We cannot understand Chinese courts' routine failure to offer relief from domestic violence as merely a matter of compliance failure. On the contrary, we should understand the judicial practices I have documented in this book as compliance success. Courts' routine failure to protect victims of domestic violence is a function of *purposeful institutional design*.

Judges' seemingly limitless discretion to assess marital quality, reconciliation potential, and evidence submitted by litigants is hardly a problem of implementation, much less an unintended consequence of institutional design flaws. It is an institutional *feature* as much as it is an institutional *bug*. The logic of the breakdownism standard is its flexibility for allowing judges to apply it in ways that support prevailing political priorities and pragmatic needs. China's enduring political priority of preventing frivolous divorce has a long legacy and has taken on renewed urgency since 2012. Judges' discretionary application of breakdownism supports the political goals of family preservation and social stability. Indeed, according to the deputy chair of the committee responsible for drafting the 1980 Marriage Law, this was precisely the legislative intent of breakdownism, which "at once maintains the principle of freedom of marriage [which includes the freedom of divorce] and also gives the courts considerable latitude" (Huang 2005:187). From a pragmatic standpoint, an overly rigid system risks breaking.

Discretionary flexibility helps judges complete their tasks and reap the rewards of hitting performance targets. As Lipsky (2010:19) puts it, "Lower-level participants develop coping mechanisms contrary to an agency's policy but actually basic to its survival."

METHODOLOGICAL IMPLICATIONS

According to Pope and Meyer (2016:289–90), "Decoupling exists because world models, which are suffused with meaning and cultural significance, interface with situated interests and practical concerns at the local level (Meyer and Rowan, 1977). Local actors may adopt the models for reasons of external legitimacy, but buffer the models from daily practices to maintain internal technical efficiency or solidarity." A key methodological task, then, is properly identifying and measuring the relevant local "situated interests" and "practical concerns." From the existing literature on the topic, we should expect that the extent to which China's commitments to women's rights, including the right to divorce, are decoupled from on-the-ground practices is a function, above all, of its links to world society, its bureaucratic capacity, and the strength and autonomy of its domestic feminist women's movement (Cole 2015; Englehart and Miller 2014; Htun and Weldon 2018; Wang and Schofer 2018). But, of course, we cannot assess the relative importance of multiple and potentially contradictory norms in a given institutional context before knowing what they are and how they work in theory and practice. As Drori and Krücken (2009:20) put it: "The research methodologies common to world society theory have not allowed for specific findings that explain different degrees of coupling or pointed to the cultural and historical specificity of the determining societal context."

Insofar as a key objective of macro-comparative cross-national research on states' promises and practices is to assess the relative importance of endogenous and exogenous influences, essential methodological ingredients include appropriate measures of endogenous influences well attuned to local contexts. Only after first inductively ascertaining the endogenous norms and practices – legal and otherwise – that pertain to a specific context would we know what to compare with exogenous world society norms and practices (Fourcade and Savelsberg 2006; Hagan, Levi, and Ferrales 2006; Halliday and Carruthers 2009). Might empirical findings in the world society literature consistently showing stronger exogenous effects be an artifact of dominant approaches

to measuring endogenous influences? Any comparison of global and local effects will necessarily privilege the former if salient endogenous norms and practices obstructing the realization of exogenous institutional prescriptions are poorly measured or altogether omitted from the analysis. We will not find what we do not know to look for; we cannot assess what we do not know to include in the assessment. For example, in their cross-national research on divorce rates, Wang and Schofer (2018:20) find that global cultural norms valorizing "individual freedom, consent, and gender equality" trump local cultural and institutional barriers to divorce, which they measure as economic development, religious tradition, mass education enrollment (all of which theoretically drive cultural values conducive to divorce), and female labor force participation (which theoretically promotes women's financial wherewithal to divorce). China's perfect score of 3 out of 3 (over a period of almost 40 years) on a "divorce law equality" index (with higher scores meaning greater gender equality, constructed from components of Htun and Weldon's [2015] "family law index") belies the endogenous Chinese legal standard of breakdownism – altogether invisible in this scholarly literature – routinely used to deny first-attempt divorces to plaintiffs, and especially to female plaintiffs, particularly when they make claims of marital violence.

