Supplementary methods S1: Assignment of protection categories to plots.

1. Plots located within a designated protected area

If an FIA field plot was located within an IUCN-labeled polygon on the map of protected areas [39] then the indicated IUCN protection category was assigned to that plot. Full versions of the following truncated descriptions of each IUCN category may be accessed at https://www.iucn.org/theme/protected-areas/about/protected-areas-categories.

> *Ia. Strict Nature Reserve:* Category Ia are strictly protected areas set aside to protect biodiversity and also possibly geological/geomorphical features, where human visitation, use and impacts are strictly controlled and limited to ensure protection of the conservation values. Such protected areas can serve as indispensable reference areas for scientific research and monitoring ...
>
> *Ib. Wilderness Area:* Category Ib protected areas are usually large unmodified or slightly modified areas, retaining their natural character and influence without permanent or significant human habitation, which are protected and managed so as to preserve their natural condition...
>
> *II. National Park:* Category II protected areas are large natural or near natural areas set aside to protect large-scale ecological processes, along with the complement of species and ecosystems characteristic of the area, which also provide a foundation for environmentally and culturally compatible, spiritual, scientific, educational, recreational, and visitor opportunities...
>
> *III. Natural Monument or Feature:* Category III protected areas are set aside to protect a specific natural monument, which can be a landform, sea mount, submarine cavern, geological feature such as a cave or even a living feature such as an ancient grove. They are generally quite small protected areas and often have high visitor value...
>
> *IV. Habitat/Species Management Area:* Category IV protected areas aim to protect particular species or habitats and management reflects this priority. Many Category IV protected areas will need regular, active interventions to address the requirements of particular specis or to maintain habitats, but this is not a requirement of the category...
>
> *V. Protected Landscape/ Seascape:* A protected area where the interaction of people and nature over time has produced an area of distinct character with significant, ecological, biological, cultural and scenic value: and where safeguarding the integrity of this interaction is vital to protecting and sustaining the area and its associated nature conservation and other values...
>
> *VI. Protected area with sustainable use of natural resources:* Category VI protected areas conserve ecosystems and habitats together with associated cultural values and traditional natural resource management systems. They are generally large, with most of the area in a natural condition, where a proportion is under sustainable natural resource management and where low-level non-industrial use of natural resources compatible with nature conservation is seen as one of the main aims of the area...

2. Plots not located within a designated protected area

If the CBI IUCN protection category was "unassigned," or if the plot was not contained in a CBI polygon, then the protection category was assigned from the actual FIA ownership type [38] for that plot as follows.

de facto Federal (FIA owner codes 11 – 25). Federal government agency public land.

de facto nonfederal (FIA owner codes 26 – 33). State or local government public land.

Unprotected family (FIA code 45). Includes private land owned by an individual, family, trust, estate, or family partnership.

Unprotected nonfamily (FIA codes 41 – 44). Includes land owned by a corporation, non-governmental or natural resources organization, unincorporated local organization (partnership, association, club), or Native American entity.

Note: For some of the CBI polygons, the conventional name of the IUCN category does not accurately reflect actual ownership. For example, the Adirondack and Baxter State parks are labeled as National parks in the CBI database.

Supplementary methods S2: Invasibility model.

We modified an earlier model [28] to model invasive exotic species presence/absence using field observations from 23,039 FIA plots (one observation per plot) that were surveyed between 2001 and 2011 [35-36,38]. Invasive plant "presence" was defined as the occurrence of at least one invasive plant at a given plot location. FIA defines invasive plants as exotic plant species of any growth form that are likely to cause economic or environmental harm [37]. Invasive plant observations were initiated on FIA plots in 2001, but not all plots in the eastern United States have been surveyed for invasive plants. As a result, the surveyed plots do not by themselves constitute a statistical basis for regional comparisons of protected area effectiveness.