To be sure, China's rising divorce rates are consistent with Wang and Schofer's (2018:16) sanguine conclusion that "the legitimation of world cultural principles at the global level can propel local change." Global scripts, including "developmental idealism," may very well contribute to values of individualism, feminism, and equal rights in China and may thus influence *individual* behavior (Boyle, McMorris, and Mayra 2002; Thornton and Xie 2016; Yu and Xie 2015) in the context of marriage and divorce despite durable local *organizational* barriers to divorce rooted in endogenous institutional norms and practices that are orthogonal to world society models. Such local organizational barriers are the focus of this book. If we know what to look for and where to look for it, we will surely find similar decoupling processes in other contexts characterized by street-level bureaucrats who, for material, ideological, political, cultural, and cognitive reasons, faithfully enforce endogenous institutional norms hostile to the very elements of world society they simultaneously champion. Such decoupling processes are all too often obscured by the local embrace of world society norms captured in macro-comparative cross-national research because they are more conspicuous on the surface veneer of institutions.

A macro-comparative cross-national research design would have obscured our ability to discern these key local forces obstructing women's freedom of divorce in China. Macro-comparative cross-national generalizations come at the potential cost of missing the story in specific cases. Conversely, contextually specific explanations, even when they get the story right in a specific case, come at the potential cost of precluding generalizations. Even when they cannot be generalized, however, idiosyncratic stories can be of enormous theoretical utility insofar as they are exceptions that prove the rule or bring to the fore less common and therefore less conspicuous processes (Emigh 1997; Lieberman 2005; Pearce 2002). Macro-comparative cross-national research on the promises and practices of nation-states is self-avowedly and variously macro-sociological, macro-social, macro-cultural, macro-institutional, macro-structural, macro-historical, and macro-phenomenological (Bromley and Suárez 2013; Drori and Krücken 2009; Fourcade and Savelsberg 2006; Frank and Moss 2017; Hallett 2010; Meyer 2010; Pope and Meyer 2016; Schofer et al. 2012; Wotipka and Ramirez 2008). Scholars in this tradition have been forthright in acknowledging the methodological limitations of their approach. They are the first to admit that their images of the ground taken from the stratosphere are at best low resolution (e.g., Frank et al. 2009:279).

Country-level indicators of state responses to human rights and gender violence have gained currency and acquired legitimacy in policy and scholarly contexts. Ironically, such indicators are themselves institutionalized myths that are loosely coupled with what they ostensibly reflect. Human rights indicators not only shame autocrats by illuminating their human rights abuses, but also serve to buttress authoritarian regimes by obscuring their human rights violations (Zaloznaya and Hagan 2012). When social scientists use indicators to assess compliance with human rights treaties, they "transform a judgment-laden process into one that appeared technical, scientific, and therefore – in a context in which the treaty bodies' authority is often in doubt – more legitimate" (Rosga and Satterthwaite 2012:306). Standardized models for indicators and rankings have emerged and spread globally according to the same isomorphic pressures social scientists use such measures to study (Erkkilä and Piironen 2018; Merry 2016; Sauder and Espeland 2009).

Cases in point are the Cingranelli–Richards Human Rights Data Project's "physical integrity rights," "women's rights," and other related ordinal measures (Alexander and Welzel 2015; Cole 2012, 2015; Cole

and Ramirez 2013; Englehart and Miller 2014; Wei and Swiss 2020). (For a review and assessment of an array of gender equality measures, see Liebowitz and Zwingel [2014] and Sundström et al. [2017].) Hafner-Burton and Tsutsui's (2005) ordinal measure of government repression of human rights and Hathaway's "fair trial index" (2002) use many of the same sources, most notably US Department of State human rights reports, which have "come to play an outsized role in academic research on human rights and state repression, as commonly used sources of cross-national data on state behavior" (Bagozzi and Berliner 2018:663; also see Gallagher and Chuang 2012; Innes de Neufville 1986). In the absence of measurement validity assessments (but with the occasional ceremonial reference to intercoder reliability to enhance their legitimacy), we are asked to take these indicators at face value as objectively accurate and free of political bias. Of similarly dubious validity and inscrutable construction is the "Caprioli index of physical security of women" measuring violence against women and available in the WomanStats database (Hudson, Bowen, and Nielsen 2011). We have only an incomplete picture of the "over 500 sources" used (Caprioli et al. 2009:5), the role of country experts, and the precise criteria for coding laws as, for example, "generally enforced" or "rarely enforced." In the same vein are the "index of sex equality in family law" (Htun and Weldon 2015, 2018:127–32; Wang and Schofer 2018) and the "index of government response to violence against women" (Htun and Weldon 2018:31–50). China's score of 12 out of 13 in "sex equality in family law" – with points for all three divorce components (Wang and Schofer 2018) – belies a key finding reported in this book: the unlikely and worsening prospects of getting divorced in court on the first attempt and women's disproportionate challenge in this regard. China's score of 2 out of 10 in "government action on violence against women" (Htun and Weldon n.d.) may accurately capture the limited on-the-ground impact of its "policy regime" but misses the broad scope and large scale of its official policy responses, including legal reforms reviewed in Chapter 2.

Similar concerns can be raised about some of the independent variables used to explain these outcomes measures. If "bureaucratic efficacy" does indeed account for variation in country-level human rights enforcement, are the International Country Risk Guide indicators used by Cole (2015), which include opaque ordinal "bureaucracy quality" and "corruption" measures constructed by research staff, sufficiently sensitive to local context? Is either tax collected as a proportion of

GDP (Englehart and Miller 2014) or government consumption as a percentage of GDP (Wang and Schofer 2018) a valid measure of state capacity in general and in the context of women's rights enforcement in particular?[1] Do such measures capture relevant endogenous institutional forces as effectively as the measures of exogenous influences against which their effects are typically compared? If we have reason to believe that the gap between the promises and practices of law stems from the concrete working conditions of street-level bureaucrats responsible for the disposition of justice, why would we assess the effects of measures of dubious relevance such as government tax revenue and consumption instead of relevant measures such as pressures from political ideology, caseloads, and performance targets? Instead of measuring what is conveniently available, should we not instead measure what we inductively ascertain from deep contextual knowledge to be salient and relevant?

As another case in point, does a code of 0 out of 2 for China in both the "strength" and "autonomy" of its feminist women's movement – constructed largely on the basis of "country and region-specific expert opinion" – sufficiently capture a relevant endogenous determinant of the effectiveness of women's divorce rights and legal protections against violence (Htun and Weldon 2018:52–58) that can be meaningfully compared with the exogenous influence of global legal norms as measured by CEDAW ratification (Htun and Weldon 2018:62–63)? Does a code for China's apparent establishment of a "low-level gender mainstreaming institution" in 1992 adequately capture the effect of endogenous "women's policy machinery" effectiveness on state responses to various dimensions of women's rights, including protection against violence through 2005 (Htun and Weldon [2018:59, 232], using codes from True and Mintrom [2001])? Scholars who use measures like these have not systematically confronted and grappled with such questions. Beyond raising measurement issues such as these, scholars have also critiqued dominant modelling strategies in this literature (Hug and Wegmann 2016).

In fairness, however, scholars who develop and use global indicators can hardly be faulted for their efforts to surmount the methodological

[1] Yasuda (2017:16–17) shows that state capacity measured as revenue as a proportion of GDP is a poor predictor of China's food safety record and introduces an alternative measure that takes country-level scale challenges into account. Lieberman (2002) and Hendrix (2010) assess the strengths and weaknesses of tax-related proxy measures of state capacity so widely used in macro-comparative cross-national research.

challenges inherent to macro-comparative cross-national research designs. "Lowest common denominator" strategies of using measures plausibly connected to the theorized mechanism of interest and available for the maximum number of cases are good-faith and unavoidable last resorts in analyses of pooled cross-sectional data from large samples of countries over long periods of time. Additional methodological strategies include the use of survey data of either general populations or sub-populations such as business managers about perceptions of, trust in, and experiences with institutions as proxies for institutional performance, including corruption. However, owing to one-size-fits-all instruments, perception bias, and desirability bias, the data they capture do not always reflect salient local institutional norms and practices (see Y. Wang [2013:108–9] for a review of such approaches). For example, scores and rankings of countries according to the World Justice Project's "rule of law index" (Urueña 2015) omit contextually specific endogenous institutional legacies such as mediation practices, which continue to dominate the Chinese justice system, including the courts (Huang 2016; Liu 2006).