Invasibility was based on the presence of any inventoried plant species because, insofar as invasive exotic plants are an indicator of unnatural species assemblages, the specific identity of a species which has invaded is relatively unimportant. The independent variables in a logistic regression model included ecological province, site productivity, distance from a road, forest cover fragmentation, and land use (Table S2.1). The earlier study [28] used the same variables, but focused on clarifying the main effects of each variable. In this study, to improve per-plot prediction accuracy we accounted for a large number of interactions by nesting the independent variables representing productivity, forest fragmentation, distance from a road, and land use within ecological province. Since the revised model differs substantially from the earlier model, it is necessary to describe the statistical results for the modified model even if the overall model results and interpretations are similar.

At least one invasive forest plant was observed on 58% of the 23,039 plots that were used to fit the logistic regression model. The observed percentage of invaded plots ranged from 8% in province M211 to 78% in province 231 (Table S2.2). The observed percentage of invaded plots increased with site productivity, fragmentation, and anthropogenic land use, and decreased with distance from a road (Table S2.3 and Figure S2.1). The fitted model was judged adequate for the purpose of comparing the forest invasibility among protection categories. The Wald X^2 test was significant ($p < 0.0001$) for each independent variable (Table S2.4). The partition table for the Hosmer and Lemeshow goodness-of-fit test [51] showed close agreement between observed and expected numbers of invaded and uninvaded plots (Table S2.5), and that test did not indicate a need for additional interaction terms ($X^2 = 6.6$; $p = 0.58$). The area under the receiver operating characteristic (ROC) curve (Figure S2.2) indicated that the model correctly classified randomly drawn pairs of invaded and uninvaded plots 76% of the time, a reasonably good rate that was significantly ($X^2 = 232$; $p < 0.001$) better than chance. At a broader spatial scale, the regional pattern of predicted, per-plot probability of invasion was similar to the pattern of observed per-county invasion rates [27] (Figure S2.3). All statistical analyses were conducted with SAS© software [52].

Table S2.1. Variables used in the logistic regression model to estimate invasibility (the probability that a forest plot was invaded).

Variable	Definition
Ecological province	Categorical variable derived by spatial overlay of plot locations and the ecoregions [32]. The thirteen classes are the provinces shown in Figure 1 of the main text.
Productivity	Categorical variable derived by condensing the FIA site class ("site index") code [38] to three productivity classes of high (FIA codes 1 and 2), medium (FIA codes 3, 4, and 5), and low (FIA codes 6 and 7).
Forest fragmentation	Categorical variable derived by condensing the value of forest area density (proportion of forest land cover) [53] in a 15.2 ha neighborhood (FAD) to three fragmentation classes of low (FAD ≥ 0.9), medium (0.4 ≤ FAD < 0.9), and high (FAD < 0.4). FAD was derived from analysis of the 2006 National Land Cover Database [54].
Land use	Categorical variable derived from the classification of landscape mosaic (a ternary classification from the proportions of agriculture, developed, and other land cover) [55] in a 590.49 ha neighborhood. The four condensed land use classes are called natural (< 10% each of agriculture and developed land cover), agriculture (≥ 10% agriculture and < 10% developed), developed (≥ 10% developed and < 10% agriculture), and agriculture & developed (≥ 10 % each of agriculture and developed). Land use was derived from analysis of the 2006 National Land Cover Database [54].
Distance from a road	Continuous variable derived by overlaying plot locations on road maps [56] and measuring the distance (m) from each plot to the nearest road of any type.

Table S2.2. Number of plots and observed percent of plots invaded, by ecological province.

Ecological province	Number of plots	Percent invaded[1]
		%
211	163	42
212	455	30
221	899	65
222	201	78
223	1,376	71
231	8,042	78
232	8,276	46
234	779	43
255	482	45
M211	158	8
M221	1,396	37
M223	255	30
M231	557	40
All	23,039	58

[1]The X^2 test rejected the null hypothesis that the rate of invasion was the same for all provinces (df = 12; $p < 0.0001$).

Table S2.3. Distribution of 23,039 sample plots and observed percent of plots invaded by productivity class, fragmentation class, and land use class.

Variable	Class	Distribution of sample plots	Percent invaded[1]
		%	%
Land use	Natural	49	46
	Developed	6	62
	Agriculture	39	69
	Agriculture & developed	7	80
Productivity	High	5	76
	Medium	79	61
	Low	16	40
Forest fragmentation	High	14	65
	Medium	49	65
	Low	37	46

[1]X^2 tests rejected the null hypotheses that the percent of invaded plots was the same for all classes within each variable (df = 2 for productivity and fragmentation; df = 3 for land use; $p < 0.0001$ for all three tests).

Table S2.4. Type 3 analysis of effects from the fitted logistic regression model of the probability that a plot was invaded. Parentheses indicate the nesting of effects. .

Effect	df	Wald X^2	*p*
Province	12	147	< 0.0001
Land use (Province)	39	657	< 0.0001
Productivity (Province)	21	504	< 0.0001
Fragmentation (Province)	26	292	< 0.0001
Distance from a road (Province)	13	88	< 0.0001

Table S2.5. The partition table for the Hosmer and Lemeshow goodness-of-fit test for the fitted logistic regression model of forest invasibility. The partition table shows the observed and expected numbers of invaded and uninvaded plots for deciles of plots sorted by increasing estimated probability.

	Invaded		Uninvaded	
Group[1]	Observed	Expected	Observed	Expected
1	446	444	1,858	1,860
2	773	752	1,531	1,552
3	883	916	1,421	1,388
4	1,082	1,083	1,222	1,221
5	1,315	1,330	989	974
6	1,478	1,447	826	857
7	1,635	1,622	669	683
8	1,784	1,804	520	500
9	1,946	1,950	358	354
10	2,067	2,061	236	243

[1]n = 2,304 for groups 1 through 9 and 2,303 for group 10.

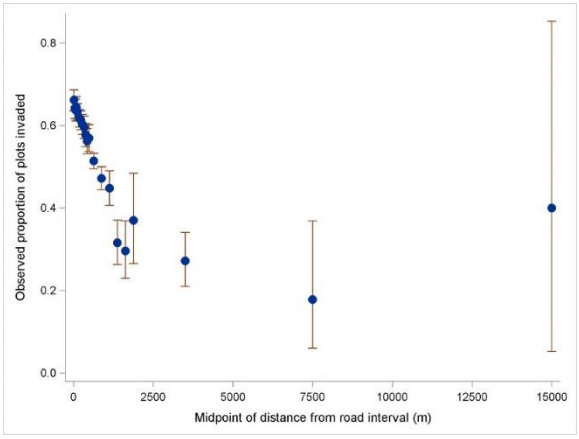


Figure S2.1. Observed proportion of plots invaded in relation to distance from a road. The plots were grouped into distance intervals and the proportion was calculated for the subset of plots in each interval. The upper and lower 95% confidence limits are indicated.

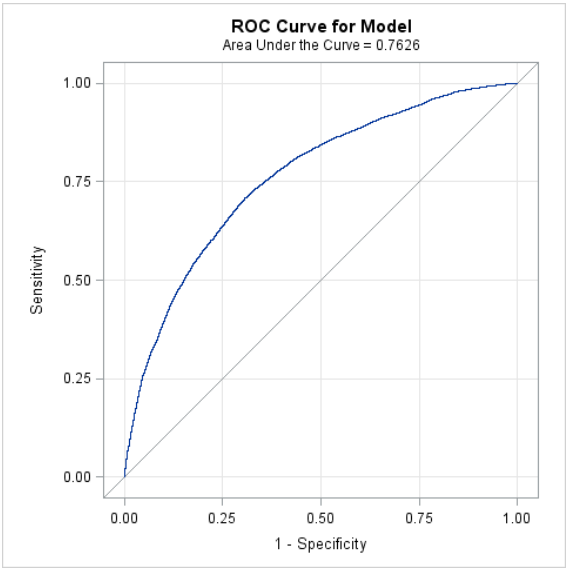
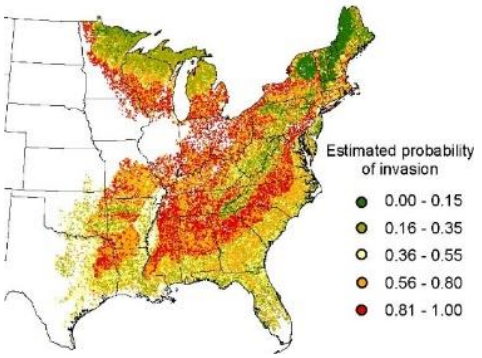
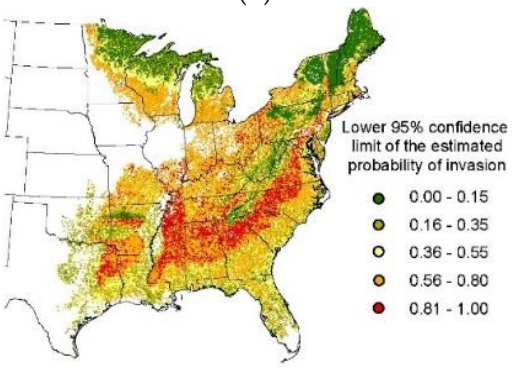