A related methodological approach to cross-national comparisons of legal systems is to analyze vignette data (e.g., answers to questions about hypothetical disputes) collected by the Lex Mundi Project from lawyers around the world (Negro and Longhofer 2018). There is even a tradition of analyzing vignette data within the field of China studies to assess variation in court performance (Gallagher 2017; Gallagher and Yang 2017; Y. Wang 2013, 2014).[2]

A further limitation of macro-comparative cross-national research is its tendency to obscure subnational variation in institutional behavior by taking the country-year as the unit of analysis (Berkovitch and Gordon 2016). A research design limited to urban courts, which handle only a relatively small share of divorce litigation, would fail

[2] Yuhua Wang's (2013, 2014) indirect proxy measure of judicial corruption exemplifies the questionable validity of some prevailing measures of court performance. Using data from a 2003 nationally representative survey of the general population in 102 counties across China, he measures local courts' levels of corruption as the proportion of a small subsample of respondents in each locale who chose the category "courts are corrupt" as their answer to why they would not go to court in the hypothetical instance of a dispute. Leaving aside the issue of whether general perceptions of corruption are valid measures of actual corruption, a perhaps more critical issue is that, owing to complex skip patterns, respondents eligible to answer this question were limited to those who indicated a willingness to pursue the resolution of a hypothetical dispute in the first place and who then indicated an unwillingness to go to court. In other words, respondents who indicated they would not pursue any form of resolution or who indicated they would seek help in court – who together form a sizeable chunk of the sample – were removed from the pool of respondents asked about their perceptions of court corruption.

to reveal the important stories at the heart of this book. Failing to disaggregate urban and rural China, or focusing only on urban China, would thus limit our ability to identify key institutional forces animating divorce litigation.

In this book, rather than relying on indirect proxy measures of the gap between judicial promises and judicial practices, I scrutinized the actual behavior of over 250 courts in two provinces, directly observed the incidence of their use of competing legal standards in real-life rulings, and empirically assessed and explained the differential impact they have on female and male litigants. There is nothing abstract about these measures. They do not come from country experts, nor do they come from readings of US government reports about the behavior of Chinese courts. A court decision to deny a female plaintiff's first-attempt divorce petition means a real woman was unable to divorce (for at least some period of time). Had I measured China's "trial fairness" according to its formal laws and international treaty commitments summarized by the US State Department in its human rights reports (Hathaway 2002), I would have arrived at very different conclusions. Had I assessed Chinese court behavior according to vignettes presented to lawyers or to ordinary citizens (most of whom had never been to court; Negro and Longhofer 2018; Y. Wang 2014), or according to panels of regional and country "experts" (Cole 2012, 2015; Hafner-Burton and Tsutsui 2005; Htun and Weldon 2018), my conclusions would likewise have been different. Contrary to much of the literature on the diffusion and local penetration of global legal norms, my conclusions about judicial behavior and trial fairness come from empirical analyses of actual judicial behavior and actual trials.

Real decisions from real courts show both that, over time, adjudicated divorce became increasingly difficult in general and was disproportionately difficult for women in particular. My empirical findings show that the wide gender gap in the probability of getting an adjudicated divorce on the first attempt is explained in large measure by five correspondingly wide gender gaps in (1) the incidence of plaintiffs with domestic violence claims, (2) the incidence of plaintiffs whose spouses withheld consent, (3) judges' responses to plaintiffs' claims of domestic violence, (4) judges' responses to plaintiffs' claims of missing spouses, and (5) judges' responses to defendants' failure to appear in court for other reasons.

We must remain mindful of the substantive trade-offs of our methodological choices for other reasons, too. In-depth case studies such

as this one are, by definition, poorly suited for generalizable research spanning wide swaths of time and place. They are also relatively narrow in the scope of the institutional issues they can address. In more concrete terms, the broader institutional issue of gender justice cannot be reduced to the specific issue of divorce practices in lower civil courts, the empirical focus of this book. Although this is a study of only one narrow slice of gender justice in one country, it provides a critical test for theories of local compliance with global norms. Is there a more likely place than the court system for the implementation of domestic laws such as China's that are so consistent with global legal norms? If a particular set of global legal norms embedded in domestic laws generally fails to penetrate the courts, perhaps we should harbor doubts about the prospects of world society penetration in other organizational contexts.

As the world's most populous country, China exerts a profound influence on the *worldwide* extent of world society penetration. All efforts to identify country-level determinants of compliance with global norms have treated small and large countries equally. No study of which I am aware paints a global portrait of the worldwide extent of local compliance with a given set of global norms. We know, for example, that rape-law reform is positively correlated with police reports of rape at the level of the country-year (Frank, Hardinge, and Wosick-Correa 2009). But we have no idea about the worldwide impact of rape-law reform. Likewise, studies of country-year correlates of the implementation of human rights (Cole 2015; Hafner-Burton and Tsutsui 2005) and women's rights (Htun and Weldon 2018) tell us little about their worldwide impact.

Two examples illustrate this point. First, insofar as rising levels of *income inequality* in most of the world coincides with rising levels of *income* in the poorest and most populous countries in the world, namely, China and India, treating all countries equally or weighting countries according to their populations lead to diametrically opposing conclusions about worldwide levels of income inequality. Either omitting China and India from the analysis or ignoring country-level populations shows an aggregate worldwide *increase* in income inequality, whereas including these two countries in a population-weighted analysis shows an aggregate worldwide *decrease* in income inequality (Firebaugh and Goesling 2004; Hung and Kucinskas 2011). Similarly, the choice to include or omit China from a pooled cross-national analysis of the feminization of legal professions would have dramatic consequences for our substantive conclusions about worldwide levels of

https://doi.org/10.1017/9781108768177.013 Published online by Cambridge University Press

lawyer feminization. Taking the unweighted country-year as the unit of analysis would vastly exaggerate lawyer feminization because doing so would treat small and large countries equally. Unusually low lawyer feminization levels in China and India suppressed the global impact of the rapid feminization of bars elsewhere in the world (Michelson 2013:1087, 1097). Although lawyer feminization has taken hold in a lot of countries around the world, results from a population-weighted analysis that includes China and India support the less-than-sanguine conclusion that "from a global perspective, the process of lawyer feminization has hardly begun" (Michelson 2013:1101). In short, the worldwide penetration of world society hinges to an important degree on the penetration of world society in China. World society theory will need to reckon with China.

WHITHER THE IMPACT OF CHINA'S 2015 ANTI-DOMESTIC VIOLENCE LAW?

The vast majority of court decisions in my collections predate China's 2015 Anti-Domestic Violence Law (Chapter 2). Nonetheless, we can look for early clues of its impact in the portion of decisions in my Zhejiang sample that were made after this new law took effect on March 1, 2016. Perhaps the grim picture I paint in this book began to change. Although all the decisions in my Henan sample predate the implementation of this law, 10,501 adjudicated divorce decisions in my full Zhejiang sample were made on or after March 1, 2016. In the solitary decision (one out of 10,501) in which this new body of law is cited, the plaintiff claimed to have been cut and injured in a knife attack by the defendant. In its written decision, the court cited this new body of law to justify denying the plaintiff's petition on the grounds that the plaintiff waited two years to file for divorce following the alleged attack and failed to submit evidence proving "frequent beatings and other violent behavior," and, above all, that the defendant both denied the plaintiff's claim of abuse and was unwilling to divorce (Decision #4687109, Haiyan County People's Court, Zhejiang Province, September 1, 2016).[3] A Chinese report cited by Amnesty International was similarly discouraging:

> 10 months after the enactment of the [Anti-Domestic Violence] law, of the 142 abuse-related divorce cases in the city of Jinan, only 14 cases

[3] Case ID (2016)浙0424民初2645号, archived at https://perma.cc/X876-G25A.

> were allowed to get divorced [sic]. The reason these 14 cases were successful were invariably the same [sic]: the accused admitted to abusing the victim. In the rest of the cases, failure was also invariably due to the same reason: the accused denied allegations of domestic violence, and judges deemed the cases to have insufficient proof. (Lu 2018)

The enactment of this special body of law, so far at least, has apparently done more to signal a symbolic commitment to combatting domestic violence than to change judicial practices on the ground.