Figure S2.2. The receiver operating characteristic (ROC) curve for the fitted logistic regression model of the probability that a plot was invaded. The ROC estimate (0.76; SE = 0.003) indicates the model correctly classifies randomly drawn pairs of invaded and uninvaded plots 76% of the time, a rate that is significantly (X^2 = 232; $p < 0.001$) better than the null hypothesis indicated by the diagonal line in the chart.

(**a**)



(**b**)



(**c**)

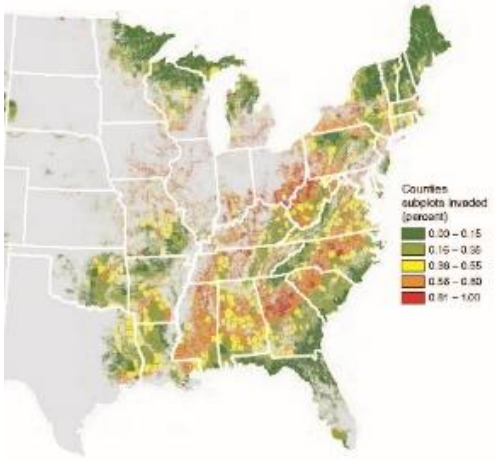


Figure S2.3. Comparisons of the regional pattern of estimated plot-level probability that a plot was invaded (**a**, **b**) with observed percentages of plots invaded by county (**c**). Note: (**c**) is reproduced from the public domain Figure 8.3 in [27]; the original caption is "National map showing percent of forested subplots with at least one nonnative invasive plant, calculated at the county level. Forest/nonforest mask applied to the conterminous United States.".

References

51. Hosmer, D.W.; Lemeshow, S. *Applied Logistic Regression*; John Wiley & Sons: New York, NY, USA, 1989.
52. SAS. *SAS/STAT Software, Version 9.4 of the SAS System for Windows*; Copyright© 2012 SAS Institute Inc.: Cary, NC, USA, 2012.

53. Riitters, K.H.; Wickham, J.D.; O'Neill, R.V.; Jones, K.B.; Smith, E.R.; Coulston, J.W.; Wade, T.G.; Smith, J.H. Fragmentation of continental United States forests. *Ecosystems* **2002**, *5*, 815–822.
54. USGS (United States Geological Survey). *NLCD 2006 Land Cover (2011 Edition)*; U.S. Geological Survey: Sioux Falls, SD, USA, 2014.
55. Riitters, K.H.; Wickham, J.D.; Wade, T.G. An indicator of forest dynamics using a shifting landscape mosaic. *Ecol. Indic.* **2009**, *9*, 107–117.
56. USCB (United States Census Bureau). *TIGER/Line Shapefiles*; U.S. Department of Commerce, U.S. Census Bureau, Geography Division, Spatial Data Collection and Products Branch: Washington, DC, USA, 2016. http://www.census.gov/geo/maps-data/data/tiger.html.