One way to assess the impact of this law is to assess the effectiveness of the system of personal safety protection orders it formalized (人身安全保护令; Chapter 2). Scholars have lamented the small number of applications for personal protection orders, low approval rates, and ineffective enforcement (J. Jiang 2019; Y. Jiang 2019). Many judges apply excessively high standards of proof to personal protection order applications even though the evidentiary standards for proving domestic violence are laxer for personal protection orders than for divorces (Du 2018:8). Many judges, rather than issuing protection orders, conduct mediation with the goal of persuading applicants to withdraw their requests (J. Jiang 2019). According to an SPC report, China's courts received 5,860 applications for personal protection orders and approved 3,718 of them between the time the Anti-Domestic Violence Law took effect in 2016 through the end of 2018 (Equality 2019). According to another government report, China's courts issued 5,749 personal protection orders through the end of 2019 (Equality 2020). These are paltry numbers considering the prevalence of domestic violence in such a large population.

Insofar as it represents a court's affirmation of domestic violence, judges should treat a personal protection order as proof of statutory wrongdoing in a divorce case. A decision to grant a divorce should thus be a no-brainer when a plaintiff submits a personal protection order as evidence of domestic violence. This is not so, however. Chapter 7 contains a case example in which a court, in its decision to deny a woman's divorce petition, disregarded a personal protection order it had issued to her only a week earlier. Chapter 3 also describes the Sisyphean plight of Ning Shunhua, whose divorce petitions were repeatedly denied after the same court had repeatedly granted her applications for protection orders. This is not an uncommon pattern. In another case, a court issued a personal protection order to a plaintiff on January 20, 2016, three days after she filed for divorce. On February 22, 2016, during the trial, the defendant stated to the court, "The plaintiff's claim

of domestic violence is not factual. I never hit the plaintiff. The only time I ever hit her is when she made a date with another man and verbally provoked me." The court affirmed the following facts: "Beginning in August 2015, owing to the plaintiff's failure to return to and reside at home, the defendant's stalking the plaintiff, and other reasons, the two sides once again got into a fight, which led to physical conflict." The presiding judge, as judges so typically do, denied the plaintiff's divorce petition after representing her allegations of domestic violence as mutual fighting and holding that she had failed to provide evidence of the breakdown of mutual affection (Decision #4151585, Cixi Municipal People's Court, Zhejiang Province, February 22, 2016).[4]

To the extent that judges' tendency to ignore domestic violence in their haste to deny divorce petitions stems from their discretion and overwork, a reduction in either or both could shift their decision-making incentives to the benefit of vulnerable women. Likewise, to the extent that courts are sensitive and responsive to public outrage (Chapters 2 and 9), ongoing academic, journalistic, and public advocacy efforts to heighten public awareness within China of gender injustice in its divorce courts may also incentivize risk-averse judges to apply the law more faithfully and equitably in support of vulnerable women. Public sympathy for women who kill their abusive husbands can result in sentencing leniency (Chapter 9). Perhaps public sympathy for women seeking to divorce their abusive husbands in court can similarly result in a narrowing of the gap between legal promises and practices. Because the Anti-Domestic Violence Law was implemented only shortly before all courts were prohibited from publishing divorce decisions online in October 2016, we will need another source of data to assess its impact henceforth.

FINAL THOUGHTS

This book has chronicled the Sisyphean struggle of contested divorce in China, identified the institutional sources of this struggle, and assessed the extent of gender inequality with respect to outcomes of this struggle. It has documented the extent to which and provided reasons why women have borne the brunt of Chinese courts' clampdown on adjudicated divorce. Generally speaking, divorce is readily attainable outside the court system if both sides are willing and can agree on all terms.

[4] Case ID (2016)浙0282民初00647号, archived at https://perma.cc/H7SS-K55W.

Courts are the only place in China to which people can take contested, unilateral, ex parte divorces. China's divorce laws on the books provide strong protections to women seeking divorce. Chinese courts, however, routinely stretch these laws beyond recognition or altogether ignore them, and in so doing subvert China's own laws and international legal commitments. The evidence is clear: an apparent claim of domestic violence has no meaningful influence on whether a court grants an adjudicated divorce. In China's divorce courts, domestic laws and global legal norms concerning violence against women have been sidelined to the point of irrelevance. By privileging competing institutional imperatives, including judicial efficiency, the preservation of marriages, and social stability maintenance, courts serve the needs of political priorities more than the needs of gender justice. As political pressure to preserve marriages has grown, so too has courts' tendency to deny divorce petitions, even when – or especially when – they include claims of domestic violence. Just as proponents of the promarriage movement in the United States, in their efforts to reduce divorce, have ignored and obscured the pervasiveness of domestic violence (Catlett and Artis 2004), China's ideology of marital preservation undermines officially proclaimed commitments to combatting domestic violence. In prevailing political discourse, marital preservation serves the party-state's larger goal of social stability maintenance (Chapter 3; also see Wang 2020). By forcibly prolonging marriages, however, judges have enabled the persistence of domestic violence, which has sometimes escalated to suicide and homicide. Policies intended to promote social harmony and stability have therefore yielded unintended opposite effects.

Chinese judges deny first-attempt divorce requests for fear that plaintiffs, particularly female plaintiffs, embellish and lie; for fear that the approval process will slow the rate at which they clear cases; and for fear that angry husbands will retaliate, resulting in "extreme incidents" of social unrest. Undoubtedly, some plaintiffs do exaggerate and altogether fabricate their claims. The degree to which judges believe they do so, however, varies by plaintiff sex. Judges, biased by gender stereotypes, give greater credence to the claims of male plaintiffs and attach greater value to the rights of male defendants. When ruling on first-attempt divorce petitions, judges seem far more fearful of supporting a case without merit than denying a case with merit. On the whole, they would rather send a woman home with her abuser or force her into hiding than to grant a divorce to a woman who wants out

https://doi.org/10.1017/9781108768177.013 Published online by Cambridge University Press

of an unhappy marriage and may have thought spinning a poignant story about abuse would improve her chances of achieving her goal. Research on the veracity of women's domestic violence claims in the United States suggests that false negatives (true cases of domestic violence unrecognized as such) far outnumber false positives (fake cases of domestic violence falsely recognized as true). Whereas men's allegations of domestic violence are often false, women's are rarely exaggerated or fabricated (Haselschwerdt, Hardesty, and Hans 2011:1705–6, cited in Jeffries 2016:10; Jaffe et al. 2008:508). Even if some plaintiffs (however few) lie in court proceedings about abuse or the whereabouts of their spouses, is it better for judges to preserve the marriage than to dissolve it? Is the judicial error of dissolving the marriage of an unhappy woman who may have lied about or exaggerated abuse claims worse than the judicial error of exposing a battered woman to ongoing abuse by prolonging her marriage against her wishes? Would it not be better to err on the side of protecting women? Only exceedingly rarely have judges availed themselves of applicable evidentiary standards allowing them to give women who make allegations of domestic abuse the benefit of the doubt.

If, in one year alone, two-thirds of China's half a million plaintiffs in adjudicated divorce decisions are women, 40% of them experience domestic violence, and 70% of their petitions are denied, then over 90,000 female abuse victims seeking divorce remain exposed to their abusive husbands, typically for an additional year. The social and public health implications are palpable: China's institutionalized norm of denying a divorce request on the first attempt has spawned a sizable population of female marital violence refugees. In Henan and Zhejiang alone, thousands of women awaiting a second or third chance for an adjudicated divorce must choose from an array of similarly horrific options: further subjection to marital violence; separation from children, aging parents, and other kin while eking out an existence a safer distance from abusive husbands; and financial vulnerability and loss of child custody from concessions made to secure a divorce either in court or the Civil Affairs Administration.

In the grand scheme of domestic violence prevention in China, a 2015 criminal justice reform that led to greater leniency in the sentences of women who killed their abusive husbands (Chapter 9) is a clear case of "too little, too late." Women helped by criminal justice reforms are far outnumbered by women harmed by the divorce twofer. By the time domestic violence reaches the criminal courts, it has passed

the point of no return. If China's leaders were serious about helping battered women, they would do their utmost to prevent domestic violence from escalating to the point of becoming matters for criminal courts. Vulnerable women seeking to divorce their abusive husbands would be better served by public authorities, including judges, who believe their allegations and provide effective intervention much further upstream. Until then, the Chinese women most in need of court protection are the least likely to get it.

https://doi.org/10.1017/9781108768177.013 Published online by Cambridge University Press